Université Sorbonne Paris Cité

Thèse de Doctorat

de l'Université Sorbonne Paris Cité

Préparée à l'Université Paris Diderot

Ecole Doctorale BioSPC (562)

Institut Pasteur / Unité des Bactéries Pathogènes à Gram-positif

Biosynthesis, role(s) and regulation of the PI-2b pilus in the hypervirulent ST-17 clone of *Streptococcus agalactiae*

Par Noémi Réka SZILI

Thèse de Doctorat de Microbiologie

Dirigée par Dr. Shaynoor Dramsi

Présentée et soutenue publiquement à Paris le 7 juin 2017

Président du jury : Prof. Isabelle Martin-Verstraete / Chercheur-Enseignant / Université Paris 7
Rapporteur 1 : Dr. Hélène Bierne / Directeur de Recherche / Institut Micalis (INRA), Jouy-en-Josas
Rapporteur 2 : Dr. Michel-Yves Mistou / Chef de Département / ANSES, Maison-Alfort
Examinateur : Dr. Anne-Marie di Guilmi / Directeur de Recherche / CEA Saclay
Directrice de thèse : Dr. Shaynoor Dramsi / Directeur de recherche / Institut Pasteur

[I have approximate knowledge
of many things, human…
- The Demon Cat of
Approximate Knowledge
of Many Things
*(Adventure Time)**]***

Acknowledgements

Completing these years as a PhD student at the Pasteur Institute was a journey on its own, one that shaped me as a person, as a scientist, and will definitely continue to shape my future. I am extremely grateful for this opportunity, especially to my thesis advisor, Shaynoor Dramsi, who chose me for this project. She helped me recognize my best skills, encouraged me to further develop these strengths, and let me discover my own path, even if following it might lead me away from the bench, for which I am wholeheartedly thankful. From our work together, I not only grew as a scientist, but also learned a lot about myself, and her scientific excellence, passion, resilience and quick wit will never cease to inspire me.

I would also like to thank Patrick Trieu-Cuot for welcoming me to his team, and his contribution to my work. I would like to express my very special appreciation and all my thanks to Bruno Périchon who was working with me and guiding me right from my first day in the unit. In addition to his invaluable work on this project, he was also like a family member in Paris, this work would not have been completed without him. I would also like to thank Élise Caliot, who provided outstanding emotional support and very fun company. She also guided me through the first steps necessary for my future career-development and encouraged me to follow my instincts. I am similarly grateful to Laurence du Merle, for tirelessly helping with experiments and always having something kind to say, her company in the lab never failed to cheer me up. Speaking of cheer and fun, I also thank Myriam, Dona, Laura, Constance, Mariana, Michel, Isabelle and Marco for the extraordinary atmosphere both in the lab and in the office. I am going to miss our talks in the coffee corner, the lunches and the Friday afternoons.

I also express my gratitude to my tutor, Nadia Izadi-Pruneyre, Christine Dugast and the entire Pasteur karate team, Monica Sala, Evelyne Turlin, Agnes Ullmann and the late Cécile Wandersmann for their friendship, encouragement and help. I would also like to thank all the members of the BBPG+ unit for the support and collaboration.

A special acknowledgement to the funding that made my work possible: This project was funded by the DIM Malinf program for 3 years, and the last months of my PhD were funded by the Department of Microbiology of Pasteur Institute.

Last but not least, I am forever grateful to my family, especially my sister any my mom, and my friends all over the world for encouraging me, cheering me along the way, and helping me believe in myself. Special thanks to my Paris squad: Nóra, Nóra, Florent, Sam, Balázs, Panni, Andrew, Altaïr, Robert, Robi, Amelia, Phillip, Fanni, and all the others, you guys were amazing, and I already miss you.

Table of Contents

List of figures and tables

Abstract

Streptococcus agalactiae (also known as Group B Streptococcus, GBS) is an opportunistic Gram-positive pathogen responsible for severe invasive infections, especially in neonates. GBS strains belonging to the ST-17 sequence type are responsible for 80% of late-onset neonatal meningitis. Genomic comparison of ST-17 strains to non-ST-17 GBS isolates revealed a few surface proteins that are characteristic of ST-17 clone, such as HvgA and Srr2, which contribute to colonization and dissemination. Similarly, the PI-2b type pilus is conserved in ST-17 strains and the main goal of this PhD project was to decipher the role of this pilus in the physiopathology of ST-17 strains.

In the first part of this work, we compared the expression of the PI-2b pilus in our ST-17 representative strain BM110, and a non-ST-17 human clinical isolate, A909. We showed that PI-2b expression, although variable, was lower in ST-17 isolates as compared to non ST-17 isolates. In the ST-17 representative strain BM110, we demonstrated that the lower expression was due to the presence of a 43-base pair (bp) hairpin-like structure in the upstream region of PI-2b, preventing read-through transcription from upstream antigen B (AgB) operon. Furthermore, gene reporter assays to characterize the Ppi-2b promoter region revealed the requirement of an extended 5' region and of GBS-specific regulatory factors to drive PI-2b transcription. PI-2b transcription was shown to be maximal at 37°C. Collectively our results suggest a complex regulation of PI-2b expression in ST-17 clinical isolates, that may confer a selective advantage in the human host either by reducing host immune responses and/or increasing their dissemination potential.

In the second part of this work, we sought to investigate the role of the putative adhesin AP1-2b, and the two accessory genes *lep* and *orf* in the biosynthesis of PI-2b pilus. We showed that both *orf* and *lep* are important for PI-2b expression. Our results suggest that Lep is a functional signal peptidase involved in the optimal processing of the major PI-2b pilin. The role of *orf* remains to be uncovered.

Résumé

Streptococcus agalactiae (Streptocoque du Groupe B : SGB) est une bactérie à Gram-positif pathogène opportuniste qui cause des infections invasives surtout chez les nouveaux nés (septicémie et méningite). Les études épidémiologiques ont montré que les SGB appartenant au complexe clonal 17 (ST-17), sont responsables de la vaste majorité des cas de méningites tardives. L'analyse d'une dizaine de génomes de SGB a permis de mettre en évidence des gènes spécifiques comme les adhésines Srr2 et HvgA qui contribuent à l'adhérence et à la colonisation des tissus chez l'hôte. On y retrouve également un pilus spécifique appelé PI-2b, qui constitua le sujet de notre étude. Ce locus PI-2b est présent et conservé dans toutes les souches ST-17 étudiées jusqu'à présent mais également dans quelques souches humaines non-ST17 comme la souche A909 ainsi que dans les souches bovines. L'objectif de ma thèse a été d'étudier l'expression, la régulation et le rôle potentiel de ce pilus dans la physiopathologie des souches de type ST-17.

Dans la première partie de ce travail, nous avons comparé l'expression du pilus PI-2b dans les souches ST-17 (souche représentative BM110) versus non-ST17 (isolat clinique A909). Nous avons montré que l'expression du locus PI-2b, bien que variable au sein du complexe ST-17, est plus faible que dans les souches non-ST17. Dans la souche représentative du ST-17, BM110, nous avons montré que la faible expression du gène *spb1* qui code pour la piline majeure du pilus Pi-2b, est dûe à la présence d'une séquence de 43 paires de base en amont du locus PI-2b, formant une structure de type « tige-boucle », qui empêche la transcription provenant du locus en amont codant pour l'antigène B (AgB). Grâce à des expériences de fusion transcriptionnelle du promoteur Ppi2b avec des gènes rapporteurs comme la GFP et lacZ, nous avons montré qu'une région étendue en 5', ainsi que des facteurs spécifiques de GBS sont requis pour une expression maximale du locus PI-2b. Le niveau maximal de transcription de pilus PI-2b est observé à 37°C. L'ensemble de nos résultats suggère une régulation complexe de l'expression du locus PI-2b, dont l'expression plus faible dans les souches ST-17, pourrait conférer aux bactéries la capacité d'échapper au système immunitaire de l'hôte et de disséminer plus efficacement.

La deuxième partie de ce travail a porté sur la caractérisation des gènes codant pour

l'adhésine putative AP1-2b et les deux gènes additionnels de ce locus, *orf* et *lep*, dans la biosynthèse du pilus. On a mis en évidence l'importance de *orf* et *lep* dans l'expression de PI-2b. Nos résultats suggèrent fortement un rôle de signal peptidase pour Lep dans la polymérisation optimale du pilus PI-2b.

Résumé substantiel

Streptococcus agalactiae (Streptocoque du Groupe B : SGB) est une bactérie commensale de la flore intestinale humaine présente chez 30-40% des individus sains. Elle est également considérée comme une bactérie pathogène opportuniste responsable d'infections invasives sévères chez les nouveau-nés (Lancefield, 1935; Congdon, 1935) et plus récemment chez les adultes âgés ou immunodéprimés (Ballard *et al.*, 2016). Ces infections invasives se manifestent sous forme de pneumonie, septicémie ou méningite, qui si elles ne sont pas traitées sont mortelles. Historiquement, SGB a été décrite tout d'abord comme un pathogène vétérinaire en particulier chez les bovins où elle est responsable de mastites ce qui lui a valu son nom « *agalactiae* » (Nocard et Mollereau, 1887). Les infections à SGB posent toujours un problème considérable pour l'industrie laitière (Jorgensen *et al.*, 2016). Depuis quelques années, SGB représente un pathogène majeur chez le poisson avec un nombre d'infections croissant dans les pisciculturesà *Tilapia*, majoritairement en Asie (Dangwetngam *et al.*, 2016).

Les souches de SGB appartenant au complexe clonal CC17 (ou séquence-type ST-17) a été désigné comme « hypervirulent » car il est responsable de plus de 80% des cas d'infections tardives chez le nouveau-né, dont la manifestation clinique principale est la méningite bactérienne (Davies *et al.*, 2004; Tazi *et al.*, 2010). Grace à l'introduction d'un protocole de dépistage du portage de la bactérie à la fin de la grossesse et de l'administration d'une prophylaxie antibactérienne aux femmes SGB-positives au moment de l'accouchement, la fréquence des infections précoces au SGB, due majoritairement aux souches de sérotype Ia, II, III et V, a fortement diminuée. Cependant, le nombre des cas de méningites tardives, qui se manifeste entre la 1ère et la 12eme semaine après la naissance, n'a pas diminué (Six *et al.*, 2015). Une meilleure compréhension du processus infectieux de SGB ST-17 est donc essentielle pour la prévention des infections tardives.

Parmi les gènes spécifiques au complexe clonal 17, en plus des adhésines Srr2 et HvgA, il y a le pilus spécifique appelé PI-2b qui a été identifié (Brochet *et al.*, 2006; Tazi *et al.*, 2010 ; Six *et al.*, 2015). A ce jour, trois îlots génomiques codant pour des pili ont été décrit chez SGB. Le locus PI-1 se trouve sur un élément mobile aussi appelé « ilôt de pathogénicité »,

alors que le locus PI-2 se situe dans le génome et existe sous deux versions différentes, PI-2a ou PI-2b selon les souches étudiées (Rosini *et al.*, 2006). L'organisation génomique de ces trois îlots est largement similaire aux autres loci hétérotrimériques codant pour des pili chez *Corynebacterium diphtheriae* (Ton-That and Schneewind, 2004; Danne et Dramsi, 2012). Les souches SGB contemporaines contiennent au moins un pilus dont le plus répandu est le pilus PI-2a, retrouvé chez 79% des isolats cliniques, suivi par PI-1 présent chez 72% des souches, tandis que PI-2b est considérablement plus rare, trouvé chez seulement 21% des souches.

Le locus PI-2b est caractéristique et conservé dans toutes les souches appartenant au complexe clonal CC17 étudiées jusqu'à présent mais également dans quelques souches humaines non-ST-17 comme la souche A909 ainsi que dans les souches bovines (Teatero *et al.*, 2016).

Les pili des bactéries à Gram-positif sont des polymères protéiques covalents formés d'une à trois sous-unités (pilines) qui sont ensuite ancrés de manière covalente à la paroi bactérienne. La polymérisation des sous-unités pilines ainsi que l'ancrage à la paroi est médié par des enzymes appelées « sortases » (Dramsi et al., 2005). Les pili de SGB sont tous composés de 3 sous-unités, dont une piline majeure formant le squelette du polymère, une piline à l'extrémité distale qui joue souvent le rôle d'adhésine et une piline d'ancrage. Le pilus PI-2b est donc composé de 3 sous-unités, le piline majeure appelée Spb1 (codé par le gène *san1518*), qui forme le polymère du pilus, la protéine d'ancrage, AP2-2b (San1516), et une adhésine putative, AP1-2b (San1519) (Rosini *et al.*, 2006). Le pilus est assemblé par la sortase SrtC1, et ancré à la paroi par la deuxième sortase du locus, Srt2 (Lazzarin *et al.*, 2015).

Dans les cas de certains pili, des gènes additionnels, souvent des régulateurs putatifs, font également partie du locus. Ape1, le régulateur de type AraC se trouve juste en amont du locus PI-1 de SGB, et participe à sa régulation (Jiang et al., 2012). De plus, il a été montré que Ape1 est sous le contrôle direct du système à deux composants, CovR-CovS, le régulateur majeur de virulence chez SGB. CovR agit comme répresseur du pilus PI-1 en se fixant à la région promotrice du pilus PI-1 (Jiang *et al.*, 2012). D'une manière similaire, la transcription du pilus PI-2a est activée par la produit du gène RogB, localisé immédiatement en amont de l'opéron PI-2a, un régulateur de la famille RALP (RofA-like regulator) (Dramsi *et al.*, 2006). L'expression de RogB est sous contrôle de Rga, une autre protéine RALP, qui règule

fortement l'expression d'autres adhésines bactériennes comme Srr1 (Dramsi *et al.*, 2012). Dans le locus PI-2b, en revanche, aucun gène régulateur putatif n'a été identifié en amont ou en aval du locus. Dans le locus PI-2b, on trouve deux gènes additionnels, *orf* et *lep* situés en amont du premier gène de piline, *ap1-2b* et dont le rôle n'est pas connu.

De nombreux pili des bactéries à Gram-positif sont impliqués dans l'adhésion aux cellules de l'hôte, ou aux protéines de la matrice extracellulaire, et participent à la formation de biofilm, et parfois à l'évasion de la réponse immunitaire. Chez SGB, PI-2a participe à l'adhésion aux cellules épithéliales du poumon humain A549 par la sous-unité AP1-2a, une adhésine contenant une domaine von Willebrand facteur Type A (VWA) (Dramsi *et al.*, 2006; Konto-Ghiroghi *et al.*, 2009). PI-2a est également important pour la formation des biofilms, où l'adhésine et la piline jouent un rôle important (Konto-Ghiorghi *et al.*, 2009; Rinaudo *et al.*, 2010).

Dans la première partie de ce travail, nous avons comparé l'expression du pilus PI-2b dans les souches ST-17 et non-ST-17. Nous avons montré que l'expression du locus PI-2b, bien que variable au sein du complexe ST-17, est plus faible que dans les souches non-ST-17. La souche BM110, un isolat clinique humain a été choisie comme souche représentative du groupe ST-17, en comparaison avec la souche A909, non-ST-17 isolée d'une infection chez l'homme et dont la séquence génomique complète est disponible montrant qu'elle possède le locus PI-2b.

Nous avons montré que l'expression du gène *spb1* codant pour la piline majeure, est 3-5 fois plus faible dans la souche BM110 comparée à la souche A909. Ceci a été quantifié par qRT-PCR ainsi que par détection avec des anticorps spécifiques de la protéine Spb1 à la surface bactérienne par dot-blot et cytométrie en flux. La comparaison génomique du locus PI-2b dans les 2 souches A909 et BM110, et plus particulièrement au niveau de la région intergénique en amont du locus PI-2b a révélé l'existence d'une séquence additionnelle de 43 paires de base (pb) dans la souche BM110, formant une structure de type « tige-boucle », qui est entièrement absente dans la souche A909. Cette même séquence est retrouvée dans le génome de 26 souches ST-17 de notre collection, et absente des 7 souches non-ST-17s dont la séquence est disponible. Suite à la délétion de cette séquence de 43 pb dans la souche BM110, le niveau de transcription des gènes de l'opéron PI-2b (*orf*, *lep*, *bp-2b*, *ap1-2b*) a augmenté d'un facteur 3-4. Dans ce mutant BM110*Δ43*, ainsi que dans la souche A909,

un produit de transcription englobant *orf* le premier gène du pilus PI-2b et le gène *sak1445* le dernier gène du locus AgB en amont, a été mis en évidence par RT-PCR. Ce produit n'a pas été détecté dans la souche BM110, ce qui suggère que l'élément de 43 pb formant une tige boucle empêche la transcription de PI-2b dit « read-through », provenant du locus AgB, et cet effet terminateur peut partiellement expliquer la différence d'expression du pilus PI-2b observée entre A909 et BM110.

Le site d'initiation de la transcription (TSS) du pilus PI-2b a été identifié 37 bp en amont du premier gène *orf*, par RNA-seq et extension d'amorce. Une boîte -10 canonique étendue et une séquence -35 semi-canonique correspondant à ce site ont été également identifiés. La fonctionnalité de la région promotrice a été étudiée en réalisant des fusions transcriptionnelles : des fragments de la région intergénique PI-2b de A909, contenant le TSS, ont été clonés dans le plasmide pTCV-GFP, en amont du gène rapporteur *gfp* et exprimés dans la souche non-PI-2b de SGB, NEM316 (ST-23). Nous avons montré que le site d'initiation de la transcription est fonctionnel, mais aussi qu'une région étendue en 5', ainsi que des facteurs spécifiques de SGB sont requis pour une expression maximale du locus PI-2b. Entre autres, la température affecte l'expression de *gfp* depuis le TSS de PI-2b, qui a été maximale à 37°C.

En effet lorsque les fusions transcriptionnelles ont été introduites dans une souche de *Lactococcus lactis*, une espèce hétérologue proche considérée comme non-pathogène. Chez *L. lactis*, aucune expression de GFP ou transcrit du gène *gfp* a pu être détecté, donc des facteurs spécifiques de SGB sont nécessaires pour la transcription du pilus PI-2b.

Nous avons également étudié l'impact potentiel du système à deux composants CovR/S dans l'expression du locus PI-2b comme cela a été montré pour le pilus PI-1 (Jiang et al., 2012). L'expression de la GFP s'est révélée 2-3X plus élevée dans un mutant *ΔcovR* de NEM316 par rapport à la souche sauvage. Une différence similaire a été démontré pour la transcription du gène *orf* et *spb1* entre le mutant BM110*ΔcovR* et BM110. Cependant, nous n'avons pas observé de différence d'expression du locus PI-2b dans les mutants ponctuels CovRD53A de NEM316 et BM110, où le site de phosphorylation de CovR est muté ce qui résulte en une activation constitutive de CovR. En revanche, dans le cas de PI-1 qui est sous la régulation directe de CovR, la mutation D53A de la protéine a un effet délétère pour l'expression du pilus. Dans le cas de PI-2b, CovR a donc plutôt un effet indirect dans

l'expression, indépendant de la phosphorylation de CovR.

L'ensemble de nos résultats suggère une régulation complexe de l'expression du locus PI-2b, dont l'expression plus faible dans les souches ST17, pourrait conférer aux bactéries la capacité d'échapper au système immunitaire de l'hôte et de disséminer plus efficacement.

La deuxième partie de ce travail a porté sur la caractérisation des gènes codant pour l'adhésine putative, AP1-2b, et les deux gènes additionnels de ce locus, *orf* et *lep*, dans la biosynthèse du pilus. L'analyse bioinformatique de *lep* a montré que la protéine Lep appartient à la famille des peptidases signal type S26A et contient le site catalytique décrit pour ce groupe (Young *et al.*, 2014). En tant que protéines de surface, la majorité des pilines possèdent un peptide signal, qui doit être clivé après sécrétion de la protéine à travers la membrane. Cependant, la présence d'une séquence signal reconnue par signal peptidase de Type I, avec un motif AxA (Auclair, Bahnu et Kendall, 2012), a également prédit *in silico* pour la piline majeure Spb1 et l'adhésine AP1-2b.

Nous avons construit un mutant délété pour le gène *lep* dans la souche A909 et BM110. Nous avons montré que l'expression ainsi que la polymérisation du pilus PI-2b était réduit de 2-3 fois dans les mutants délétés pour *lep*. La surexpression de la piline majeure, Spb1, dans ce mutant, indique que, en absence de *lep*, l'efficacité de la polymérisation des pili est diminuée. La complémentation du mutant *lep*, en surexprimant Lep à partir d'une construction plasmidique, a restauré la piliation à un niveau comparable à la souche sauvage correspondante indiquant que Lep code potentiellement pour une signal-peptidase impliquée dans la maturation de la piline Spb1 et contribuant à la biosynthèse optimale du polymère.

La délétion du gène *orf* a également été obtenue mais les phénotypes observés n'ont pas pu être complémentés par la surexpression du gène *orf* en trans. Malgré nos efforts, aucune production d'une protéine Orf n'a été détectée. Néanmoins, cette organisation génomique, où une petite *orf* chevauchant un autre gène de type *lep*, est conservée dans d'autres opérons pili chez les Streptocoques, et donc le rôle de ces gènes reste à étudier.

Introduction

During the past decade, several bacterial species that are normal constituents of the human microbial flora have been shown to act as opportunistic pathogens under certain conditions. These infections usually concern immune-compromised individuals, the elderly and the newborn, and are also often associated with other diseases, such as different types of cancer. The exact factors, however, that allow commensal bacteria to turn pathogenic are so far poorly defined. Prevention strategies and risk-assessment guidelines are therefore difficult to establish, especially because the role of these bacteria as flora-constituents in healthy persons is not well understood. This is the case for *Streptococcus agalactiae* (also known as GBS for Group B *Streptococcus*), which is part of the intestinal and vaginal flora of healthy humans, but is also an emerging cause of infections in elderly persons and a well-known neonatal pathogen responsible for severe invasive infections and mortality worldwide.

A. *Streptococcus agalactiae* (or Group B Streptococci)

Streptococcus agalactiae is the species designation for streptococci belonging to Lancefield group B and thus is also known as Group B *Streptococcus* (or GBS) (Lancefield, 1933). This species was first isolated from a bovine mastitis infection in 1887 (Nocard and Mollereau, 1887). Rebecca Lancefield established a serological classification system for *Streptococci* in 1933, which identifies *Streptococcus agalactiae* as the only Group B Streptococcus species (Lancefield, 1933). She was also the first to isolate *S. agalactiae* from a human vaginal sample in 1935 (Lancefield, 1935) and, the same year, this species was first associated to neonatal disease and described as a human pathogen (Congdon, 1935; Fry, 1938).

General Features

S. agalactiae is a Gram-positive species with low GC% belonging to the *Streptococcaceae* family that includes the *Streptococcus*, *Enterococcus* and *Lactococcus* genera. These genera share several common characteristics: they are non-sporulating Gram-positive cocci with aerobic or facultative anaerobic metabolism, and are all catalase-negative. They are often referred to as lactic acid bacteria, as during their metabolism they produce lactic acid from glucose fermentation, leading to the acidification of their growth medium (Brown and Collins, 1977).

GBS are usually cultivated in rich media such as Todd-Hewitt, M17, Columbia or Brain Heart Infusion. *S. agalactiae* form chains of 4 to 15 coccoid (or sometimes ovoid) cells, each of 0.6-1.2µm diameter, the chains are immobile and sometimes capsulated. Colonies on horse-blood agar are gray-white, 3-4mm diameter, flat and mucoid, exhibiting a narrow β-hemolysis, sometimes visible only when colonies are removed from the agar. GBS also produces an orange pigment during anaerobic growth on Granada medium (Rosa-Fraile *et al.*, 1996) and a pore-forming toxin named CAMP (Christie, Atkins and Munch-Petersen, 1944; Landwehr-Kenzel and Henneke, 2014). The CAMP factor is an extracellular protein that, in the presence of the β toxin of *Staphylococcus aureus*, produces synergistic hemolysis when grown on sheep blood agar. GBS can be differentiated from other streptococci by a combination of tests including specific type B-antigen testing, CAMP, bile esculin reaction, and bacitracin sensitivity.

As Gram-positive bacteria in general, *S. agalactiae* is naturally resistant to polymixins and first-generation quinolone-type antibiotics (intrinsic resistance). Due to mobile genetic elements, *S. agalactiae* can also acquire extrinsic resistance to several antibiotics. For example, 80% of GBS clinical isolates are tetracycline-resistant, mainly due to the *tet(M)* gene carried on the conjugative transposon Tn*916* (94% prevalence in France) (Poyart *et al*., 2000), and more rarely due to other *tet* genes (Da Cunha *et al.*, 2014). 20% of GBS isolates are resistant to macrolides and, in these strains, the ribosomal target of the antibiotics is modified by a methylase, usually coded by either *erm(B)* or the *erm(TR*) genes. In a few rare cases, macrolide resistance is conferred by the *mef(A)* gene that codes for an active efflux system (De Mouy *et al*., 2001, Poyart *et al.*, 2000). About 10% of *S. agalactiae* strains are resistant to kanamycin and neomycin, and less than 5% of the isolates are resistant to streptomycin (Poyart *et al.*, 2000).

The group B antigen is a conserved complex multi-antennary carbohydrate structure composed of 4 oligosaccharides (rhamnose, galactose, N-acetylglucosamine and glucitol), which is covalently linked to the cell wall (Pritchard *et al.*, 1984). It is not considered to be highly immunogenic because it is partly masked by the capsule, another carbohydrate polymer linked to the cell wall.

Streptococcus agalactiae are classified into ten different serotypes (Ia, Ib, II-IX) based on differences in their capsular polysaccharide antigens that are designated as type antigens. The first three serotype-antigens (serotype I, II and III) were also identified by Rebecca Lancefield (Lancefield and Freimer, 1966). The serotype I has been subsequently divided into 2 subtypes, Ia and Ib, and 6 other serotypes have been identified so far, based on different arrangements of four monosaccharides. Glucose, galactose, and N-acetylneuraminic acid (often referred to as sialic acid) are found in all known capsule serotypes, N-

acetylglucosamine is also present, except in the VI and VIII serotypes. In contrast, serotype VIII contains a rhamnose repeating unit (Cieslewicz *et al.*, 2005).

Each antigen has a backbone formed by repeating units of two (Ia, Ib), four (II), or three (III, IV, V, VII, VIII) monosaccharides, mostly glucose and galactose, to which one or two side chains are linked. Sialic acid is the exclusive terminal side chain sugar, linked to a glucose residue by an α-2-3 bond (Chaffin *et al*., 2000), except for the type II polysaccharide, which contains a terminal galactose. This structural diversity is explained by the genetic diversity of the polysaccharide capsule biosynthesis cluster associated to different serotypes (Cieslewicz *et al.*, 2005). Interestingly, the asialylated type III polysaccharide is immunologically identical to *Streptococcus pneumoniae* type-14 group antigen (Lindberg, Lönngren and Powell, 1977).

The GBS capsule is an immunogenic molecule and capsular serotyping is the method of choice in epidemiological studies. Most human GBS isolates are capsulated and the sialylated capsular polysaccharide is an important virulence factor that prevents activation of the alternative complement pathway by inhibiting complement deposition, and therefore promotes immune escape. The capsule type III is the richest in sialic acid, therefore, it confers a high virulence potential to these strains: 60-80% of severe neonatal infections are due to serotype III GBS strains (Schuchat, 1999; Weisner *et al.*, 2004).

Experimental evidences suggest that the group B carbohydrate and the CPS are linked independently to cell wall peptidoglycan by covalent bonds (Deng *et al.*, 2000). Immuno-electron images revealed the presence of abundant CPS on Lancefield prototype strains Ia, II, and III, whereas less dense capsular polysaccharides were found on type Ib strains (Kasper and Baker, 1979). Capsule expression in GBS is regulated by growth stage (Paoletti, Ross, Johnson, 1996; Ross, Madoff, Paoletti 1999).

Recently, new studies using atomic force microscopy (AFM) and transmission electron microscopy (TEM) aimed to analyze the relative distribution of CPS and Group B antigen on the bacterial surface. These nanoscale experiments showed that, in WT bacteria, a very thin group B antigen layer (the “pellicle”) covers the cell wall peptidoglycan which is overlaid by an external 15–45 nm thick layer of capsule polysaccharide. The capsule was over-expressed in a mutant unable to synthesize antigen B (*ΔgbcO*), thus expression of these two surface glycopolymers is probably coordinated in GBS. AFM topographic imaging revealed the exposure of cell wall peptidoglycan in a non-capsulated mutant *(ΔcpsE*) and its organization as 25 nm wide bands running parallel to the septum, while similar disruption of AgB production had no similar effect. The capsule therefore constitutes an effective protective barrier for the peptidoglycan layer from environmental exposure (Beaussart *et al.*, 2014). Moreover, unlike mutants unable to synthesize group B antigen, null-mutants of capsule production were not affected in growth, morphology or peptidoglycan synthesis (Rubens *et*

al., 1987; Rubens *et al.*, 1993). GBS strains, designated as non-typeable because they do not express one of the 10 CPS-specific antigens are characterized by molecular methods such as PCR. Other DNA-based methods are used for analyzing the diversity of GBS, such as PFGE, random amplified polymorphism, *Eco*RI ribotyping, or the most recently developed Multi-Locus Sequence Typing (MLST). The MLST method is based on the sequence analysis of 7 housekeeping genes of *S. agalactiae* (Jones *et al*., 2003).

Nowadays, beyond capsular serotyping, MLST is the most frequently used method for molecular characterization of the strains. The seven housekeeping genes used in this method are *adhP*, *atr*, *sdhA*, *pheS*, *glnA*, *glcK* and *tkt*, showing about 5-11 different allelic variants per gene, the combination of which is gives a distinct pattern for different MLST types. Initially 29 STs have been identified, however, since then, due to the increased number of GBS draft genomes, this number is constantly growing. Interestingly, strains belonging to the same ST were sometimes found to belong to different capsular serotypes (Jones *et al.*, 2003), which suggests that recombination events can occur at the capsular locus, causing serotype-switching, which was later confirmed on the molecular level (Martins, Melo-Cristino and Ramirez, 2010).

The first complete genome of *S. agalactiae* strain NEM316 was published in 2002 (Glaser *et al.*, 2002). Analysis of two fully sequenced genomes (A909 and 2603V/R) together with 5 draft genomes (COH1, 515, CJB111, H36B and 18RS21) and revealed a stable core genome of about 1800 genes with 11-14 interspersed mobile islands (Tettelin *et al.*, 2005). Some of these islands harbor several known virulence genes, including the cytolysin (*cyl*) operon and the *lmb-scpB* locus indicating that these regions have an important role in strain diversity and pathogenicity. Very recently, using a high coverage transposon-based random mutagenesis approach, and collating results from three different libraries, Hooven and colleagues defined the essential genome of GBS strain A909 (serotype Ia, ST-7). 13.5 % of genes were shown to be essential and 1.2 % to be critical, which have a role in mainly fundamental cellular housekeeping functions, such as acyl-tRNA biosynthesis, nucleotide metabolism, and glycolysis (Hooven *et al.*, 2016).

As mentioned above, capsular serotyping is not necessarily a precise and reliable approach to address the genetic diversity of *S. agalactiae*. Tettelin and colleagues further indicate that due to frequent horizontal gene transfer events in GBS, the seven genes arbitrarily selected for MLST might be misleading for revealing the real genetic diversity and phylogeny, which can only be addressed by whole-genome sequencing (Tettelin *et al.*, 2005; Springman *et al.*, 2009; Sorensen *et al.*, 2010; Salloum *et al.*, 2011). For the time being, however, serotyping and MLST typing of GBS strains is an invaluable tool for understanding streptococcal colonization and virulence, especially in the case of such highly clonal ST-s such as ST-17.

ST-17 and CC17 of GBS

The population of invasive GBS isolates shows high clonality as isolates from infections in both adults and newborns mainly belong to five major genetic lineages, ST-19, ST-17, ST-1, ST-23, and ST-9 complexes (of serotype Ia, Ib, and II to V) as determined by MLST in several studies (Luan *et al.*, 2005). In all such studies, the ST-17 clone, almost uniquely containing serotype III isolates, have been shown to be responsible for the vast majority (75-90%) of late-onset neonatal meningitis cases (see corresponding chapter) and what therefore designated hypervirulent (Davies *et al.*, 2004; Luan *et al.*, 2005; Jones *et al.*, 2006; Manning *et al.*, 2009).

A broader phylogenetic analysis using sequences of 15 housekeeping genes and virulence-associated patterns later revealed a genetically heterogenous GBS core-population, with the occasional emergence of virulent lineages as a result of recombination events that affect bigger segments of the genome (Sorensen *et al.*, 2010). The ST-17 clone, however, was strongly identical with the corresponding CC17 clonal complex, confirming its high homogeneity (Brochet *et al.*, 2006; Springman *et al.*, 2009; Sorensen *et al.*, 2010). The results also suggested that relying solely on MLST for typing GBS can be misleading, and also revealed that although a previous study suggested otherwise (Bisharat *et al.*, 2004), CC17 (or ST-17) have not emerged from bovine-specific ancestry (Sorensen *et al.*, 2010).

Certain homogenous virulent lineages, such as CC17 and two distinct clusters belonging to CC23 are exclusively adapted to their specific hosts, humans and cattle, respectively. Homogeneity of the CC17 clone could be a result, at least partially of its relatively recent emergence, as the number of GBS-associated neonatal infections rose drastically in the 1960 (Sorensen *et al.*, 2010). By a combined phylogenetic and antibiotic resistance marker analysis, it has been revealed that among all six human-invasive CCs, CC17 has the lowest rate of recombination, suggesting that it represents the youngest clone among these relatively recent clones. The emergence of GBS-associated infections coincided with the widespread use of tetracycline in medical practice, that led to the replacement of the human-colonizing GBS population by a selection of tetracycline resistant (TcR) clones. This suggests that the currently human invasive GBS clones were selected through the extensive use of tetracycline, which is also supported by the clinical observation that most human isolates are resistant to this compound (Da Cunha *et al.*, 2014).

As later detailed in this study, the CC17 clone codes for certain specific or highly characteristic surface proteins. Beyond these, the chromosomal structure also carries certain distinctive features, most importantly a significantly low presence of spacers compared to other CCs (Lier *et al.*, 2015). Although the recombination-related variation of the CC17 core genome is very low, certain pathogenicity islands are systematically acquired or lost. One of

such events is the replacement of the Pilus Island 1 with antimicrobial resistance determinants (Teatero *et al.*, 2016).

As rapid detection of CC17 can be crucial in prevention of neonatal infections, a specific PCR-based screening protocol has been developed. The rapid-test relies on a CC17-specific allele of the *gbs2018* gene, coding for a surface adhesin (Brochet *et al.*, 2009), and provides reliable and accurate identification of CC17 bacteria in clinical samples (Lamy *et al.*, 2006). The importance of this detection method is also highlighted by the growing number of studies reporting capsular switching of GBS strains (Luan *et al.*, 2005). While these events were relatively rare for CC17 that consistently carried serotype III surface antigens, recently serotype IV CC17 isolates have been described (Bellais *et al.*, 2012).

Asymptomatic carriage in humans

Streptococcus agalactiae is found as part of the intestinal and genital microbiota of 15 to 35% of healthy adults. The primary reservoir of colonization is the intestine in both men and women. GBS has been isolated from the small intestine of adults on several occasions (Barnham, 1983) and found to be associated with infections following surgery of the upper or lower intestinal tracts (Easmon *et al.*, 1986). Human-colonizing strains were shown to belong mostly to serotypes III (24%), Ia (21%), V (18%), and Ib (17%) in Europe and North America. Later, by MLST typing, asymptomatic strains have been shown to belong mainly to the sequence types ST-1 and ST-19, and partially to ST-23, although this latter clonal complex also contains invasive isolates (Jones *et al.*, 2003). Serotype III isolates, and especially ST-17 clones are more often related to invasive disease than to asymptomatic carriage (Davies *et al.*, 2004).

Vaginal colonization in adult women is periodical in time, a longitudinal cohort study of non-pregnant women in the 1970s showed that 50% of culture-negative women become positive carriers during follow-up at three 4-month intervals (Meyn *et al.*, 2002). Similarly, nearly 50% of pregnant women positive for vaginal colonization at delivery have had negative antenatal culture results. The predictive value of a positive prenatal culture result is highest (73%) in women with vaginal and rectal colonization and lowest (60%) in women with rectal colonization only. In agreement with these statistics, sample collection from both rectal and vaginal sites has been shown to be the most reliable screening method for GBS colonization (Dillon *et al.*, 1982). Sample collection and analysis at 1 to 5 weeks before delivery is fairly accurate for predicting colonization status at delivery (Yancey *et al.*, 1996), and is generally recommended as prevention methods.

GBS infection in neonates

Streptococcus agalactiae emerged as the leading cause of bacterial neonatal infections in Europe and North America during the 1960s via the selection and worldwide dissemination of only a few clones. This expansion was shown to be preceded by the insertion mobile elements conferring tetracycline resistance (TcR) to these GBS clones, therefore it was proposed that the common use of tetracycline from 1948 onwards led to the observed emergence of a subset of strains as main cause of GBS infections (Cunha *et al.*, 2014).

Newborns are usually considered to acquire GBS via ingestion or inhalation of the bacteria during the passage through the vaginal tract of a colonized mother or *in utero* when bacteria ascend vertically to reach the amniotic fluid. Longitudinal comparison of mother-baby pairs showed that 30 out of 35 colonized neonates were born to culture-positive carriers, while 5 to non-carriers or women whose initial culture samples were negative (Berardi *et al.*, 2013). Molecular typing analysis confirmed strain identity in all mother-infant pairs, which suggests direct maternal transmission in these cases. However, the colonization of children of non-carrier mothers raises questions about other possible transmission routes. In this study, 6 of 83 culture-positive carrier mothers had a positive milk culture (Berardi *et al.*, 2013). A comparative study of literature also suggested infected breast milk as a source of GBS (Filleron *et al.*, 2013). A comparative genomic analysis of strains from infected neonates and their mothers also proposed this route, but also suggested that it is more likely that GBS detected in breastmilk is actually a retro-infection from the oral bacteria of the newborn due to breastfeeding, and the ingestion of infected breastmilk might be responsible for LOD (Almeida *et al.*, 2015).

Infants born to colonized mothers become colonized themselves in 50-70% of the cases, and only 2-3% of the colonized newborns develop a symptomatic invasive infection. The two distinct clinical manifestations of neonatal GBS infections are the early-onset disease that occurs during the first week of life, and the late-onset disease that manifests from one week up to three months after birth.

Early onset disease (EOD)

The characteristic clinical manifestation of early-onset disease is usually respiratory troubles due to severe pneumonia, often accompanied by septicemia. Meningitis is observed in only about 10% of EOD cases (Zangwill *et al.*, 1992). EOD is either caused by an ascending *in-utero* infection after the passage of the bacteria through the amniotic membrane, especially in case of early rupture of the membranes, or more often by the inhalation and ingestion of vaginal secretions by the newborn during the passage through the birth canal.

After the implementation of a prevention protocol based on systematic screening of pregnant women and an antibiotic prophylaxis administered to the colonized mothers a few hours prior to birth, the incidence of EOD drastically dropped. While at the time of the introduction of the protocol in the early '90s, EOD cases represented 80% of all GBS infection, it decreased to 50% by the end of the 2010s, while the number of LOD cases remained unchanged. In the USA, the current frequency of EOD is 0.4 per 1000 live births.

Preterm birth was identified as the major risk factor for adverse birth outcomes and especially of EOD. More precisely, infection with vaginal GBS *in utero* was recently shown to be a significant cause of preterm birth. GBS hyaluronidase was suggested to plays a crucial role in ascending infection and preterm birth (Vornhagen *et al.*, 2016).

GBS capsular serotypes associated to EOD cases correspond to those described to be the most prevalent in vaginal colonization, with a similar proportion of strains of serotypes Ia, II, III and V in Europe and North-America (Harrison *et al.*, 1995; Blumberg *et al.*, 1996). Serotype VIII GBS is mainly isolated from pregnant women in Japan (36%) (Lachenauer *et al.*, 1999), while infections due to this serotype remain sporadic elsewhere (Paoletti *et al.*, 1999; Ekelund *et al.*, 2003).

Late onset disease (LOD)

Late onset GBS infections occur from one week to up to three months after birth, with approximately the same frequency, although some researchers suggest a distinction between late-onset and very late-onset infections (Cantey *et al.*, 2014). The LOD syndrome is different from early-onset streptococcal infections in its clinical manifestations: the two most common manifestations of LOD are meningitis and bacteremia without a focus. About 10% of the late onset infections are fatal and up to 50% of the neonatal meningitis cases lead to severe neurological sequelae, including mental retardation, deafness, seizures, speech and language delay, etc. (Levent *et al.*, 2010). While the antibiotic prophylaxis protocol significantly decreased the occurrence of EOD cases, the prevalence of LOD infections remained stable over the past 20 years, suggesting and important difference in the acquisition and/or the pathogenesis of the two syndromes.

The majority of the LOD cases are attributed to serotype III GBS infections, among which ST-17 strains are the most prevalent (Tazi *et al.*, 2012). The ST-17 complex has more precisely been associated to up to 85-90% of the late onset meningitis cases, and was therefore designated as the hypervirulent serotype of *S. agalactiae* (Davies *et al.*, 2004).

GBS infection in adults

GBS is also responsible for endocarditis and bacteremia in non-pregnant adults, especially in the elderly and immunocompromised individuals (Faro, 1981). The clinical manifestations vary from septicemia, meningitis and pneumonia, to less severe arthritis or urinary tract infections. 70% of these cases are attributable to GBS serotypes Ia, III, and V (Farley, 2001). Among these, serotype V recently emerged as the most common capsule type recovered from non-pregnant adults with invasive GBS disease (Blumberg *et al.*, 1996). Infections by *S. agalactiae* also occur in pregnant women and mothers *postpartum*, in these cases the infection is usually localized to the reproductive tracts and uterus, and can result in endometriosis and/or infertility if untreated (O'Higgins *et al.*, 2014). About 33% of yearly GBS infections occur in non-pregnant adults, and 15% in pregnant women (Farley *et al.*, 1993), but the frequency of invasive GBS infections among the elderly is increasing (Ballard *et al.*, 2016).

GBS infection in animals

Bovine GBS infections

GBS was first described as animal pathogen isolated from bovine mastitis (Nocard and Mollereau, 1887), and before the use of antibiotics in farming, it was responsible for up to 90% of all mastitis cases. While nowadays epidemics of bovine mastitis are rare, sporadic cases often occur. The prevalence of *S. agalactiae* in North American and North European dairy herds is still around 10% and is responsible for important losses in the dairy industry (Jorgensen *et al.*, 2016).

The presence of GBS in the udder can be asymptomatic colonization, or subclinical mastitis, but it can also provoke the painful swelling of the udder and the production of abnormal milk containing pellets and aggregates. In the most severe cases, atrophy, udder fibrosis and the complete arrest of lactation can ensue (Keefe, 1997). The bovine gastrointestinal tract (rumen) and the udder environment were both shown to be reservoirs of *S. agalactiae*. 2 different possible transmission cycles were proposed: an environmental contagion cycle via the milking machine and a cow-to-cow oro-fecal transmission cycle (Jorgensen *et al.*, 2016).

The relationship between GBS strains of human and bovine origin has been subject of debate for years. However, the use of whole genome sequencing to compare strains from bovine sources and human neonates indicates that these lineages are unrelated. Phylogenetic lineage analysis determined the existence of an "ancestral" lineage of bovine GBS, from which human and bovine strains developed through host-specific adaptation

(Bohnsack *et al.*, 2004). Consequently, substantial biochemical, serologic, and molecular differences exist between human and bovine strains (Finch and Marin, 1984), most importantly bovine strains display the unique ability to use lactose as carbon source (Richards *et al.*, 2011). Comparative virulence gene analysis of bovine and human pathogenic strains also confirmed a strong host-adaptation, supporting the idea that *S. agalactiae* isolated from humans and bovines are generally unrelated and belong to separate populations (Emaneini *et al.*, 2016).

The majority of bovine invasive GBS strains belong to serotype Ia, or are non-typable due to the loss of the capsule. Based on MLST typing, these strains belong mainly to ST-61 or ST-63 sequence type (Sorensen *et al.*, 2010; Rosinski-Chupin *et al.*, 2013).

Fish infections

From the 1980s, GBS has been described as an emerging pathogen in fish farming. One of the most affected fish species is Tilapia, but other fishes and amphibians can also be infected. Symptoms are difficult to describe in fish, infections usually manifest as sudden death on the population level, but analysis of the infected fish showed that the main sites of infection are the brain and the kidneys (Dangwetngam *et al.*, 2016).

S. agalactiae infections in fish are predominantly caused by beta-hemolytic strains of ST-7, or by non-hemolytic strains of ST-260. Comparison of the genome sequence of ST-260 and ST-23 strains derived from fish, cattle or humans revealed the presence of genomic elements that are unique to fish-infecting subpopulations of *S. agalactiae*. These genes were found mainly on mobile genetic elements, and encode proteins that potentially provide fitness in the aquatic environment (Delannoy *et al.*, 2016). Phylogenetically, these fish isolates clustered both with human and bovine strains, suggesting a diverse origin and heterogeneity of this population. Furthermore, it has been shown that certain human and bovine GBS isolates are capable of causing disease in tilapia, indicating a possible route of contamination of fish farms from human or ruminant hosts (Chen *et al.*, 2015).

B. Virulence factors of GBS

Virulence factor is a generic term that regroups all molecular components in a pathogenic bacterium that contributes to its survival in the host environment and that promotes infection. These functions include molecules necessary for increasing the efficiency of colonization in different organs, invading host cells and tissues (especially in the case of intracellular pathogens) and persisting by getting nutrients, immunosuppression and evasion of the host immune responses.

Virulence factors can be categorized using different criteria, here I chose to differentiate essentially between surface-expressed virulence factors (usually – but not exclusively – adhesins), and a handful of well characterized GBS virulence factors that I chose to present below.

The capsule (CPS)

The capsule is one of the main antigenic determinants of GBS. It is also a key virulence factor, as it interferes with the complement response, helping GBS to escape killing by immune cells (Edwards *et al.*, 1982). Serotype III GBS strains generally resist well complement-mediated phagocytic killing if specific antibodies are not produced. The serotype III highly sialylated capsule glycopolymer was shown to prevent complement (C3) deposition, reducing phagocytic killing. Uncapsulated bacteria and mutants with a capsule lacking sialic acid (Sia) bound drastically more C3 on their surface than wild type bacterial cells, where deposition occurred mainly via the alternative pathway. In contrast to the wild type bacteria, capsule-deficient mutants were efficiently killed by human leukocytes, revealing a correlation between capsule production level and survival to phagocytic killing (Marques *et al.*, 1992).

Interestingly, the position and linkage of terminal sialic acid (Sia) residues in GBS capsule polymers resemble the pattern of cell surface glycans often coating human cells. Therefore, it was proposed that bacterial sialylated molecules are mimicking the host proteins decoration to avoid recognition by the immune system. Immunoglobulin-like lectins (Siglecs) are expressed on the surface of leukocytes, and certain are responsible for binding Sia-capped glycans in the host as a part of a pathway that reduces inflammatory response (Chang *et al.*, 2014). Different GBS strains were shown to interact with these Siglecs with an efficiency that was dependent on the degree of capsulation and the sialylation state of the capsule polysaccharide (Carlin *et al.*, 2007; Chang and Nizet, 2014). As the serotype III capsule is the most sialylated variant, this GBS serotype is the most resistant to phagocytic killing in humans. A recent study confirmed that hypervirulent ST-17 strains contained more

sialylated molecules as compared to ST-1 and ST-19 isolates of serotype III (Lartigue *et al.*, 2011). Interestingly, even among ST-17 strains, differences have been shown in the transcription of the *neuD* gene, which codes for the enzyme carrying out the last step of capsule glycopolymer sialylation, according to the isolation site (brain versus vagina), indicating that controlling the capsular Sia content can actually be a means for GBS to adapt to certain host environments (Pailhories, Lemire and Lartigue, 2013).

GBS capsule was also shown to modulate immune responses in mouse dendritic cells, which are the main antigen-presenting cells to other components of the host immune system. Dendritic cells internalize both capsulated and non-capsulated GBS cells, however the capsule allows long-term intracellular bacterial survival. Internalization of GBS occurs via the clathrin-mediated endocytosis, whether or not the capsule is present, capsulated bacteria, however, are also internalized in a lipid raft-dependent pathway that could partially account for their prolonged intracellular survival. Furthermore, internalized GBS also affected pro-inflammatory chemokine, as well as pro-inflammatory and Th1 cytokine production in a manner that was dependent on the origin (bone marrow or spleen) of the dendritic cells (Lemire *et al.*, 2012a; Lemire *et al.*, 2012b). These results foreshadow the existence of an intricate interplay between GBS capsule and actors of the host immune response, mainly phagocytes, resulting in selective modulation of immune response mechanisms.

Cytotoxins

The pore-forming toxins expressed by GBS are one of the most important factors during infection, as by lysing the host cells, they promote deeper tissue colonization and systemic dissemination. Cells expressing no or low amounts of pore-forming toxins were shown to be attenuated for virulence. As mentioned in chapter 1, *Streptococcus agalactiae* codes for two known pore-forming toxins: β-hemolysin/cytolysin (β-H/C) and the CAMP factor.

β-hemolysin/cytolysin

As its name suggests, β-hemolysin/cytolysin (β-H/C) promotes invasion of both epithelial and endothelial barriers, including the gut or lung epithelium and the blood–brain barrier, inducing organ failure via promoting local inflammation and host cell death (Rajagopal, 2009). For example, GBS β-H/C was also shown to promote neuronal damage in the case of central nervous system infections, without triggering a primary inflammatory response, via inducing caspase-independent apoptosis (Reiss *et al.*, 2010). When GBS is internalized by host cells, β-H/C also contributes to crossing of the phagosomal membrane due to its pore-forming activity. The release of GBS or GBS components, such as bacterial DNA, into the host

cytosol triggers the upregulation of the type I interferon (IFN), leading to increased macrophage activation (Charrel-Dennis *et al.*, 2008).

β-haemolysin/cytolysin was first identified to be genetically encoded by the large *cyl* operon, containing all the genes involved in hemolysin expression (Spellerberg *et al.*, 1999). The β-H/C activity was subsequently shown as encoded by the *cylE* structural gene (Pritzlaff *et al.*, 2001). However, Whidbey and colleagues recently demonstrated for the first time that hemolytic and cytolytic activity of GBS is due to the ornithine rhamnolipid pigment and not due to a pore-forming protein toxin (Whidbey *et al.*, 2013). This explains why the expression of β-H/C always coincided with the presence of a red polyenic carotenoid pigment; for example, deletion of the *cylE* gene results in non-hemolytic and non-pigmented bacteria, while a hyper-hemolytic phenotype is usually accompanied by hyper-pigmentation.

While the deletion of this factor confers an important selective disadvantage to mutants in colonization and disease (Rodriguez-Granger *et al.*, 2015), naturally non-hemolytic and non-pigmented isolates of GBS are frequently associated with invasive infection (Six *et al.*, 2016). In such strains, mutations were found both in the *cyl* operon, and in the *abx1* gene, encoding a partner of the two-component global virulence regulator system, CovR/S. Therefore, while β-H/C production and hemolysis is usually associated with GBS virulence, it is not entirely indispensable in human invasive infections (Six *et al.*, 2016).

The orange pigment, also referred to as granadene as it becomes visible when GBS is cultured on specific Granada™ medium (De La Rosa *et al.*, 1983), was shown to promote resistance to reactive oxygen species due to its structure rich in unsaturated bonds. Several carotenoid pigments were shown to quench ROS in numerous known pathogenic bacteria (Rajagopal, 2009).

CAMP factor

The CAMP factor is a secreted pore-forming toxin, which was suggested to be important for GBS pathogenesis, as the first in vivo rabbit and mouse experiments showed that both purified CAMP or sub-lethal doses of GBS supplemented with CAMP were lethal. The mechanism of pore-formation was also confirmed by microscopic observations, where CAMP molecules were shown to oligomerize on the target membranes thus creating pores, and it was also revealed that this reaction requires binding to host surface receptors, as it is the case for most cytotoxins. However, a mutant deleted for the *cfb* gene encoding the CAMP cytotoxin did not exhibit reduced virulence nor alteration of any virulence-associated phenotypes. These results question the role of CAMP in virulence (Hensler *et al.*, 2008).

Given that both CAMP and β-H/C are pore-forming toxins, it is also likely that this latter may compensate for the absence of CAMP factor during systemic infection. There are a few

host niches where β-H/C activity is diminished, the most important being the human lung, where the dipalmitoyl-phosphatidyl-choline (DPPC) surfactant component is antagonistic to β-H/C. Rajagopal and colleagues therefore speculate that CAMP factor expression may exclusively be important in very specific host niches (eg. those high in DPPC), while generally GBS preferentially utilizes β-H/C (Rajagopal, 2009).

SodA

Production of reactive oxygen species (ROS) in the phagosomes of immune cells is an important mechanism of pathogen-killing by host immune defense. As mentioned earlier, the granadene carotenoid pigment is important for GBS survival in highly oxidative environments. GBS also encodes a superoxide dismutase, SodA, to ensure resistance to ROS and host immune evasion. GBS SodA belongs to the Mn^{2+} dependent superoxide dismutase family, whose main activity is converting singlet oxygen radicals or superoxide anions to more stable molecular oxygen and H_2O_2. These molecules are subsequently metabolized by catalases or peroxidases to further reduce the oxidative. SodA is therefore a key factor in enabling GBS to resist oxidative stress in the host during infection, more effectively resisting phagocytic killing by dendritic cells and macrophages (Poyart C *et al.*, 2001; Rajagopal, 2009).

DltA

One of the first lines of host defense is the production of cationic antimicrobial peptides (CAMPs, not to be confused with the CAMP factor), such as defensins and cathelicidins, by several host cells-types including epithelial cells, keratinocytes and leukocytes as a part of the innate immune response. Since these molecules display a strong cationic charge, they can interact and damage the negatively charged cell wall through electrostatic interactions. It was observed that in the presence of CAMPs, bacteria resist to these molecules by incorporating D-alanine into their cell wall to reduce the overall surface charge (Weidenmaier *et al.*, 2003; McBride and Sonenshein, 2011).

The elements necessary for D-alanine incorporation into lipoteichoic acid in GBS are coded by the *dltABCD* operon. *DltA* encodes the D-alanine-D-alanyl carrier ligase, that carries out the final step of the incorporation process. Mutants deleted for *dltA* bound significantly higher amount of CAMPs on their surface and were more susceptible to killing. The first proposed explanation for the role of D-alanylation of surface lipoteichoic acid was lowering the net bacteria surface charge, and thus reducing the efficiency of electrostatic binding of CAMPs (Poyart *et al.*, 2001; Poyart *et al.*, 2003). However, recent studies revealed that lipoteichoic acid D-alanylation provides resistance only to a specific subset of CAMPs, depending on their length and charge density. The reduced bacterial surface charge was not

proven to effectively decrease the number of CAMPs bound to the cell surface, and instead, D-alanylation most likely increases cell wall density and surface rigidity via conformational changes in the lipoteichoic acid structure, thus reducing the penetration of CAMPs. Decreased cell wall flexibility and permeability could provide protection against liner CAMPs, instead of reducing electrostatic interactions between these molecules and the bacterial surface (Saar-Dover *et al.*, 2012).

Ectonucleotidases

In order to avoid recognition and killing by the host immune system, bacteria developed several strategies. To mute proinflammatory responses, for example, GBS can interfere with the host purinergic signaling pathway. This pathway mainly relies on the balance of extracellular ATP and extracellular adenosine as effector molecules. Host cells secrete ATP in response to cell damage or infection, which acts as a proinflammatory signal, while adenosine antagonizes this effect (Vitiello *et al.*, 2012). Interestingly, human neonatal blood contains a higher concentration of adenosine, polarizing their innate immunity toward an anti-inflammatory response (Kollmann *et al.*, 2012), which makes them more vulnerable to GBS infections (Firon *et al.*, 2014).

NudP

The cell wall-bound LPxTG protein, NudP, has been identified in GBS bovine infections as highly immunogenic protein factor. Later it was revealed to have ecto-5'-nucleoside diphosphate phosphohydrolase (NudP) activity, that makes it capable of hydrolyzing ribo- and deoxyribonucleoside 5'-di- and -monophosphates (but not 5'-triphosphates). NudP enzymatic activity was shown to contribute to bacterial evasion of clearance by leukocytes on processes dependent on extracellular adenosine concentrations. In animal models, loss of NudP activity also impacted virulence and colonization, confirming its role in subverting the host immune response (Firon *et al.*, 2014).

CdnP

GBS produces another important signaling nucleotide, cyclic-di-AMP, which is released in the cytosol of infected macrophages to stimulate the production of type-I interferon (IFN). GBS also expressed a surface-bound LPxTG ectonucleotidase, CdnP, which hydrolyzes the c-di-AMP, resulting in decreased IFN production. Indeed, CdnP deficient mutants were more efficiently killed by immune cells due to higher host IFN-β production. CdnP can be considered as a virulence factor that contributes actively to the dampening of host immune responses (Andrade *et al.*, 2016).

Adhesion factors

Among the surface-bound adhesins of GBS, most of them were shown to promote the binding to extracellular matrix (ECM) components, while a few other may interact with specific host receptors, I have chosen to present a few examples below.

The Alp protein family and Alpha-C proteins

Alpha and alpha-like proteins are very abundant on the surface of *S. agalactiae* and were shown to confer protective immunity. The Alpha-like protein family is comprised of six allelic variants, Cα, Alp1, Alp2, Alp3, Rib, and Alp4 (formerly known as Alp2, Alpha-C, Rib, R28 and Epsilon, as described by Lindahl and colleagues) (Maeland *et al.*, 2015; Lindahl, Stålhammar-Carlemalm and Areschoug, 2005). All known Alps are conserved and stable throughout the different serotypes of GBS, despite their modular nature: they form mosaic structures on the bacterial surface, with the N- and C-terminal of the Alp proteins possessing different immunogenic specificity, which is responsible for their high immunogenicity (Maeland *et al.*, 2015). Most Alps contain tandem repeating elements with variable lengths that allow for high antigenic variation and a low repeat number might be important for immune-evasion during infection (Landwehr-Kenzel and Henneke, 2014).

Most GBS strains code for only one of the Alp family proteins, for example, among the most common human colonizing and invasive serotypes, Rib is mainly expressed by serotype III strains (including ST-17 isolates), while Alpha-C proteins are mainly found in serotypes Ia, Ib and II (Madoff *et al.*, 1991; Stålhammar-Carlemalm, Stenberg, Lindahl, 1993).

Except for Alpha-C proteins (ACP), the role of Alpha-like proteins in streptococcal virulence is poorly understood beside their strong immunogenic activity that could make them potential targets for vaccine development. ACPs, however, were shown to bind to α1ß1 integrins, which is an important mechanism for promoting human epithelial cell invasion (Bolduc and Madoff, 2007). ACP also promotes GBS invasion of the cervical epithelium, mainly by interacting with glycosaminoglycan on the host cell surfaces, promoting bacterial internalization (Rajagopal, 2009).

Fbs proteins

The Fbs fibrinogen-binding protein group in GBS is constituted of FbsA, FbsB and FbsC, named in alphabetical order according to their time of discovery. Their distribution among GBS clonal groups is different.

The initially identified FbsA and FbsB proteins were shown to bind to both immobilized and soluble fibrinogen, and before the discovery of FbsC, GBS interactions with fibrinogen

were uniquely attributed to these two factors. FbsA and FbsB are structurally unrelated. FbsA is a cell wall-anchored LPxTG protein, while FbsB is a secreted protein. FbsC, identified more than a decade later, is also covalently linked to the cell wall by an LPxTG anchor, but its protein structure displays no significant similarity to that of FbsA (Buscetta *et al.*, 2014).

FbsA triggers large-scale aggregation of human plasma fibrinogen into a polymer-like network, leading to a formation of fibrinogen-coating on the bacterial cell wall that might confer protection against phagocytosis (Pierno *et al.*, 2006). FbsA also induces the fibrinogen-dependent aggregation of platelets (Schubert *et al.*, 2002), which also affects fibrin-dependent thrombogenesis (Pietrocola *et al.*, 2006). These effects, together with FbsA promoting GBS growth in human blood suggest an important role for this factor in GBS-induced endocarditis (Pietrocola *et al.*, 2005). The domain responsible for fibrinogen binding has been identified as a tandem repeat region of 16 biding binding units, and even a truncated form of FbsA, containing only one such region has been shown to bind to its target, confirming the high binding efficiency of this protein (Schubert *et al.*, 2002).

FbsA was also shown to be important in adherence and invasion of human epithelial cells, such as A549 lung epithelial cells, and human brain microvascular endothelial cells (HBMEC) (Schubert *et al.*, 2004; Tenenbaum *et al.*, 2005). Interestingly, when heterogeneously expressed in *L. lactis*, FbsA only promoted adhesion to host cells (Schubert *et al.*, 2004), suggesting that there might be an other fibrinogen-binding protein responsible for host-cell invasion, probably, but not necessarily FbsB (Gutekunst *et al.*, 2004; Schubert *et al.*, 2004). A subsequent analysis also questions the potential co-involvement of FbsA and FbsB in the same fibrinogen-binding reaction, as while FbsA uniquely binds human fibrinogen, the FbsB N-terminal may bind to human fibrinogen, but its C-terminal is specific to the bovine version of the substrate. This suggests that the respective roles of FbsA and FbsB might be dependent on the specific host (Devi and Ponnuraj, 2010).

The third Fbs, FbsC, was found in the majority of sequenced clinical GBS isolates, but a frameshift mutation was found in the hypervirulent ST-17 strains. FbsC was shown to be important in both adhesion to and invasion of human epithelial (Caco2 intestinal and A549 pulmonary) and endothelial (hCMEC/D3 brain) cells. It also contributes to *in vitro* biofilm formation by GBS NEM316 strain. Finally, the *fbsC* deletion mutant was impaired in brain colonization and attenuated in an *in vivo* mouse model (Buscetta *et al.*, 2014).

The findings of Buscetta and colleagues also highlight the importance of the distribution of Fbs genes in different GBS serotypes or MLST-types. While they found that FbsC was not expressed in the ST-17 hypervirulent strains, another study revealed that FbsA and FbsB expression was controlled by a specific regulatory factor in ST-17 strains (Al Safadi *et al.*, 2011). However, this is most likely not the reason behind the increased fibrinogen binding

activity of ST-17 strains, as since then another group of fibrinogen-binding factors, Srr1 and Srr2 (Serine-rich repeat) has been discovered, and the ST-17-associated Srr2 allele has a magnitude stronger binding ability (as detailed later) (Seo *et al.*, 2013; Six *et al.*, 2015).

Srr1

Srr1 (for serine-rich repeat) is the first glycosylated protein identified and characterized in GBS (van Sorge *et al.*, 2009; Mistou *et al.*, 2009), as the homolog of GspB protein in *S. gordonii*. Srr1 is encoded in the *secA2* locus which includes genes encoding the SecA2 machinery necessary for the secretion of this protein and several glycosyltransferases involved in Srr1 glycosylation. The glycosylation of Srr1 contributes to its resistance to proteolytic attack and modulates bacterial adherence and virulence (Mistou *et al.*, 2009). Srr1 directly binds the Aα chain of human fibrinogen via a specific binding region (BR domain). The affinity of this binding has a Kd of $2.1 \times 10^{(-5)}$, measured for recombinant Srr1 in vitro (Seo *et al.*, 2012; Seo *et al.*, 2013).

A role for Srr1 in GBS virulence was demonstrated in a mouse model using the GBS strains NCTC10/84 and NEM316 (van Sorge *et al.*, 2009; Mistou *et al.*, 2009; Seo *et al.*, 2012). Subsequently, Srr1 was shown to promote endocarditis in a rat infection model, due to its adhesion to blood platelets via fibrinogen-binding (Seo, Xiong, Sullam, 2013). Srr1 binding to fibrinogen also enhanced GBS adhesion to human cervical and vaginal epithelial cells (Wang *et al.*, 2014).

SfbA

SfbA (for Streptococcal Fibronectin-Binding protein A) is a cell wall anchored protein, which has been identified as a major fibronectin-binding protein of GBS (Mu *et al.*, 2014). SfbA binds to purified fibronectin *in vitro*, and mediates invasion of human brain endothelial cells (hBMEC). Heterologous expression of SfbA in *Lactococcus lactis* a non-invasive bacterium, was shown to promote fibronectin binding and efficient entry into hBMEC cells. Fibronectin binds to host integrins receptors expressed by hBMECs which were shown to be important for bacterial invasion. Thus, GBS most likely uses binding to fibronectin as a mimicry mechanism to induce cellular uptake (Mu *et al.*, 2014). It was also shown to be necessary for invasion of astrocytes of the central nervous system for the ST-17 COH1 strain, although it was not involved in adhesion (Stoner *et al.*, 2015). SfbA therefore seems to be an important factor in GBS pathogenesis, and especially for blood-brain barrier penetration. Type-related distribution of *sfbA* is yet to be analyzed, but the initial study identified the gene in 5 strains of different sequence types, including the ST-17 strain COH1, the ST-23 strain NEM-316 and the ST-7 strain A909. They also reported different levels of fibrinogen-

binding for these strains but they did not necessarily correlate with *sfbA* transcription levels in the same isolates (Mu *et al.*, 2014).

PbsP

Another blood component often recruited to the bacterial cell surface is plasminogen, which is subsequently converted into plasmin by host enzymes, allowing the bacteria to access deeper tissues and distant organs through plasmin proteolytic activity. Indeed, Plasminogen-coated GBS displayed increased adhesive and invasive properties to hBMECs even in an early stage of infection, and they subsequently induced lysis hBMECs due to a plasmin-like proteolytic activity. Plasmin(ogen)-treated GBS caused meningitis neonatal mice more effectively than untreated bacteria, confirming the role of plasminogen-binding in crossing the blood-brain barrier and meningitis development (Magalhaes *et al.*, 2013). PbsP is an LPXTG protein identified recently for its ability to bind plasminogen) (Buscetta *et al.*, 2016). PbsP was shown to be conserved among most GBS lineages, and serves as the major plasminogen-binding factor in non-ST-17 strains (Buscetta *et al.*, 2016).

Lmb

Another ECM constituent is laminin, which becomes accessible to pathogenic bacteria upon the damaging or disruption of the endothelial structures. The first laminin-binding protein, Lmb, in GBS was identified in 1999, based on sequence homology with known laminin-binding factors (Spellerberg *et al.*, 1999). Lmb is a lipoprotein expressed in operon with another LPXTG protein which is a C5a peptidase that cleaves the complement factor C5a. The involvement of Lmb in laminin binding was confirmed using purified laminin (Spellerberg *et al.*, 1999). During the infectious process, Lmb largely contributes to invasion of human brain microvascular endothelial cells (HBMECs) (Tenenbaum *et al.*, 2007).

The crystallographic structure of Lmb revealed a metal binding crevice, that can maintain a zinc ion coordinated to three histidine residues. These residues are essential to laminin-binding activity, most likely through maintaining the correct functional confirmation of Lmb (Ragunathan *et al.*, 2013). Due to its unique Zn-binding activity, Lmb was later identified as a Zn ABC-type solute-binding protein, that in combination with a translocon system, form a zinc-ABC transporter which is the main zinc-uptake mechanism of GBS. This system is essential for bacterial growth and correct morphology in zinc-limited environments, such as human body fluids (Moulin *et al.*, 2016).

BibA

The gene *gbs2018* was first identified as one of the most upregulated genes in a *covR* transcriptional regulator mutant (Lamy *et al.*, 2004). Based on allelic variations of this gene, a simple PCR method to detect specifically ST-17 strains was proposed (Lamy *et al.*, 2006).

BibA contributes to bacterial survival in human blood and to resistance to opsonin-dependent phagocytotic killing, by specifically binding to the human C4-binding protein and thus interfering with the activation of the complement pathway. It also contributes to adherence to human cervical and epithelial cells (Santi *et al.*, 2007). Accordingly, in a rat infection model, higher BibA expression correlated with increased GBS virulence (Santi *et al.*, 2009). The expression of BibA is strictly regulated by the two-component regulator of virulence CovR, temperature and pH (Santi *et al.*, 2009a; Rajagopal, 2009).

ST-17 specific surface proteins

As it has been pointed out earlier, different serotypes or clonal complexes of GBS often distinctively code for certain virulence factors or specific alleles thereof, while others are more diversely expressed. This is the case for members of the ST-17 complex as well, which, due to the very strong clonality of this group, all code for certain specific surface proteins that are not or only rarely can be found in GBS isolates of any other origin. The genes coding for these factors have been identified by sequence alignment of a diverse set of GBS genomes (Brochet *et al.*, 2006). Among the ST-17 specific genes, two are coding for distinct surface proteins, *srr2* and *hvgA*, former a variant of *srr1*, latter a second allele of the *bibA* locus in non-ST-17 strains. These two genes to date have been solely identified in ST-17 strains. Also, characteristic to ST-17 strains is the PI-2b type pilus island, an allelic variant of the PI-2a pilus locus, but it has been equally found in several non-ST-17 bovine isolates, as detailed in the corresponding chapter (chapter D and E).

Srr2

The ST-17 specific allelic variant of the *srr* locus, *srr2* codes for the major fibrinogen-binding LPxTG surface protein of these strains (Brochet *et al.*, 2006). Both Srr variants bind the fibrinogen Aα chain, the binding affinity of Srr2 is, however, one magnitude higher than that of Srr1, with a KD of 3.7×10^{-6}. This difference in activity can be explained by the differences in the “latch” segment of the binding domain that binds fibrinogen via a "dock, lock, and latch" mechanism (Seo *et al.*, 2013).

The higher fibrinogen-binding affinity of Srr2 could account for the increased virulence associated with ST-17 strains. Furthermore, in contrast with Srr1, Srr2 was also revealed to

bind to human plasminogen and plasmin (Six *et al.*, 2015). These proteins are important components of the fibrinolysis pathway, and via Srr2 binding to these elements, bacterial aggregates are formed which increases bacterial uptake by phagocytes. Similar to the high sialic acid content of the serotype III capsule, Srr2 also contributes to increased intracellular survival of the phagocyted bacteria. In a murine model of meningitis, Srr2 was shown to contribute to bacterial persistence, dissemination and invasiveness, including the breaching of the blood-brain barrier (Six *et al.*, 2015).

HvgA

HvgA was identified as a ST-17–specific surface-anchored LPxTG protein, where the name stands for "hypervirulent GBS adhesin", as it was shown to largely contribute to ST-17 hypervirulence. While the specific target molecule of HvgA-mediated bacterial adhesion is not known, it was shown to promote adhesion to a plethora of different human cell lines, including intestinal epithelial cells, choroid plexus epithelial cells, and HBMECs of the blood-brain barrier (BBB). In a murine oral infection experiment, mortality to HvgA expressing ST-17 strains was also significantly higher than with a *ΔhvgA* isogenic mutant (Tazi *et al.*, 2010).

Similarly, in competitive colonization experiments, HvgA-expressing ST-17 strains outcompeted by far both non-ST-17 strains and the *ΔhvgA* mutant, suggesting that the presence of HvgA confers a selective advantage already at this early step of the infection. Subsequently it also contributes to the breaching of the intestinal barrier and bacterial dissemination in the blood.

The quantity of bacteria retrieved from the brains of infected mice was also significantly higher for HvgA-expressing strains, suggesting that HvgA is also needed for the invasion of the central nervous system (CNS), which corroborates the *in vitro* HBMEC adhesion results. This invasion mechanism was shown to be independent of acute inflammation-induced local tissue damage. Furthermore, bacteria were retrieved from choroid plexus cells, meninges, and brain microcirculation veins, as observed in histopathological samples of fatal human neonatal late onset meningitis cases (Tazi *et al.*, 2011). HvgA is, in conclusion, a critical virulence factor in GBS infections of the CNS, especially in the neonatal context and largely contributes to specific ST-17 hypervirulence.

C. Virulence regulation in GBS

GBS is a commensal bacterium that resides in the intestinal and genital tracts of humans. However, under certain circumstances, it is able to switch into a life-threatening pathogen able to translocate epithelial barriers and disseminate in blood to reach for example the central nervous system. Therefore, expression of virulence factors must be tightly controlled and finely tuned.

In Gram-positive bacteria, bacterial responses to environmental conditions are often mediated by two component systems (TCS). A typical TCS is constituted of a membrane-bound histidine protein kinase as a sensor and a cytosolic response regulator. Signal transduction occurs through autophosphorylation of the kinase which then transfers the phosphoryl group to the regulator. Phosphorylation (usually at a conserved aspartate residue) induces conformation change of the response regulator molecule that alters its DNA binding affinity to the target promoters.

In silico genomic analysis of the GBS NEM316 strain revealed the presence of 20 sensor histidine kinase elements and 21 predicted response regulators (Glaser *et al.*, 2002) while only 17 TCS were found in the genome of strain 2603V/R (Tettelin *et al.*, 2002). This number of potential two-component systems is higher than in related organisms: to date, 14 such regulatory systems have been described in *S. pneumoniae*, and 13 in *S. pyogenes*, which suggests that *S. agalactiae* may have higher adaptive capacity to different environments that might explain its wider host-range potential (Glaser *et al.*, 2002).

Out of these potential TCS, only four have been characterized in more or less detail in GBS, namely the CovR/CovS, DltR/DltS, RgfC/RgfA and CiaR/CiaH systems (Rajagopal, 2009). In addition to these TCS, a few “stand-alone” regulators have been shown to play a role in virulence. RovS was suggested to regulate the expression of fibrinogen-binding factors, especially FbsA (Pérez-Pascual *et al.*, 2015), while the RgA/RogB system is responsible for the regulation of the expression of two other surface-expressed virulence factors in GBS NEM316, the fibrinogen-binding adhesin Srr1 and PI-2a type pili, that will be described in detail later (Chapter E). Among these systems, CiaR/CiaH and especially CovR/CovS were already proven to play a role in GBS virulence.

CiaRH is important for GBS survival inside neutrophils and macrophages, and more generally in resistance to oxidative stress and susceptibility to antimicrobial peptides although the target genes are not yet fully identified (Quach *et al.*, 2009; Mu *et al.*, 2016). Its activity, however is of lower importance than the CovR/S system, which is essential for most known mechanisms of GBS virulence as a master switch of an intricate regulatory network, as detailed below.

CovR/S

The best studied two-component regulatory system in GBS is the CovS/CovR, which was revealed to be the global regulator of virulence-associated factors. The two genes that code for CovS and CovR in GBS are part of a seven-gene operon (Lamy *et al.*, 2004), which is positively by CovR as revealed by microarray analysis (Lembo *et al.*, 2010). A mutant deleted for both genes was severely affected in on phenotype, including increased hemolytic activity and pigmentation and reduced CAMP activity. The *ΔcovSR* mutant was severely impaired for virulence in a rat sepsis model with an LD(50) increased by 3-logs (Lamy *et al.*, 2004).

Comparative transcriptomics using macro-arrays identified 76 genes potentially repressed and 63 up regulated by the CovS/R TCS. Subsequently, a microarray analysis revealed that CovR positively regulates its own expression and modulates the expression of another 153 genes (Lembo *et al.*, 2010). Similar results were obtained by comparative mass-spectrometry analysis of the exoproteome of GBS NEM316 strain and its isogenic *ΔcovRS* mutant. Out of a set of 53 identified proteins, 41 were found to be affected by the deletion of the CovR/S system, revealing drastic changes in the bacterial surface (Papasergi *et al.*, 2013).

CovR/S-deletion strongly influence GBS adherence to human epithelial cells of different origins, such as vaginal, cervical and respiratory epithelia, adhesion to ECM components and *in vitro* biofilm formation, confirming the master regulator role of this TCS in the expression of virulence genes (Jiang *et al.*, 2005; Park *et al.*, 2012).

While phosphorylation of CovR by CovS occurs at the D53 aspartate residue, a eukaryotic-like serine/threonine kinase Stk1 was shown to be able to phosphorylate CovR at the T65 threonine residue (Lin *et al.*, 2009, Lembo *et al.*, 2010). Phosphorylation at Thr65 impairs phosphorylation of CovR at Asp53 by CovS and thus modulate CovR activity. Phosphorylation by two independent enzymes, the co-regulated CovS and the independently expressed Stk1 seems to co-modulate the intricate system of CovR-dependent virulence gene regulation (Rajagopal *et al.*, 2006; Henneke and Landwehr-Kenzel, 2014).

A multi-spanning Abi-domain transmembrane protein, Abx1, has been identified as a regulator of the CovS histidine kinase. Abx1 is essential for β-hemolysin/cytolysin production, as it regulates the *cyl* operon, as well as several other virulence factor-coding genes in a CovR/S-dependent manner. Interestingly, Abx1 overexpression activates virulence gene expression via CovS, even in absence of a specific external signal, while absence of Abx1 leads to accumulation of CovS locked in its kinase-competent state. Abx1 was confirmed to form a signaling complex with CovS via a direct protein-protein interaction, further contributing to the “fine-tuning” of GBS signal processing and virulence gene regulation (Firon *et al.*, 2013).

D. Pili in Gram-positive bacteria

Pili have been identified on the surface of Gram-negative bacteria since the early 1950s, and were described as regular polymeric structures formed by the non-covalent association of protein subunits (designated as pilins). Due to their relative thinness, Gram-positive pili are more difficult to visualize and thus were discovered later than their Gram-negative counterparts. The first pilus-like structures in Gram-positive bacteria, were detected in 1968, on the surface of *Corynebacterium renale* by electron microscopy (Yanagawa *et al.*, 1968). Subsequently, similar structures were described in *Actinomyces naeslundii* (Cisar, 1988; Yeung & Ragsdale, 1997; Yeung, 1998), *Corynebacterium diphtheriae* (Ton-That and Schneewind, 2003; Gaspar and Thon-That, 2006), *Streptococcus parasanguinis* (Ton-That and Schneewind, 2003), and *Streptococcus salivarius* (Handley, Carter and Fielding, 1984; Weerkamp, van der Mei and Liem, 1986).

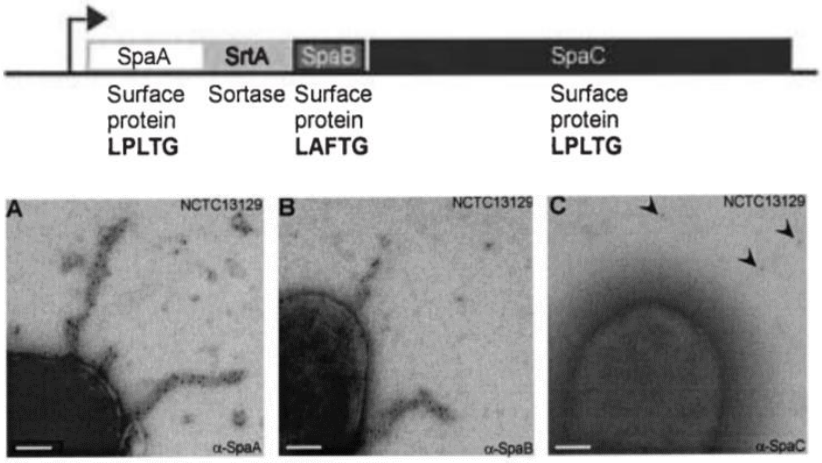


***Figure 1* | Genomic organization of the SpaA pilus locus in Corynebacterium diphtheriae.** *The three spa genes code for the respective pilin subunits and srtA code for a class C sortase involved in pilus polymerization. The arrow indicates the predicted promoter. As seen on immunogold labelled C. diphtheriae strain NCTC13129 by electron microscopy A) the major pilin SpaA can be detected as a long pilus filament on the cell surface, B) the anchoring subunit, SpaB localizes to the basal site of the pilus, and C) the putative adhesin, SpaC only decorates the tip of the pili. This organization is typical for trimeric pili in Gram-positive bacteria. Image adapted from Ton-That & Schneewind 2003.*

A major discovery occurred in 2003 with the identification and molecular characterization of the first genetic locus encoding a pilus in *C. diphtheriae* (Ton-That and Schneewind, 2003), opening the way for pilus loci discovery through *in silico* analysis (Figure 1). Pili were also identified in the three major human pathogenic Streptococcus species: *Streptococcus pyogenes* (GAS, Group A *Streptococcus*) (Mora *et al.*, 2005), *Streptococcus agalactiae* (GBS, Group B *Streptococcus*) (Lauer *et al.*, 2005; Dramsi *et al.*, 2006; Rosini *et al.*, 2006) and

Streptococcus pneumoniae (Barocchi *et al.*, 2006), and subsequently in *Streptococcus gallolyticus* (Danne *et al.*, 2011). While several studies identified two distinct types of pilus-like surface appendages in Gram-positive bacteria: short rods, also called fibrils, in *Streptococcus oralis* and *Streptococcus gordonii*, and longer flexible filamentous structures as observed in *Corynebacterium* species, and pathogenic Streptococci, the term "pilus" refers uniquely to these latter in Gram-positive bacteria (Telford *et al.*, 2006).

The biochemical composition of pili was first studied in *A. naeslundii* by Yeung and colleagues who identified the presence of an LPxTG-containing sorting signal in the sequence of the two pilins coded in this species, FimA and FimP (Yeung and Ragsdale, 1997; Yeung, 1998). The sortase-dependent assembly of pilus subunits into covalently linked polymeric filaments was then described in detail in 2003 for *C. diphtheriae* (Thon-That and Schneewind, 2003). Subsequently, *in silico* search for operons containing sortase and LPXTG-protein coding genes led to the rapid expansion of pilus discovery in Gram-positive bacteria (Telford *et al.*, 2006; Scott and Zahner, 2006; Mandlik *et al.*, 2008), which all share a similar organization (Figure 2).

This covalent structure and assembly process is unique to Gram-positive bacteria. The sortase enzymes responsible for pilus polymerization belong to the class C of sortases (Dramsi, Trieu-Cuot and Bierne, 2005), except for *S. pyogenes* pili that are polymerized by a class B sortase (Kang *et al.*, 2011). The covalent anchoring of the pilus polymer to the cell wall peptidoglycan is then carried out by the housekeeping sortase A (SrtA) or another sortase.

All Gram-positive pili share a similar structure and are constituted of two or three different pilus subunits covalently linked to form a hair-like structure of approx. 3 nm diameter and 0.1 to 5 μm in length on the bacterial surface (Danne and Dramsi, 2012). Pilus subunits can also be found on the surface as monomers, homo- and/or heterodimers on the bacterial surface (Chang *et al.*, 2011). In the typical heterotrimeric pilus, one of the three pilin subunits constitute the backbone of pilus filament often referred to as the major or backbone pilin, one of the ancillary pilins is usually found along the backbone filament or at its tip, while the other ancillary pilin is localized close to the bacterial surface, and shown to serve as anchor (Nobbs *et al.*, 2008).

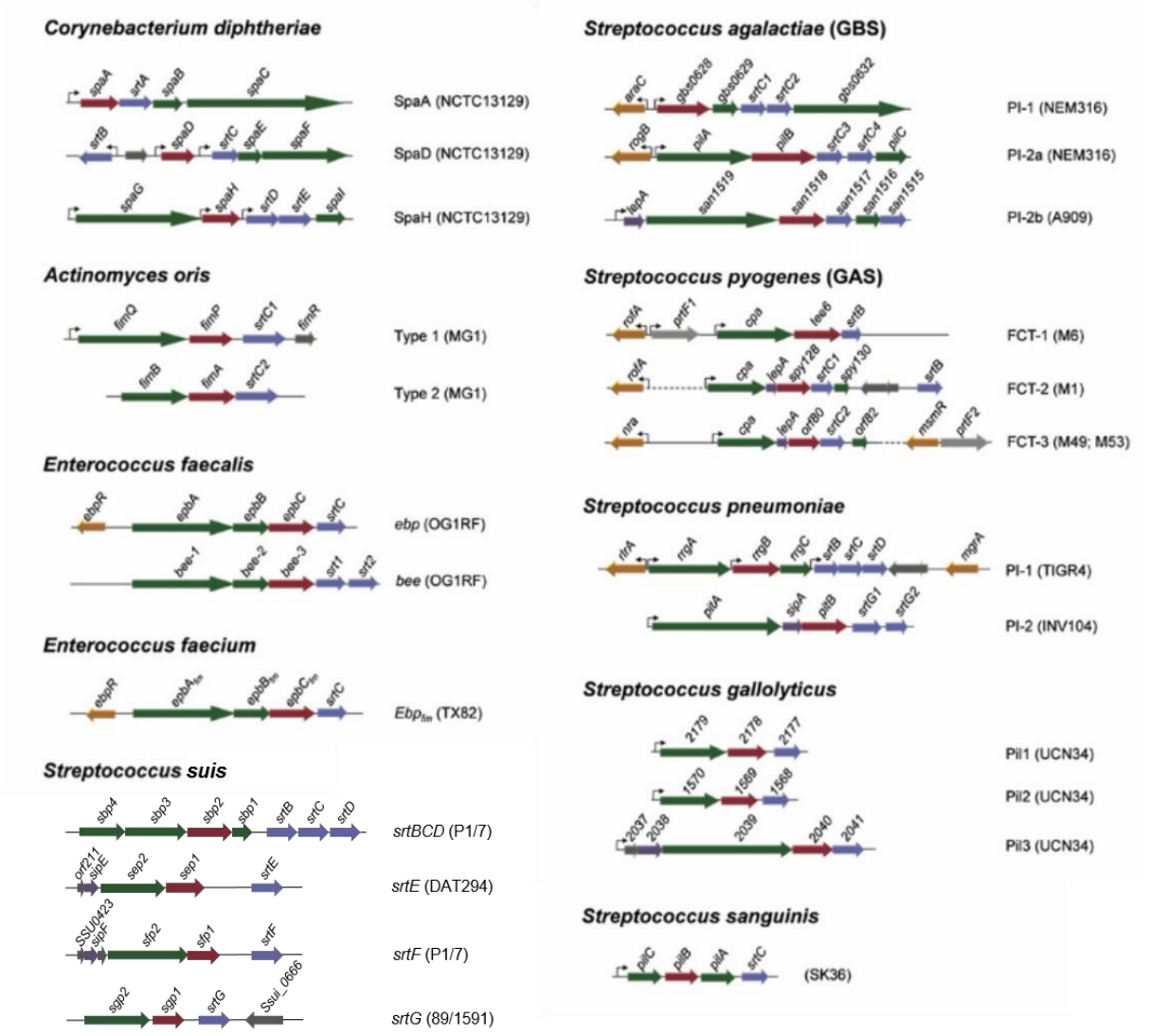


Figure 2 | Organization of pilus loci in different Gram-positive bacteria. *Genes coding for sortases involved in the pilus assembly are marked in light blue. The putative adhesin is marked in green, other pilus subunits in red, while pilus-associated regulators appear in yellow, putative signal peptidase-like chaperone in purple, and grey denotes genes with unknown function. Arrows mark the predicted transcription start sites. The currently used name of the pilus and strain are indicated on the right. Adapted from Danne and Dramsi, 2012.*

General mechanism of pilus-assembly

The precise steps of the pilus polymerization and anchoring were described in first in *C. diphtheriae* (Ton-That and Schneewind, 2003). As described step-by-step in Figure 3, first the LPxTG-containing pilus subunits are secreted in a Sec-dependent manner and remain cell associated via their C-terminal membrane-spanning domains (Bae and Schneewind, 2003; Mazmanian *et al.*, 2001). They are subsequently recognized by a pilus-specific sortase, cleaved between the threonine (T) and glycine (G) residue of the LPxTG motif, and linked to the catalytic pocket of the pilus-specific sortase enzyme by a thioester bond. Subsequent polymerization of the pilus subunits in a head-to-tail manner is also mediated by the same sortase enzymes. The threonine (T) residue of the C-terminal LPxTG of one subunit is linked to the amino-terminal lysine (K) side-chain of the VYPKN pilin motif of a neighboring pilin subunit by an amide bond. After several steps of addition of backbone pilin subunits at the base of the pilus, the pilus polymer is anchored to the peptidoglycan through incorporation of the minor anchor subunit (Figure 3). Anchoring to the peptidoglycan is usually carried out by a different sortase enzyme, either the housekeeping sortase A (SrtA) coded outside of the pilus locus, or a second pilus-specific sortase (Swierczynski and Ton-That, 2006).

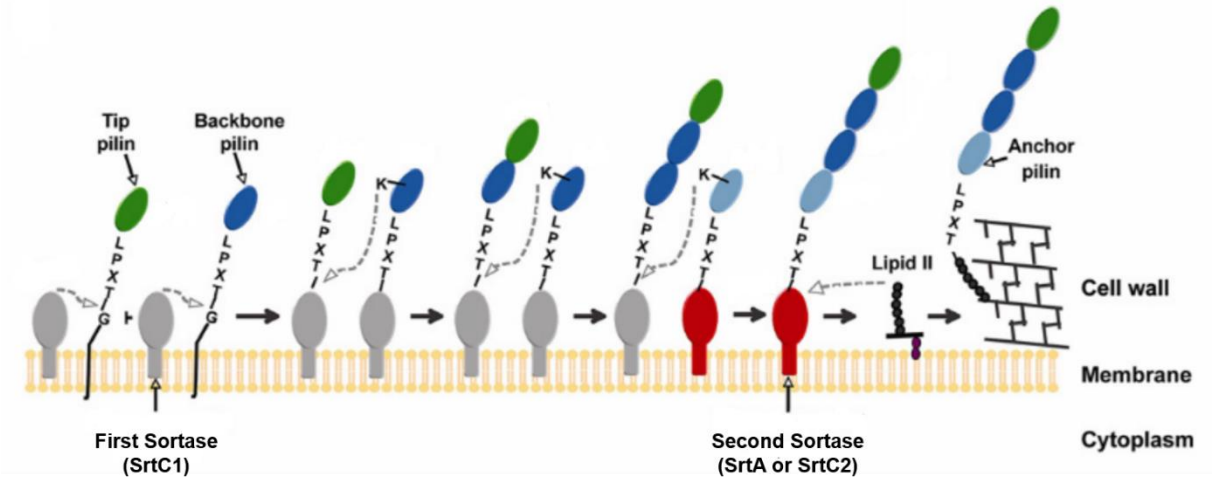


Figure 3 | General mechanism of pilus assembly and anchoring to the cell wall. *The LPXTG sorting signals of the backbone pilin (BP) and the tip pilin (AP-1) are cleaved by the pilus-specific C-type sortase (SrtC1) at the threonine, generating an acyl-enzyme intermediate. Subsequently, an intermolecular isopeptide bond is formed between AP-1 and BP as a result of the nucleophilic attack by the lysine (K) side chain in the BP pilin motif. This catalyzes further pilus polymerization with the addition of BP subunits. Incorporation of the anchoring pilin AP-2, is carried out either by the housekeeping sortase SrtA, or an other C-type sortase, resulting in the termination of pilus polymerization and its covalent attachment to the cell wall peptidoglycan.*

Beside the anchoring signal, major pilin subunits contain a so-called VYPKN pilin motif at their N-terminus (Yamaguchi and Matsunoshita, 2004). Although the sequence of the pilin motif can vary, the lysine residue is conserved and was suggested to be essential for the polymerization reaction (Ton-That, Marraffini and Schneewind, 2004). Another characteristic element of pilins is the so-called E-box that contains a conserved glutamic acid, which is

often necessary for the folding of the individual subunits, but is not universally present in such molecules (Cozzi *et al.*, 2012).

Accessory pilins also contain the C-terminal LPxTG sorting sequence but the pilin motif is often absent, additionally, unlike the major pilins, accessory subunits do not form long polymeric filaments. Therefore, it has been suggested that an accessory subunit, usually AP-1, initiates the pilus polymerization and is the first pilin to be incorporated, which defines its localization at the pilus tip (Danne and Dramsi, 2012). In heterotrimeric pili, the second, AP-2, accessory pilins are most often found at the base of the pilus shaft, serving as an anchor. Based on this model, the length of the pilus is determined by the abundance of pilus subunits that are available at the surface, notably the ratio of backbone pilin subunits to anchor pilin levels in a stochastic manner (Nobbs *et al.,* 2008; Swierczynski and Ton-That, 2006). In simpler pili constituted of only 2 subunits, the major pilin forms the pilus backbone while the only accessory pilin is located at the distal end (tip) of the filament (Budzik *et al.*, 2009; Mishra *et al.*, 2007).

Further insight was gained to the function of these motifs when first pilin crystal structure was determined for Spy0128, the major pilin from *S. pyogenes*, to 2,2 Å resolution (Kang *et al.*, 2007). This also led to the establishment of a paradigm which is still being held through subsequent structural analyses of pili: a modular construction based on repeated immunoglobulin-like domains containing intramolecular isopeptide bonds (Kang and Baker, 2012). These bonds, forming between Lys and Asn (or Asp) side chains within the pilin subunits, stabilize the long pilus molecules, protecting them against proteolysis, heat and mechanical stress such as shearing (Kang and Baker, 2009; Hendrickx *et al.*, 2012). The glutamic acid of the E-box motif of the individual pilin molecules is also involved in stabilizing these internal isopeptide bonds (Cozzi *et al.*, 2012). Interestingly, unlike most gram-positive major pilins, Spy0128 does not contain the canonic VYPKN pilin motif, but the molecule still contains a lysine residue that is essential for its structure (Kang *et al.*, 2007).

Later, other major pilins, although largely differing in size and sequence, have been shown to be built similarly of tandem Ig-like domains of two types: CnaA and CnaB domains, as first described in the cell-surface adhesin Cna of *S. aureus* (Zong *et al.* 2005; Kang and Baker, 2012). While the *S. pyogenes* Spy0128 pilin has a relatively simple architecture with two CnaB domains (Kang *et al.*, 2007), major pilins are often more complicated structures of several Cna domains. Tip pilins are structurally more specialized in accordance with their adhesin roles, containing extra domains with adhesive functions that are connected to the pilus shaft via a Cna domains stalk (Kang and Baker, 2012).

Pilus functions

As long, highly stable surface appendages, pili and their components have the potential to be the first bacterial proteins that come into contact with host cell surfaces and tissue factors. Thus, pili were shown to play important roles during adhesion and colonization of host tissues.

Adhesion to eukaryotic cells

Gram-positive pili were first shown to be important in adhesion host cells in *C. renale*, which adheres to kidney and bladder epithelial cells (Honda and Yanagawa, 1975 and 1978).

The first pilin that was proven to play an important role in adhesion to host epithelial cells was identified in our lab: the minor subunit AP1 of the pilus type PI-2a of *S. agalactiae* promotes adhesion to lung epithelial cells (Dramsi *et al.*, 2006). Subsequently, pilus-mediated adhesion to host epithelial cells was also demonstrated in other streptococcal species.

For example, pilus structures were shown to be necessary for adhesion to pharyngeal cells in the *S. pyogenes* M1_SF370 strain (Abbot *et al.*, 2007; Manetti *et al.*, 2007). Among the pilus proteins, the Spy0125, or Cpa, tip pilin was shown to be specifically driving the adhesion to pharyngeal cells, while the pilus shaft is most likely needed for its accessible exposure in the cell surface. The N-terminal of Cpa harbors a thioester-containing functional domain, which is necessary for forming specific covalent linkage with its host receptor domains. Such thioester domains have been identified in cell wall proteins of other Gram-positive bacteria that might play a similar role in infection (Linke-Winnebeck *et al*., 2013).

Similarly, in *S. pneumoniae*, the PI-1 pilus was also shown to be necessary for adhesion to A549 lung epithelial cells (Barocchi *et al.*, 2006). The overexpression of the minor pilin, RrgA, increased adhesion under static conditions, suggesting that the pilus shaft *per se* is not indispensable for adhering to host cells (Nelson *et al.*, 2007). The possible role of the pili in pneumococcal virulence, however, is questionable, as the PI-1 pilus is only present in 25-30% of *S. pneumoniae* clinical isolates (Barrochi *et al.*, 2006).

Apart from allowing the optimal display of the pilus associated adhesin beyond the capsule, one may ask if the backbone protein could have other functions. In GBS, it was suggested that backbone pilins of both PI-1 and PI-2a type pili contribute to paracellular crossing of epithelial membranes, as described in the corresponding chapter (Pezzicoli *et al.*, 2008).

Adhesion to extracellular matrix proteins

Accessory pilins (AP-1) have been described as adhesins in several bacterial species (Jones *et al.*, 1995; Dramsi *et al.*, 2006; Mandlik *et al.*, 2007), and several of them were identified as MSCRAMMs (Microbial Surface Components Recognizing Adhesive Matrix Molecules). These are bacterial surface proteins that specifically interact with components of the host extracellular matrix, such as collagen-, fibrinogen-, fibronectin- or heparin-related polysaccharides (Patti *et al.*, 1994). These adhesins therefore mediate binding to extracellular-matrix (ECM) proteins of the host and are thus crucial for colonization and invasion.

In *S. gallolyticus*, it was shown that the Pil1-associated adhesin, Gallo_2179 or Acb efficiently binds to collagen type I, the major component found in cardiac valves, and collagen type IV, usually enriched in the basal lamina of precancerous tissue formations (Sillanpaa *et al.*, 2009; Danne *et al.*, 2011). Interestingly, this adhesin is homologous to collagen-binding proteins of other Gram-positive species, such as Cna of *Staphylococcus aureus* and Ace of *Enterococcus faecalis*, but while these proteins are monomeric surface-anchored LPXTG proteins. The pilus-associated expression of Acb in *S. gallolyticus* might confer an advantage for efficient colonization of the cardiac valves, thus contributing to the development of infective endocarditis (Danne *et al.*, 2011).

The adhesin subunit of the Pil3 *S. gallolyticus* pilus was shown to promote the binding to mucins and to fibrinogen, therefore contributing to the colonization of the intestinal site (Martins *et al.*, 2016). These results highlight the importance of different pili specifically expressed at different host sites by the same bacterial cell during the infectious process.

Pilus-associated streptococcal adhesins can sometimes also be found as individual surface proteins, such as the RrgA minor pilin in *S. pneumoniae*. RrgA is essential for pneumococcal adherence and disease even if pilus polymers are not produced (Nelson *et al.*, 2007), as it interacts with human fibronectin, collagen I and laminin, but this adhesion is more efficient in vivo when the complete pilus structure is produced (Hilleringmann *et al.*, 2008).

Biofilm formation

Biofilms are generally made up of different bacterial communities (one or several species) integrated in a protective adherent matrix (composed of DNA, proteins and extracellular polysaccharides) that they produce during biofilm formation. This biofilm matrix environment enhances bacterial resistance to antibiotics, and therefore poses an important hygienic and epidemiological problem (Beloin *et al.*, 2008).

The dental plaque represents the most important bacterial biofilm environment in the human body, which is composed of more than 500 bacterial species, with a majority of Gram-positive bacteria, such as *Actinomyces* species and oral streptococci, including *S. oralis*, *S. mitis*, *S. sanguinis* and *S. gordonii* (Jakubovics, Yassin and Rickard, 2014). The colonization of the dental surface occurs in two distinct steps. The bacteria listed above represent mainly the primary colonizers that initially bind to the teeth. The tooth surface is covered in a pellicle that consists of a variety of host-derived molecules that can be recognized by the surface molecules of these early colonizer bacteria. The primary dental biofilm enables the inclusion of other species in this structure, providing a protective environment for the growth of several, mainly Gram-negative species (Palmer *et al.*, 2003).

Pili, expressed by several Gram-positive oral bacteria, have been recently shown to be important surface factors in Gram-positive biofilm formation. The FimA type pili of *Actinomyces naeslundii*, for example, play and important role in the adhesion to the tooth enamel in the formation of dental plaque biofilms in coaggregation with oral streptococci, most often *S. oralis* and *S. sanguinis* (Mishra *et al.*, 2010). *S. sanguinis*, one of the most abundant dental plaque-forming species, also relies on its pili for bacterial adhesion and biofilm formation on saliva-coated surfaces, as the tip pilin binds α-amylase, the major component of whole saliva (Okahashi *et al.*, 2011).

Pili of *S. pyogenes* have also been shown to be implied in bacterial aggregation in liquid media, microcolony formation and adhesion to human pharyngeal tissues and mature biofilm formation (Manetti *et al.*, 2007). Interestingly, the contribution of *S. pyogenes* pili in biofilm formation is dependent not only on the pilus FCT type, but also the strain M-type and on growth conditions (Lembke *et al.*, 2006, Manetti *et al.*, 2010). The FCT-1 type pili of the M6 strain TW3558 has been shown to contribute to biofilm formation, but to hinder bacterial aggregation (Kimura *et al.*, 2011). Conversely, the FCT-3 type pilus of serotype M49 strain was not involved neither in aggregation nor in biofilm formation (Nakata *et al.*, 2009).

Involvement of pili in biofilm formation has been shown in several other Gram-positive species, such as *S. pneumoniae*, *S. agalactiae* (see later), *S. gallolyticus*, *A. naeslundi* and *A. oris*, and enterococci like *E. faecalis* and *E. faecium* (Danne and Dramsi, 2012). However, the molecular mechanisms underlying this process remain poorly understood.

Immune system evasion

As highly exposed surface appendages, pili have been shown to modulate the host immune responses, through interaction with specific host surface molecules, facilitating bacterial survival and dissemination.

Pneumococcal pili, for example, not only contribute to adhesion to host tissues, but also affect inflammatory responses in the host, triggering higher TNF-α and interleukin-6 pro-inflammatory responses than non-piliated strains (Barocchi *et al.*, 2006). The pilus-associated adhesin RrgA triggers increased phagocytotic uptake of pneumococci by macrophages in a complement receptor 3 (CR3) dependent manner in vitro, suggesting a specific interaction between these molecules. Similarly, in vivo, RrgA contributes to the dissemination of streptococci in both intra-nasal and intraperitoneal mouse infection models. RrgA expression also significantly increases intracellular streptococcal survival, confirming that RrgA-CR3 interaction mediated phagocytosis by macrophages is an important mechanism for pneumococcal systemic dissemination (Orrskog *et al.*, 2012).

RrgA is also recognized and bound by TLR-2 (toll-like receptor 2) on the surface of host leukocytes, especially in their oligomerized form. This interaction is specifically mediated by a surface-exposed 49-amino acid motif. TLR-2 binding also leads to the activation of TNF-α inflammatory responses in an in vivo mouse model (Basset *et al.*, 2013). RrgA is therefore a key component of the activation of host inflammatory response by interacting independently with both TLR-2 and CR-3 receptors.

Immunogenic properties

As long, protruding surface elements, pili are likely targets of the host immune system and antibody production. Immunogenicity of pili have already been described for several Gram-positive species, as even in non-invasive carriage, anti-pilin IgG antibodies can be isolated from the blood of the host (Ahmed *et al.*, 2014; Dangor *et al.*, 2015). This makes pili an attractive target antigen for vaccine production for active immunization against Gram-positive bacterial infections.

In the case of *S. pneumoniae*, both active and passive immunization with recombinant RrgA pilus adhesin leads to protection against a lethal challenge with the TIGR4 strain in mice (Gianfaldoni *et al.*, 2007). RrgA exists in two cladic variants, but this variation has no effect nor on RrgA adhesive properties nor on the recognition of the protein by the immune system, as antibodies against the two types offer cross-protection (Moschioni *et al.*, 2010). While some suggested using RrgA for vaccine production as anti-RrgA antibodies are not only highly protective, but also hinder bacterial adhesion to host tissues (Ahmed *et al.*, 2014; Amerighi *et al.*, 2016), the pilus islet 1 is only coded in about 25-30% of the clinical isolates, therefore such a vaccine could not provide universal protection against pneumococcal infections.

Different difficulties arise in the case of *Streptococcus pyogenes*, as its pili are not only highly diverse, but major subunits were also shown to contain variable extended structures

and loops. These structures are often targeted by host IgG production, but the diversity in size and structure of the subunits could make the production of a potential pilus-based vaccine against GAS infections quite challenging (Young *et al.*, 2014).

A similar approach against nosocomial catheter-associated infections by enterococci and especially *E. faecalis* in mice has been proven more successful. Catheter-associated biofilm formation by this pathogen is dependent on the EbpA pilus-associated adhesin, which binds to the fibrinogen layer that is deposited on the catheter surface after its introduction in the urinary tract. Antibodies against either whole EbpA or only its specific fibrinogen-binding domain prevented bacterial deposition and biofilm formation on the abiotic surface, preventing the development of enterococcal urinary tract infection (Flores-Mireles *et al.*, 2015). In this case, the low variability of the pili and the sequence of its subunits and the precisely localized nature of the infection are in favor of the development of a pilus-specific antibody-based prevention strategy. It also highlights the how the feasibility of a pilus-based vaccination approach can largely depend on the target species and the characteristics of the infection they are associated with.

Regulation of pilus expression

While pili represent important colonization factors in most Gram-positive bacteria, maintaining them also represents a considerable fitness cost, and in some cases, could hinder dissemination. Opportunistic pathogenic bacteria also face several different environments during the infectious process, representing a wide range of temperature, pH and oxidative conditions, that require constant adaptation. While at certain host sites, high expression of pili might be advantageous, at other sites, their presence may be deleterious. Therefore, a fine-tuned regulation of pilus-expression in response to environmental signals is necessary to optimize colonization, persistence and dissemination in the host. The modulation of pilus-expression can either be mediated by protein transcription regulators, or can occur due to other specific regulatory elements, such as specific sequence regions or RNA products.

Heterogenous pilus expression

Heterogeneous pilus expression within a bacterial population was first described in *C. renale* and *Corynebacterium pilosum*, where strongly piliated Pil^+ and non-piliated Pil^- cells were observed in the same culture (Ito *et al.*, 1987). Hiramune and colleagues postulated that the shifting between piliated and non-piliated bacteria occurs via phase variation in *C. pilosum* (Hiramune *et al.*, 1991).

A similar heterogeneous expression was observed for PI-1 pilus in *S. pneumonia*e. Two research groups simultaneously showed that this heterogeneity is a result of bistable PI-1 expression in the studied strains (Basset *et al.*, 2011, De Angelis *et al.*, 2011). Sequencing the PI-1 locus in the piliated and non-piliated cells of the same bacterial population revealed no genetic differences, therefore excluding the possibility of a phase-variation-based regulatory mechanism in this case. Instead, the bistable expression of PI-1 pili in *S. pneumoniae* was subsequently proven to be dependent on a positive-feedback loop controlling the expression the transcriptional activator of PI-1 expression, *rlrA* (Basset *et al.*, 2012).

Heterogenous expression of Pil1 pilus was also observed in *S. gallolyticus* (Danne *et al.*, 2014). Interestingly, the underlying molecular mechanism combines phase variation and transcriptional attenuation at the single cell level. Sequencing of the promoter region of hyper-piliated variants of *S. gallolyticus* strain UCN34 always revealed modifications of the number of GCAGA repeats in the promoter region compared to the UCN34 original sequence. Translation of a leader peptide made of 22 GCAGA repeats whose length can change (by addition/deletion of one GCAGA repeat) was shown to block the formation of a hairpin like structure leading to higher pilus expression (Danne *et al.*, 2014). A similar regulatory mechanism was suggested to be involved in the expression of Pil3 type pili of the same species, as a similar region of varying numbers of tandem repeats and an attenuator-like hairpin structure is also present in the promoter region of this locus (Martins *et al.*, 2016).

Response to environmental factors

One of the most important environmental signal is temperature, as it changes both among certain colonization sites within the host (eg. skin, outer genital surface, blood, etc.) and as a response to infection (fever). FCT-3 type pilus expression in *S. pyogenes* (M49), for example, increases from 20% to 47% when the temperature is lowered from 37°C to 30°C. As 30°C is closer to human skin temperature, this finding also suggests that pili might have a specific role in skin-colonization or infection, while their presence is less required in internal infections (Nakata *et al.*, 2009). Conversely, in *E. faecium* (E1165), PGC-1 and PGC-3 type pili were shown to have an optimum for polymerization at 37°C, while at 21°C, only monomers of the major pilins can be observed at the cell surface. This suggests that during prolonged environmental survival of *E. faecium*, pili are not expressed, but pilus biosynthesis is turned on once the bacteria contaminate the host or indwelling medical devices, most likely to promote adhesion and colonization (Hendrickx *et al.*, 2008).

Another important environment signal is local pH, in the case of *S. pyogenes*, the pH conditions at the relevant infection sites are pH 5-7 at the oral cavity, pH 6.4-6.9 in the nasopharynx and pH 4.2-5.9 on the skin surface. During its growth, *S. pyogenes* also acidifies

its environment as a result of fermentative sugar metabolism. This local acidification in vitro was responsible for increased biofilm formation on abiotic surfaces and human epithelial cells by FCT-types 2, 3, 5 and 6 and certain FCT-4 strains due to enhanced expression of pilus proteins (Manetti *et al.*, 2010). The regulation of the metabolic pathway leading to acid production and pilus expression activation under low pH conditions appear to be both under the control of the same YvqE regulator (Isaka *et al.*, 2016). Although the additive effects of intrinsically low pH at certain human sites and bacterial-metabolism generated environmental acidification are yet to be elucidates, environmental pH certainly affects the expression of most pilus types of *S. pyogenes*, in order to facilitate biofilm and microcolony formation at the right moment during infection.

Sometimes other environmental factors can also affect pilus expression, for example, upon presence of human serum in the growth medium, *E. faecalis* strain OG1RF expresses almost 4-times more Ebp-type pili, and cells of the same species recovered from a rat endocarditis site show even stronger piliation (Nallapareddy *et al.*, 2011). Expression of the same *ebp* locus also responds positively to increased bicarbonate concentration in the medium but not that of CO_2. Bicarbonate is present in most mammalian host environments, and due to its buffering capacities, it is mainly found at acidic sites where it contributes to the maintenance of the local pH homeostasis (Bourgogne *et al.*, 2010). The *S. pyogenes* pilus FCT-3 was shown to be upregulated under oxidative stress conditions, independently of PerR, the main GAS peroxide response regulator. This effect was shown to be mediated by the pilus-specific regulator MsmR, in a serotype-dependent manner (Grifantini *et al.*, 2011). These mechanisms most likely help the bacteria quickly adapt to the new niches encountered in the human body during its dissemination.

Specific and other transcriptional regulators

Most Gram-positive pilus loci contain a gene coding for at least one stand-alone transcriptional regulator, usually transcribed independently of the pilus operon (Danne and Dramsi, 2012) (marked in yellow on Figure 3). The regulatory proteins coded by these genes have been assigned mainly to two major protein families: the AtxA/Mga regulator superfamily and the AraC/XylS-type regulator family.

The first superfamily contains the EbpR regulator of the *E. faecalis ebp* pilus operon (Singh *et al.*, 2007) as well as the RALP (RofA-Like Protein) regulator family (Kreikemeyer *et al.*, 2002). The family is named after RofA, the first such regulator described in *S. pyogenes*. This species codes for three further regulatory proteins belonging to this family, that were shown to modulate the expression of numerous proteins such as hemolysins, MSCRAMMs, proteases and other virulence regulators in this bacterium, including pili (Kreikemeyer *et al.*, 2011). RofA itself is coded divergently upstream in the proximity of the FCT-1 and FCT-2 type

pilus operons (in M6 or M1 strains, respectively), and acts as a positive regulator of their transcription. Its binding to its target DNA occurs via a specific 17bp recognition site in the proximity of the regulated genes (Granok *et al.*, 2000), including in its own promoter region, as RofA expression is positively autoregulated. Although at first, RALPs have been described as exclusively positive regulators, later negative regulation of virulence genes by these factors have also been observed (Kreikemeyer *et al.*, 2002, Beckert *et al.*, 2001). In *S. pyogenes*, RofA is coded divergently upstream in the proximity of the FCT-1 and FCT-2 type pilus operons (in M6 or M1 strains, respectively), and acts as a positive regulator of their transcription.

Additional RALPs regulating pilus expression in Gram-positive bacteria include the EbpR proteins in *E. faecalis* and *E. faecium*, the RogB activator of PI-2a pilus expression of GBS NEM316 (as described later), and Nra, the second RALP identified in *S. pyogenes* as a regulator of the FCT-3 pilus locus. These regulatory factors are all coded upstream of the pilus genes and transcribed inversely (Figure 3) (Danne and Dramsi, 2012).

The expression of FCT-3 pili in *S. pyogenes* is, beside the upstream coded Nra, also regulated by a second protein factor, MsmR. This protein, however, belongs to the AraC/XylS-type regulator family, which is coded downstream of the FCT-3 locus and transcribed divergently. This regulator group was named after AraC, identified as an arabinose-response mediator, which contains characteristic DNA binding domains that can be identified in all AraC/XylS-type regulators. While most standalone pilus-specific regulators are activators, Nra and MsmR regulate pilus expression in opposite ways where the polarity of regulation depends on the serotype per the differences in the transcriptional circuitry (Luo *et al.*, 2008). As all AraC-type regulators, MsmR binds directly to the corresponding promoter regions, and regulate the transcription of not only the pilus core operon but also of Nra, that will have an adverse effect on pilus transcription (Kreikemeyer *et al.*, 2011).

A similar organization with one RALPs regulator coded upstream of the pilus locus and an AraC-type regulator found downstream can be observed for the PI-1 pilus locus of *S. pneumoniae* TIGR4. However, in this case, the RALP, RlrA is the major regulator, while MgrA coded by the downstream gene has milder repressive effects on pilus gene transcription (Hemsley *et al.*, 2003). The only AraC-type regulator of pilus transcription coded upstream of its target pilus operon identified to date is Ape1 in *Streptococcus agalactiae* NEM316, which modulates the expression of its target, the PI-1 pilus operon, as a part of a bigger regulatory network (Jiang *et al.*, 2012). Similarly, most pilus-associated transcriptional regulators do not function as isolated regulatory blocks, they are integrated and cross-linked in a regulatory network of virulence factors.

Beyond the examples described above, a few unique regulatory mechanisms have also been identified in certain species. For example, in *E. faecalis*, *ebp* pilus transcription is regulated by the major quorum-sensing system alongside with several other virulence factors in a coordinated manner. In the same species, a putative RNase coding gene was shown to affect the *ebp* operon transcription, suggesting that pilus expression can also be regulated at a prost-transcriptional level (DebRoy *et al.*, 2014). In *S. pyogenes* (M1), a small FasX RNA was revealed to be an important repressor of pilus gene transcription, as it binds to and destabilizes pilus operon mRNA and interferes with the translation of minor pilin Cpa (Liu *et al.*, 2012). It also positively regulates the expression of streptokinase, a secreted virulence factor of *S. pyogenes* (Kreikemeyer *et al.*, 2001). In one unique case, a pilus subunit has also been suggested to affect the expression of pilus genes: the RrgA accessory pilin in *S. pneumoniae* is a negative regulator of pilus expression (Basset *et al.*, 2011).

Although the most important regulators of bacterial virulence factor expression are two component systems (TCS) that consist of a receptor-kinase component and a pleiotropic effector molecule, the involvement of major TCSs in the regulation of pilus expression has so far only been described as indirect. For example, in *S. pneumoniae*, the effect of two different TCSs on pilus expression is achieved via the pilus-specific regulator RlrA that acts in a TCS-dependent manner (Rosch *et al.*, 2008).

This diversity of possible mechanisms for regulating pilus expression highlights the importance of a tight and adaptive control of the biosynthesis and display of these surface-exposed virulence factors. Understanding the regulatory mechanisms governing pilus expression and the adaptive processes they are involved in could also shed light on the function of these structures in the host colonization process and infection.

E. GBS pili

In GBS, two types of pilus islands have been identified through genomic analysis (Lauer *et al.*, 2005; Rosini *et al.*, 2006); which are composed of genes coding for partially homogeneous LPXTG proteins and sortase-coding genes (pilus islands 1, 2a, and 2b). PI-1 is coded in a pathogenicity island, while the PI-2 locus is part of the stable bacterial genome. The sequence of the pilin subunits corresponding to the same function are highly dissimilar among the different loci, suggesting specific roles for each pilus in GBS colonization and/or virulence. At least one of the three loci is present in all GBS clinical strains tested to date.

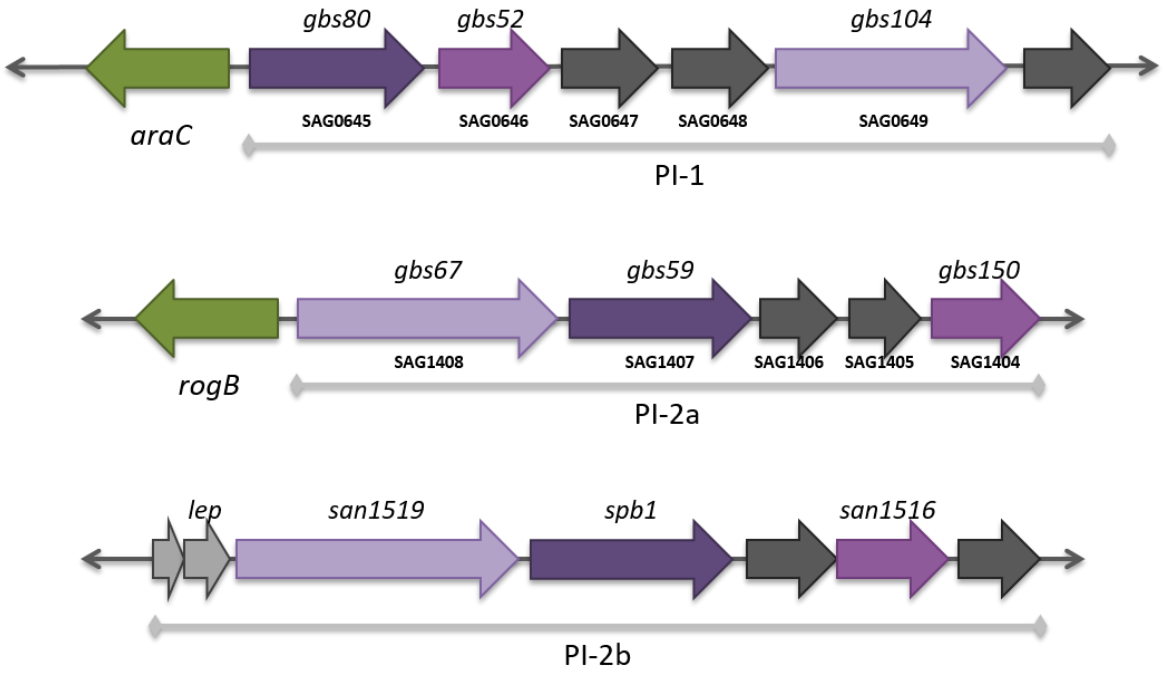


Figure 4 | Genomic organization of pilus loci in Streptococcus agalactiae. *Schematic representation of the three GBS pilus islands, with LPXTG-containing pilus subunits represented with purple arrows, and sortases coding genes in dark grey. Known transcriptional regulators are represented in green, while conserved genes of unknown function appear in light grey. For PI-1 and PI-2a gene names appear as in the assembled genome of GBS 2603V/R, for PI-2b genes as in the COH1 strain. Adapted from Rosini et al., 2006.*

Distribution of pilus loci in GBS

At least one pilus island has been shown to be present in all GBS strains studied so far (Martins *et al.*, 2013). PI-2a was found to be the most widespread locus, found in 79% of the studied strains, and PI-1 was expressed in 72% of known GBS isolates as of 2012 (Jiang *et al.*, 2012), whereas the frequency of PI-2b was considerably lower, found only in 21% of the isolates (Martins *et al.*, 2013). PI-1 can be found alongside either of the PI-2 variants in GBS strains, however, the presence of PI-2a and PI-2b is mutually exclusive. To date, strains

containing both PI-1 and PI-2a were found to be the most frequent, although this ratio could shift with the increasing numbers of published GBS (draft) genomes.

While most human-derived strains contained PI-1, it was never detected in bovine GBS isolates, which usually possessed a unique PI-2b allele. Pathogenic strains isolated from neonatal infections more frequently had both PI-1 and a PI-2 variant than maternal colonizing strains, latter characteristically being the PI-2b allele in the hypervirulent CC-17 clones (Springman *et al.*, 2014). PI-1 is found at a variable part of the GBS genome (Lauer *et al.*, 2005), flanked by direct repeats, and in the strains, that do not code for this locus, the two conserved genes flanking the PI-1 island are contiguous, indicating that this island was most likely acquired by horizontal gene transfer (Telford *et al.*, 2006). Loss of this PI-1 locus in pathogenic ST-17 isolates has been observed at an increasing frequency, and the pilus island is most often replaced by antibiotic resistance genes (Campisi *et al.*, 2016), which also suggests that this locus is not essential for GBS pathogenicity in humans.

Beyond the distribution of pilus loci among GBS isolates, other additional conditions, such as mutations in regulatory factors can affect which loci are expressed in which strain. For instance, the serotype III strain NEM316 has loci coding for both PI-1 and PI-2a, but as a result of a frameshift mutation in the *araC* (*sag0644*) regulator gene, the PI-1 locus is not expressed in this strain, whereas in the GBS strain JM9130013, the gene coding for the RogB regulator of PI-2a contains a frameshift, resulting in only PI-1 being expressed, although this strain also carries both pilus islands (Dramsi *et al.*, 2006). The serotype V strain, 2603V/R also harbors inactivating mutations in RogB, and produces no detectable amounts of PI-2a, and was thus extensively used in studies targeting the function and regulation of the PI-1 locus.

PI-1

The PI-1 locus, found at a variable pathogenicity island in the GBS genome, contains three structural protein genes (*gbs80*, *gbs52* and *gbs104*) and two sortase genes of partially redundant functions (Figure 4). The two sortases SAG0647 and SAG0648, encoded in the locus are specific for only one of the accessory pilins, while the major pilin, GBS80 (SAG0645) can be polymerized by either, the formation of the pilus backbone is thus not dependent on only one sortase (Rosini *et al.*, 2006). Two PI-1 components, the major pilin GBS80 and the putative adhesin accessory pilin GBS52 have been partially crystallized and their structural organization was revealed (Vengadesan *et al.*, 2010; Krishnan *et al.*, 2007). Interestingly, the GBS52 subunit also contains a pilin-motif-like sequence which might facilitate its incorporation in the pilus structure (Krishnan *et al.*, 2007).

PI-1 functions

In a first study, recombinant AP1-1 accessory pilin subunit of PI-1 was shown to bind to cells of different human lineages, including Caco2 colorectal, A549 lung epithelial and ME180 cervical cells. Deletion of AP1-1 but not that of BP-1 was also shown to reduce the adhesion of the COH1 GBS serotype III strain to ME180 cells by 35-40% (Pezzicoli *et al.*, 2008). Subsequent studies, however, using the 2603V/R serotype V strain, failed to reproduce these results, using the same target cell lines, with the addition of VK2 vaginal cells (Jiang *et al.*, 2012). This apparent conflict might have occurred because the initial adhesion level of the wild-type COH1 and 2603V/R cells to the studied cell lines is already below 5% in both studies, in which a 40% change is most likely of very low relevance.

The GBS52 adhesin subunit, that contains two Cna domains was suggested to be involved in bacterial adherence to A549 lung epithelial cells via its second Can domain (Krishnan *et al.*, 2007). The pilus has also been shown to mediate adhesion to immobilized salivary glycoproteins in three different GBS strains (Brittan and Nobbs, 2015). Additionally, Jiang and colleagues also reported that PI-1 inhibits uptake and intracellular killing of GBS by monocyte-derived macrophages (but not by neutrophils) (Jiang *et al.*, 2012).

To date, most studies addressing roles of PI-1, however, have only been carried out in vitro, using either purified substrates or human cell lines, and were carried out in a limited set of strains. The GBS2603V/R strain expresses only the PI-1 locus, while COH1 expresses a second pilus, PI-2b, which might result in different specialized roles for the same locus in these two isolates, that are also associated to significantly different clinical phenotypes. The role in infections caused by ST-17 strains also questionable, as strains belonging to this clone usually express variably low levels of PI-1, and the locus is frequently lost in these strains.

Regulation of PI-1

Based on studies conducted with the GBS strain 2603V/R that only expresses PI-1 due to a frameshift mutation in the PI-2a expression regulator gene *rogB*, three factors regulating PI-1 expression have been identified: i) Ape1, an AraC-family type regulator, ii) the CovR/S two-component system, and iii) environmental pH.

Ape1 (for Activator of Pilus Expression 1) is coded by the *sag0644* gene in the GBS strain NEM316, directly upstream of the PI-1 and transcribed divergently, as it is often observed for pilus-specific regulators (Figure 4). The C-terminal part of Ape1 is composed of two AraC-like DNA binding domains, while its N-terminal serves as a dimerization or ligand-binding domain (Dramsi *et al*., 2006). It also shows 72% amino acid sequence identity to MsmR, an AraC/XylS type pilus-specific positive transcriptional regulator of *S. pyogenes* (Nakata, Podbielski and Kreikemeyer, 2005).

The CovR/S two-component system effector CovR was shown to repress the expression of all PI-1 components at the transcriptional level, which effect is partially, but not entirely mediated by CovS (Jiang *et al.*, 2012). Ape1 expression is also under the control of the CovR/S system, therefore its effect on PI-1 expression is also mediated by this system. In a mutant deleted for *covR*, *ape1* transcription is de-repressed and the increased amount of Ape1 regulator can activate transcription of the PI-1 locus. In the GBS wild type context, the two mechanisms most likely play nonredundant roles in regulating and fine-tuning PI-1 expression with CovR acting as a direct repressor even in the absence of Ape1, as both CovR and Ape1 can bind to the PI-1 promoter DNA (Jiang *et al.*, 2012). To date, PI-1 is the only streptococcal pilus whose expression is tightly regulated by the CovR/S two-component system.

Regulation by PI-1 is also modulated by environmental conditions, especially different levels of pH. Higher PI-1 gene transcription was observed at an acidic pH of 5.0 than at neutral pH (Santi *et al.*, 2009; Jiang *et al.*, 2012). Although a first study attributes this effect to higher Ape1 transcription at lower pH (Santi *et al.*, 2009), in the more recent study, increase in PI-1 expression at lower pH was also observed in *ape1* and *covR* mutant strains, indicating that the effect of pH on PI-1 expression is at least partially independent of these two regulators (Jiang *et al.*, 2012).

PI-2a

PI-2a is the more widespread allele of the PI-2 locus, encoding for 3 structural pilus components and 2 sortase genes (Figure 4) (Dramsi *et al.*, 2006; Rosini *et al.*, 2006). Recently, it was shown that PI-2a pili from GBS strain NEM316 were recognized by sialic-acid specific lectins, and it was proposed that the PilA adhesin subunit located at the tip of the pilus was sialylated (Morello et al., 2015). However, the sugar-modified amino acid(s) were not identified by Mass-Spectrometry. Two putative N-glycosylation sites were deduced from *in silico* analysis of PilA protein sequence. Mutagenesis of these two asparagine into glutamine indicate a potential role of the glycosylation in PilA stability (Morello *et al.*, 2015).

1- Functions of PI-2a

1-1 Adhesion to host cells and ECM components

Dramsi and colleagues first showed that deletion of the accessory pilin of the PI-2a locus, AP1-2a or PilA, reduces the adhesion of the NEM316 GBS strain to A549 lung epithelial cells (Dramsi *et al.*, 2006). This interaction is mediated mainly by the adhesin subunit, via a 353-amino-acid von Willebrand factor type A (VWA) domain, but the pilus shaft formed by BP-2a is necessary for efficient surface display of AP1-2a beyond the capsule (Konto-Ghiorghi *et al.*,

2009). Under flow-conditions that better mimic physiological conditions, pilus integrity becomes critical for GBS adhesion to lung epithelial cells, confirming the importance of intact pili in infection. This result also sheds light on the limitations of conventional static models for adhesion studies (Konto-Ghiorghi *et al.*, 2009). Interestingly, it was also revealed that during adhesion to A549 cells, AP1-2a can also act as a single SrtA-sorted LPxTG surface adhesin, independently of its role in pilus formation (Dramsi *et al.*, 2006). PI-2a was also shown to be involved in attachment of the GBS strain NCTC10/84 to vaginal epithelium using another GBS strain in a human female cervico-vaginal tract model as well (Sheen *et al.*; 2011).

It was proposed that PilA adhesin binds to collagen type I, promoting bacterial interaction with the $\alpha_2\beta_1$ integrin resulting in activation of host chemokine expression and neutrophil recruitment during infection, which is a critical step in GBS infection of mice (Banerjee *et al.*; 2011). However, these results were challenged by our group that did not reproduce the binding of PilA nor PI-2a pilus to collagen type I (Dramsi *et al.*, 2012a). Several clinical and contemporary isolates of GBS were tested showing significant binding of GBS to fibrinogen but not to collagen type I (Dramsi *et al.*, 2012a).

1-2 Biofilm formation

PI-2a also mediates biofilm formation in GBS, on both biotic and abiotic surfaces. Interestingly, the AP1-2a adhesin, was shown to be dispensable for biofilm formation, while the anchoring pilin, AP2-2a, has a key role in the process. Deletion of the major pilin abolishes biofilm formation; indicating that properly polymerized and correctly anchored pilus structures are necessary for biofilm forming (Konto-Ghiorghi *et al.*, 2009; Rinaudo *et al.*, 2010).

1-3 Interaction with the immune system

In vivo experiments in mice indicated that the major PilB pilin of GBS strain NCTC10/84 serotype V provide protection against killing by macrophages and neutrophils (Maisey *et al.*, 2008). Surprisingly, overexpression of the major pilin PilB in *L. lactis* conferred virulence to this otherwise non-pathogenic species in a mouse model (Maisey *et al.*, 2008). Furthermore, it was shown that PilB contributes to resistance to cationic antimicrobial peptides. Although PilB of NCTC10/84 and PilB of GBS strain NEM316 are highly similar, our laboratory did not reproduce these results. No role of PilB could be demonstrated in macrophages nor in

antimicrobial resistance (Papasergi *et al.*, 2011). In conclusion, the role of PilB in innate immunity is still an open question.

2- PI-2a regulation

The regulatory gene located immediately upstream of PI-2a locus is *rogB* (Figure 4), a member of the RofA-like regulator protein family (RALPs). RogB was first shown to activate the transcription of *gbs1478* (*pilA*) and *gbs1477 (pilB)* encoding PI-2a adhesin and major pilin respectively, as well as other virulence genes in the 6313 strain (Gutekunst *et al.*, 2003). Later the role of RogB as an activator of PI-2a expression was also shown in GBS strain NEM316 (Dramsi *et al.*, 2006). The major two component system CovR/S system, which regulates PI-1 expression, is not involved in PI-2a regulation (Samen *et al.*, 2011). Interestingly, RogB is under the positive regulation of Rga, another member of the RofA-like protein (RALP). Rga is an important regulator of several GBS virulence factors, as it is necessary for the optimal expression of not only PI-2a proteins but also Srr1 and FbsA, and its deletion leads to impaired virulence (Samen *et al.*, 2011). Indeed, Rga is the master regulator of PI-2a expression and RogB playing a minor role (Dramsi *et al.*, 2012b)

PI-2b

The PI-2b locus in GBS is the second allelic variant of the type 2 pilus. Its organization is very similar to PI-2a and PI-1, with three pilus subunit-coding genes: the major pilin *spb1* (*san1518*), and two accessory pilins, the putative adhesin *san1519* or *ap1-2b* and the anchoring subunit *san1516* or *ap2-2b* (Figure 4) (Lazzarin *et al.*, 2015). Unlike in PI-1 and PI-2a, two additional overlapping genes are found at the beginning of the PI-2b operon; the first gene (*orf*) of unknown function and the second gene (*lep*) encoding a putative LepA-type signal peptidase.

PI-2b distribution is restricted to hypervirulent ST-17 human clinical isolates of GBS serotype III and a few other GBS strains of serotype Ia such as A909. It is also the only pilus locus ubiquitously conserved in bovine strains (Springman *et al.*, 2014).

Prior to the characterization of PI-2b as a pilus-coding locus in GBS (Rosini *et al.*, 2006), Spb1, the PI-2b backbone subunit was identified as a surface invasin (hence named Surface Protein of GBS 1, Spb1), and was shown to promote invasion of human epithelial cells *in vitro* by GBS strain COH1 (Adderson *et al.*, 2003).

Distinct features of PI-2b pilus

A recent study clearly identified the role of the accessory pilin Ap2-2b (San1516) as the anchoring subunit covalently attached to the cell wall. The contribution of the two sortases encoded in this locus was also determined. Interestingly, only SrtC1 was shown necessary for pilus assembly, mediating the incorporation of both the major subunit Spb1 and the putative adhesin Ap1-2b into the pilus shaft (Lazzarin *et al.*, 2015). While SrtC1 belongs to class C sortases, the second sortase of the locus, Srt2, was clearly different and shown to be involved in pilus attachment to the cell wall (Lazzarin *et al.*, 2015).

In addition to the critical cysteine residue for sortase function, Srt2 contains two additional cysteines. Substitution experiments of these two other cysteines revealed that both play a role in Srt2 activity, $Srt2_{C192A}$ showed increased, while $Srt2_{C115A}$ reduced LPXTG-processing activity compared to the WT protein. It was proposed that due to the presence of multiple cysteine residues, intramolecular disulfide bonds can form under oxidative conditions, which result in the reversible inactivation of the enzyme. It was thus tempting to hypothesize that a redox switch could regulate Srt2 enzymatic activity, in contrast with the lid conformation change mechanisms in C-type sortases (Lazzarin *et al.*, 2015). These results imply that there could be a possible diminution in PI-2b surface display in response to highly oxidative conditions encountered in the host.

Crystal structure of over a total of 502 amino acids of the major pilin Spb1 (lacking the first N-terminal 184 amino acids) has been published recently. Two domains with an IgG-like fold organization stabilized by internal isopeptide bonds were found, and the residues important for Spb1 polymerization were identified (Cozzi *et al.*, 2015).

PI-2b function

To date, only a few papers has addressed the biological role of PI-2b pili in GBS. As mentioned above, Spb1 was first identified as an invasin in cultured epithelial cells (Adderson *et al.*, 2003). The same group showed that deletion of *spb1* in an ST-17 clinical isolate led to decreased adhesion and invasion to several epithelial cell lines (human pulmonary A549, human colonic C2Bbe1, and cervical HeLa cells) (Adderson *et al.*, 2003).

As in neonatal infections an important line of host-defense is opsonin-independent phagocytosis by activated macrophages, Chattopadhyay and colleagues compared the phagocytosis of a broad set of GBS clinical isolates by J774A.1 murine macrophages. As reference, they used the PI-2b expressing wild type GBS 874391 strain (serotype III), and an isogenic *spb1*-deficient mutant. They found that GBS strains lacking *spb1* gene were phagocytosed less efficiently, as compared to *spb1* expressing strains (Chattopadhyay *et al.*, 2011). However, the differences reported are small and highly variable.

In the same study, Spb1 was shown to enhance phagocytosis of GBS by both murine and human macrophages but not to contribute to intracellular survival (Chattopadhyay *et al.*, 2011).

The role of Spb1 in adhesion to host epithelial cells was also suggested using pilin-specific antisera on a broad set of GBS isolates of different serotypes. Surprisingly, for certain serotype VII strains, Spb1-specific antiserum inhibited adherence and invasion of A549 cells. They also reported a significant decrease in adherence to A549 cells for *spb1* deletion mutants of an undisclosed strain (Sharma *et al.*, 2013). Solid conclusions, however, are difficult to draw from this study especially because bacterial strain characterization at the genotype level is lacking.

Goals of the Project

Streptococcus agalactiae (GBS) is a major cause of late-onset neonatal infections, and its ST-17 clonal complex in particular is responsible for the majority of late-onset neonatal meningitis cases. While the frequency of early onset GBS-related neonatal infections decreased after the introduction of an antibiotic prophylaxis protocol, the rate of late-onset GBS-cases remained unaffected. Genomic comparison of ST-17 strains to non-ST-17 GBS isolates revealed several surface proteins that are characteristic to the ST-17 clone, among which HvgA (Tazi *et al.*, 2010) and Srr2 (Six A *et al.*, 2015) have been shown to be important for the hypervirulent phenotype. The PI-2b type pilus is also characteristic to ST-17 strains, but can also be found in a few non-ST-17 human isolates and bovine strains. Pili have been shown to contribute to adhesion to host cells in several Gram-positive bacteria (Danne and Dramsi, 2012) and therefore are generally regarded as important colonization factors.

The goal of this PhD project was to understand the role of this PI-2b pilus in the hypervirulent phenotype of ST-17 strains, mainly by studying in detail its regulation and biosynthesis. In the first part of the results, we compared the expression of the PI-2b pilus in our ST-17 representative strain BM110, and a non-ST-17 human clinical isolate, A909. Since no regulatory gene has been identified in the vicinity of the locus, we aimed to identify potential ST-17-specific regulatory elements and factors (Results Part I, publication).

As the role of several PI-2b genes has not been studied in detail, in the second part of this work, we sought to understand the role of the putative adhesin AP1-2b, and the two accessory genes *lep* and *orf* in the biosynthesis of PI-2b pilus (Results Part II).

Results

Part 1. - Regulation of PI-2b pilus expression

Publication: Regulation of PI-2b pilus expression in hypervirulent Streptococcus agalactiae ST-17 BM110

Noémi Szili, Bruno Périchon, Laurence du Merle, Isabelle Chupin-Rosinski, Myriam Gominet, Claire Poyart, Patrick Trieu-Cuot, and Shaynoor Dramsi

To investigate potential ST-17-specific regulators of the PI-2b locus, we compared its expression in our representative ST-17 human isolate BM110, and the non-ST-17 human strain A909.

Altogether, our results indicate that in BM110, and possibly other ST-17 strains, PI-2b expression is strongly regulated by the interplay of several factors. The resulting relatively low level of pilus expression might confer a selective advantage to these strains in certain human infections by facilitating systemic dissemination and/or evasion of the host immune response.

Regulation of PI-2b pilus expression in hypervirulent Streptococcus agalactiae ST-17 BM110

Bruno Périchon[1,2Π], Noémi Szili[1,2Π], Laurence du Merle[1,2], Isabelle Chupin-Rosinski[3], Myriam Gominet[1,2], Claire Poyart[4, 5, 6, 7, 8], Patrick Trieu-Cuot[1,2], and Shaynoor Dramsi[1,2*].

[1] Institut Pasteur, Biology of Gram-positive Pathogens Unit, Paris, France

[2] Centre National de la Recherche Scientifique (CNRS) ERL3526

[3] Institut Pasteur, Unité Ecologie et Evolution de la Résistance aux Antibiotiques

[4] Équipe Barrières et Pathogènes, Institut Cochin, Inserm 1016, CNRS UMR 8104

[5] Université Paris Descartes, Paris, France

[6] Assistance Publique-Hôpitaux de Paris, Hôpitaux Universitaires Paris Centre-Site Cochin,

[7] CNR Streptocoques, Hôpitaux Universitaires Paris Centre, site Cochin, AP-HP

[8] DHU Risques et Grossesse, AP-HP

[Π] These authors have contributed equally to this work

* corresponding author

Keywords: PI-2b pilus, S. agalactiae, GBS, Group B Streptococcus, regulation

Summary

The widely spread *Streptococcus agalactiae* (also known as Group B Streptococcus, GBS) "hypervirulent" ST-17 clone is strongly associated with neonatal meningitis. The PI-2b locus is mainly found in ST-17 strains but is also present in a few non-ST-17 human isolates such as the ST-7 prototype strain A909. Here, we analyzed the expression of the PI-2b pilus in the ST-17 strain BM110 as compared to the non-ST-17 A909. Comparative genome analyses revealed the presence of a 43-base pair (bp) hairpin-like structure in the upstream region of PI-2b operon in all ST-17 genomes, which was absent in the 8 non-ST-17 strains carrying the PI-2b locus. Deletion of this 43-bp sequence in strain BM110 resulted in a 3- to 5-fold increased transcription of PI-2b. Characterization of PI-2b promoter region in A909 and BM110 strains was carried out by RNAseq, primer extension, qRT-PCR and transcriptional fusions with *gfp* as reporter gene. Our results indicate the presence of a single promoter (Ppi2b) with a transcriptional start site (TSS) mapped 37 bases upstream of the start codon of the first PI-2b gene. The large operon of 16 genes located upstream of PI-2b codes for the group B carbohydrate (also known as antigen B), a major constituent of the bacterial cell wall. We showed that the hairpin sequence is a transcriptional terminator located between antigen B and PI-2b operons. In A909, increased expression of PI-2b probably results from read-through transcription from antigen B operon. In addition, we showed that an extended 5' promoter region is required for maximal transcription of *gfp* as a reporter gene in *S.*

agalactiae from Ppi2b promoter. Gene reporter assays performed in *Lactococcus lactis* strain NZ9000, a related non-pathogenic gram-positive species, revealed that GBS-specific regulatory factors are required to drive PI-2b transcription. PI-2b expression is up-regulated in the BM110*ΔcovR* mutant as compared to the parental BM110 strain but this effect is probably indirect. Collectively, our results indicate that PI-2b expression is regulated in GBS ST-17 strains, which may confer a selective advantage in the human host either by reducing host immune responses and/or increasing their dissemination potential.

Introduction

Group B *Streptococcus* (GBS; also known as *Streptococcus agalactiae*) is the leading cause of severe invasive neonatal infections worldwide. Clinical manifestations include pneumonia, septicemia, and meningitis occurring immediately after birth (0-6 days) referred as Early Onset Disease (EOD) or after the 1st week of life (7-90 days) referred as Late Onset Disease (LOD). GBS is a commensal gram-positive bacterium commonly found in the gastrointestinal and genital tracts of healthy individuals. For EOD, transmission to newborns probably originates from colonized mothers, by inhalation of GBS-contaminated amniotic or vaginal fluid during delivery. For LOD, the mode of transmission and the infection route remain poorly defined.

Several epidemiological studies have pinpointed a remarkable association of serotype III sequence type (ST) 17 GBS strains with meningitis, particularly during LOD [1-3]. These strains belonging to the clonal complex 17 (CC17) have historically been designated as “hypervirulent”. Deciphering the molecular bases of their higher pathogenicity constitute an important focus of our research.

Notably, two specific surface adhesins known as HvgA and Srr2 were shown to enhance the capacity of ST-17 strains to cross the blood-brain barrier [4,5]. In addition, two pilus islands known as PI-1 and PI-2b are found in the majority of ST-17 strains [6-8]. Interestingly, a recent analysis of CC17 strains phylogeny using whole genome comparison revealed the loss of pilus island 1 in about 15 % of ST-17 strains while PI-2b is present in 100% of these strains [9]. Although ubiquitous in the human ST-17 lineage, the PI-2b pilus is also found in a few non ST-17 human isolates and in most bovine strains [10]. It encodes three pilin subunits (AP1, Spb1, AP2) and two sortases (SrtC1 and Srt2) whose functions have been deciphered recently [11]. Two additional genes, *orf* and *lep*, are present at the beginning of the PI-2b operon encoding a conserved hypothetical protein and a putative signal peptidase, respectively. The backbone pilin Spb1 whose crystal structure has been solved recently [12] was proposed to enhance phagocytosis and survival in macrophages [13].

In this work, we carried out a detailed study of PI-2b expression in the ST-17 strain BM110 and in the non ST-17 strain A909. We observed a very low expression of PI-2b in BM110 as compared to A909 under laboratory conditions. Comparative genomics indicate the presence of a hairpin structure in the intergenic region upstream from PI-2b in all ST-17 strains, which is missing in non ST-17 strains. Deletion of this 43-bp sequence in strain BM110 increased PI-2b transcription by 3- to 5-fold. Lastly, our data suggest a complex regulation of PI-2b expression, being indirectly mediated by CovR and other GBS specific regulatory factors.

Results

Expression of major pilin Spb1 is lower in BM110 as compared to A909

We previously characterized the PI-2a locus encoding a heterotrimeric pilus in the ST-23 strain NEM316 [14,15]. Like its allelic counterpart, the PI-2b pilus carries 3 genes coding for the structural subunits (*spb1*, *ap1* and *ap2*), 2 genes (*srtC1* and *srt2*) expressing Srt enzymes, and two additional genes (*orf* and *lep*) encoding a small conserved gene of unknown function and a putative signal peptidase, respectively (Fig. 1A).

Surface distribution of the major PI-2b pilin Spb1 in strains A909 and BM110 was analysed by immunofluorescence. Spb1 level on the bacterial surface was higher in A909 than in BM110. As shown for other cell-wall anchored proteins, Spb1 preferentially accumulates at the cell poles in A909, whereas in BM110 only 2 to 4 distinct spots per bacterium were observed (Fig. 1B). As expected, Spb1 protein was undetectable on the surface of A909*Δspb1* and BM110*Δspb1* (Fig. 1B). Spb1 expression at the single cell level was further characterized using flow-cytometry (Fig. 1C). As shown in Fig. 1C, Spb1 is expressed homogenously in both strains, and Spb1 level is about 3 to 5 times higher in A909 than in BM110.

Finally, expression of PI-2b was analysed at RNA level by quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) using primers Spb1/Spb2 specific of *spb1* gene (Table S2). As negative controls, we used the mutants BM110*Δspb1* and A909*Δspb1* and the housekeeping *gyrA* gene for normalization. Transcription of *spb1* was about 6-fold higher in A909 as compared to BM110 (Fig. 1D).

Collectively, these results demonstrated that expression of Spb1 was higher in A909 than in BM110. This 4- to 6-fold difference in the levels of transcription and surface protein expression suggests a difference in the regulation of PI-2b expression in these two strains.

Genomic comparison of BM110 and A909 PI-2b locus

Alignment of the PI-2b locus sequence and its immediate genomic environment in BM110 and A909 revealed a strong conservation (Fig. 2A). Except for *ap1* encoding the minor pilin of unknown function, the other genes of the PI-2b locus are highly conserved encoding proteins displaying from 98.5 to 100 % identity. According to the published A909 genome sequence, AP1 exhibits a premature STOP codon at amino acid 280, leading probably to a non-functional protein. However, re-sequencing of the gene *ap1* (*sak1441*) in strain A909 revealed a missing base at position 839 which suppresses the premature STOP codon (our unpublished data). As opposed to the PI-1 and PI-2a pili characterized in GBS, no regulatory gene was found in the vicinity of PI-2b locus (Fig. S1). Immediately upstream from the PI-2b operon lies another large operon of 16 genes (*sak1438-1460* in A909) proposed to encode the group B carbohydrate also known as antigen B (AgB) [16].

Comparison of the intergenic region, located between the last gene of the AgB cluster (*sak1445*) and the first gene of the PI-2b locus (*sak1444*), revealed 84% identity between A909 and BM110. The main difference was the presence of a 43-bp sequence in BM110 located at 191 bp upstream from the *orf-lep* segment that was entirely missing in A909 (Fig. 2A). This element contains a 17-bp inverted repeat forming a hairpin-like structure which is predicted to be a transcriptional terminator (ΔG= -12.8 kcal/mol) by the ARNold finding terminators web server (http://rna.igmors.u-psud.fr/toolbox/). A similar element of 46-bp (ΔG= -20.04 kcal/mol) was identified as the transcriptional terminator of antigen B in the allelic PI-2a locus of strain NEM316 (Fig. S1).

Analysis of PI-2b intergenic region in 26 ST-17 and 8 non-ST-17 strains showed that all available ST-17 strains possessed this 43-bp sequence, which was absent in all non-ST-17 strains.

Deletion of the 43-bp hairpin increases BP expression in BM110

To test the role of this 43-bp sequence in PI-2b expression, a precise deletion of this element was constructed on the chromosome of WT BM110 strain by homologous recombination in a two-step allelic replacement process. This procedure allowed us to select at the same time the desired mutant BM110*Δ43* or an isogenic strain with a wild-type sequence, referred to hereafter as the BM110bWT control strain. In flow-cytometry experiment, the BM110*Δ43* mutant showed a 3- to 4-fold increase in Spb1 surface level when compared to the parental BM110 WT (Fig. 2B). Western blotting of cell-wall extracts from stationary cultures demonstrated the increased level of pilus polymers in the mutant BM110*Δ43* as compared to the BM110 WT or the bWT using specific polyclonal antibodies

against the major pilin Spb1. Control strains deleted for *spb1* and for *ap1* were used to show the specificity of the antibodies (Fig. S2).

This difference in protein expression results from a 3- to 4-fold increase in transcription levels of *spb1* and other PI-2b genes (*orf*, *lep*, and *ap1*; data not shown) in the *Δ43* mutant, as revealed by qRT-PCR analysis using specific primers (Fig. 2C). The wild type revertant strain (bWT) behaves exactly as the parental wild type BM110.

Evidence for PI-2b read-through transcription from antigen B operon in GBS A909

RT-PCR analyses revealed the presence of a read-through transcript encompassing the last gene (*sak1445*) of AgB operon and the first gene (*orf*) of PI-2b locus in strain A909 but not in strain BM110 (Fig. 3). This correlates with the absence of a predicted rho-independent terminator sequence in strain A909.

Mapping of transcription start sites of the PI-2b operon

TSSs were identified at the genome level in GBS strains A909 and BM110 using a differential RNA-seq approach as previously described for strain NEM316 [17]. A common TSS was identified 37 bp upstream of the start codon of the first gene (*orf*) of the PI-2b locus in both strains (Fig. 4A). In addition, increased transcription was detected in the intergenic region lying between the AgB and PI-2b operons in strain A909, but not in strain BM110 (Fig. 4A), which further suggests read-through transcription in strain A909. In contrast a sharp decrease in coverage is observed in strain BM110 upstream from the 43-bp sequence and is in agreement with the proposed role of this hairpin as transcriptional terminator of AgB operon.

Specific mapping of PI-2b TSS in A909 and BM110 was also performed by primer extension analysis. A single signal located 37 bp upstream from the translation initiation codon of *orf* was detected in both strains (Fig. 4B). The sequence lying upstream of this TSS displayed a canonical extended -10 (TGATATAAT) and semi-canonical -35 (ATGAGT) boxes separated by 16 bp (Fig. 4C).

A large upstream region is required for maximal transcription of PI-2b in GBS

To map precisely the Ppi2b promoter region, fragments encompassing different domains of A909 PI-2b intergenic region (Fig. 5A) were cloned upstream from the reporter gene encoding *gfp* using the vector pTCV-GFP, a low-copy number plasmid and assayed for GFP

expression in the non-PI-2b GBS strain NEM316 (ST-23). The various constructs are drawn schematically on Fig. 5A. Plasmid pTCV4 was used as negative control. Quantification of GFP fluorescence was performed using flow-cytometry in exponentially growing bacteria in TH broth supplemented with erythromycin for plasmid maintenance.

A detectable GFP expression was seen with a 125-bp fragment containing the promoter Ppi2b and about 60 bp upstream from the TSS (pTCV1). Higher level of GFP expression was observed with a larger 406-bp fragment including the whole intergenic region of Ppi2b (pTCV3) (Fig. 5B). No detectable GFP expression was seen with a fragment encompassing the region upstream without Ppi2b (pTCV2) which indicated that this segment did not contain another uncharacterized promoter.

GBS specific factors are required for maximal PI-2b transcription

To determine if transcription driven from PI-2b promoter region was dependent on GBS-specific regulatory factors, the various transcriptional fusions (pTCV1 to 4) were introduced into *Lactococcus lactis* strain NZ9000, a non-pathogenic gram-positive species belonging to the *Streptococcaceae* family. As shown in Fig. 4B, pTCV1 and pTCV2 gave a signal similar to the negative control. A very weak GFP fluorescence activity was measured with the largest promoter fragment (pTCV3). This result suggests that GBS-specific regulatory factors lacking in *L. lactis* are required for maximal transcription of PI-2b locus.

CovR was previously shown to be the master regulator of virulence genes in GBS [18,19]. CovR binds to target DNA sequences when phosphorylated at amino acid D53 by cognate histidine kinase CovS [20]. To assess the contribution of CovR in PI-2b expression, we introduced our *gfp* plasmid constructs in NEM316*ΔcovR* and CovRD53A mutants (Fig. 5C). A 2- to 3-fold increase in GFP expression driven from the Ppi2b promoter was consistently observed in NEM316*ΔcovR* as compared to WT NEM316 (Fig. 5C). However, this effect was not apparent in CovRD53A mutant (Fig. 5C). To test the role of *covR* directly in strain BM110, qRT-PCR analyses comparing mRNA levels of PI-2b first gene (*orf*) and AgB last gene (*san1522*) were carried out in BM110 WT, BM110*ΔcovR* and BM110CovRD53A. As shown in Fig. 5C, a 2- to 3- fold increase of PI-2b*orf* was seen in the *ΔcovR*, but not in CovRD53A mutant, as compared to the WT BM110. In contrast, the levels of *san1522*, encoding the last gene of antigen B, appeared unchanged in BM110 WT and *covR* mutants.

Since CovR was shown to directly repress PI-1 pilus transcription in GBS strain 2603V/R [21], we examined PI-1 mRNA levels in BM110 WT, BM110*ΔcovR* and BM110CovRD53A by qRT-PCR. In agreement with Jiang's results, we showed that transcription of the gene *san0698* encoding the PI-1 major pilin in BM110 increased 4-fold in the *ΔcovR* strain and 7-

fold in the CovRD53A mutant (Fig. S4). Collectively, these results demonstrate a direct role of CovR in PI-1 regulation in GBS strain BM110 and suggest an indirect role of CovR in PI-2b expression.

Discussion

Pili are important virulence factors for many gram-positive bacteria. In *S. agalactiae*, three different pilus loci have been described, PI-1, PI-2a and PI-2b, the latter two being mutually exclusive allelic variants [14,22]. The PI-1 locus is present on a mobile genetic element while PI-2 is located on the core genome [7,23]. The PI-2b pilus was primarily associated with the epidemiologically relevant clinical isolates belonging to CC17, but was not strictly restricted to this "hypervirulent" lineage causing the majority of neonatal invasive diseases. Two genomic studies investigating several contemporary CC17 isolates from Canada and South China demonstrated conservation of the PI-2b locus but loss of the PI-1 pilus which is replaced by novel mobile genetic elements encoding determinants of antimicrobial resistance [9,24]. Unlike the other two pilus loci PI-1 and PI-2a, no regulatory gene was found in the vicinity of PI-2b and the regulation of PI-2b expression has not been studied so far. It is worth mentioning that several animal isolates possess a highly similar PI-2b locus, including most bovine strains, and several fish isolates [25].

In this work, we aimed at deciphering the molecular bases of PI-2b regulation to uncover the possible role of PI-2b in the pathophysiology of ST-17 strains. Comparison of PI-2b expression between the ST-17 human clinical isolate BM110 and a non-ST-17 human pathogen A909 revealed a 5-fold higher level of PI-2b transcripts in strain A909, which translates into a 5-fold increase of PI-2b major pilin on bacterial surface. The PI-2b locus is highly similar in these two strains, apart from a 43-bp sequence forming a 35-bp hairpin structure located in the intergenic region, 184 bp upstream from the first PI-2b gene. This predicted hairpin can be found in all the available genome in the NCBI database of ST-17 isolates while missing from all non-ST-17 isolates that carry the PI-2b locus, suggesting a conserved regulatory role in ST-17 strains. Immediately upstream from PI-2b pilus locus lies the putative operon of 16 genes encoding the group B carbohydrate (also known as antigen B) as predicted by a bioinformatic analysis [16]. RNA-seq and qRT-PCR analyses indicate that the hairpin found between the loci encoding antigen B and PI-2b can serve as a transcriptional terminator in strain BM110, and its absence in A909 results in read-through transcripts. A single TSS, located 37 bp upstream from the start codon of the first PI-2b gene (*orf*), exhibiting close-to-canonic -10 and -35 boxes was identified in both GBS strains A909 and BM110 using RNA-seq and primer extension analyses.

The biological relevance of this Ppi2b promoter in PI-2b transcription was assessed at both transcriptional and protein levels using reporter plasmids carrying *gfp* as a reporter gene. The various constructs containing or not Ppi2b were cloned upstream from the reporter gene and expressed in the non-PI-2b GBS strain NEM316. Our results showed that the entire AgB-PI-2b intergenic region is needed for maximal GFP transcription. A smaller region containing Ppi2b was also able to promote GFP expression albeit at lower levels than with the entire upstream region. Expression of these reporter constructs in *L. lactis* strain NZ9000, a related non-pathogenic gram-positive species, revealed that GBS-specific regulatory factors are required to drive PI-2b transcription. Only a very weak gene reporter expression was observed in *L. lactis* NZ9000 using the entire PI-2b promoter region. No GFP activity could be detected with any of the reporter constructs in *L. lactis* at 30°C, although it is the optimal temperature for *L. lactis* growth and for GFP folding (data not shown).

The overall effect of the 43-bp hairpin presence, shown to be a transcriptional terminator, is to reduce PI-2b expression on the bacterial surface of ST-17 strains as compared to non-ST17 strains, which include one human A909 and several bovine isolates in our collection [26]. We speculate that low levels of PI-2b convey a selective advantage to ST-17 invasive isolates in a human host most likely by reducing host immune responses and/or increasing the dissemination potential of ST-17 strains. In line with the second hypothesis, approximately 90% of the weak biofilm producers belonged to CC17 and CC19 which are responsible for the majority of neonatal and adult invasive infections respectively [25]. Interestingly, bovine strains were found to form much stronger biofilms when compared to the human strains. It is thus tempting to speculate that high expression of PI-2b locus in bovine strains could contribute to biofilm formation and that this increased colonization is associated to commensalism state.

We hypothesize that the presence of the 43-bp sequence forming a hairpin-like structure and its negative effect on PI-2b expression is part of the host-adaptation of the hypervirulent ST-17 lineage. Future research will focus on identifying environmental signals or host specific cues modulating PI-2b expression in ST-17 strains.

Experimental Procedures

Bacterial strains, plasmids, and growth conditions

The bacterial strains and plasmids used in this study are listed in Table S1. GBS ST-17 clinical strains isolated from invasive infections (i.e. from blood culture, cerebro-spinal fluid or other sterile sites) were obtained from the French National Reference Center for Streptococci in Paris (http://www.cnr-strep.fr) [5]. GBS NEM316, capsular serotype III ST-23, A909, capsular serotype Ia ST-9 and GBS BM110 capsular serotype III ST-17 are well-characterized isolates from human with invasive infections. Non-ST-17 GBS isolates were previously described [10,26]. *Escherichia coli* DH5α (Gibco-BRL) was used for cloning experiments.

S. agalactiae strains were cultured in Todd Hewitt (TH) broth or agar (Difco Laboratories, Detroit, MI) at 37°C in standing filled flasks and *E. coli* in Luria-Bertani (LB) medium. *Lactococcus lactis* NZ9000 was grown in M17 medium supplemented with 1% glucose at 30°C or 37°C. Antibiotics were used at the following concentrations: for *E. coli*, erythromycin, 150 μg ml^{-1}, kanamycin, 50 μg ml^{-1}; for *S. agalactiae*, erythromycin, 10 μg; for *L. lactis*, erythromycin, 5 μg ml^{-1}.

General DNA techniques

Standard recombinant techniques were used for nucleic acid cloning and restriction analysis. Plasmid DNA from *E. coli* was prepared by rapid alkaline lysis using the Qiaprep Spin Miniprep kit (Qiagen). Genomic DNA from *S. agalactiae* was prepared using the DNeasy Blood and Tissue kit (Qiagen). PCR was carried out with Phusion Taq polymerase as described by the manufacturer (Finnzymes). Amplification products were purified with QIAquick PCR purification kit (Qiagen) and verified by sequencing.

Real-time quantitative PCR

Bacterial strains pre-cultured overnight in TH broth at 37°C in standing filled cultures were diluted to OD_{600}= 0.05 in 25 ml TH and grown at 37°C until early exponential growth phase (OD_{600}=0.3). After centrifugation of 20 mL (4°C, 7000 rpm), total RNA extraction was performed using the FastRNA PRO™ BLUE Kit (MP Biomedicals). RNA quality was assessed by agarose gel electrophoresis, residual DNA from the RNA samples was removed with the TURBO™ DNase kit (2U/μl, Ambion™ / Thermo Fischer Scientific) and extracts were stored at -80°C. RNA concentrations were measured using Nanodrop 2000c (Thermo Fischer Price). Reverse transcription (RT) was performed with the iScript™ cDNA synthesis kit (Bio-Rad). Specific primer pairs were designed to obtain a predicted amplicon size of 170 to 220 bp (see Table S2) and quantitative PCR (qPCR) was carried out with EvaGreen Universal qPCR Supermix (Bio-Rad) in a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad). Relative

gene expression levels were calculated with the ΔΔCq method [27] where expression values were normalized with the expression of the housekeeping gene gyrA and to a control sample by the CFX Manager™ Software v3.0 (Bio-Rad). Each assay was performed in triplicate with three independent cultures.

Primer extension reactions

Total RNA was used as template for primer extension reaction using a radiolabeled specific primer complementary to a sequence located downstream from the putative PI-2b promoter 2bEA3 (Table S2). For detail, synthetic oligodeoxynucleotides were 5'-end-labelled with [γ-^{32}P]-ATP (110 TBq mmol^{-1}) using T4 polynucleotidekinase. 15 µg of RNA and 2 pmol of labelled oligo-nucleotide were annealed in a total volume of 18µl of reverse transcriptase buffer (50 mM Tris-HCl; 8 mM $MgCl_2$; 30 mM KCl; 1mM DTT; pH 8.5). The mixture was incubated for 3 min at 65°C. and then 1µl (25 U) of avian myelo-blastosis virus (AMV, Boehringer) reverse transcriptase and 1µl of all dNTPs mix (20 mM each) were added. After 30 min at 42°C, reactions were stopped by the addition of 5µl of a solution containing 97.5% deionized formamide, 10 mM EDTA, 0.3% xylene-cyanol, and 3.3% bromo-phenol blue.

The corresponding Sanger DNA sequencing reactions were carried out by using the same primer and a PCR-amplified fragment containing the PI-2b upstream region (primer pair 2bEA1 and 2bEA2, table S2) with the Sequenase PCR product sequencing kit (USB).

Genome-wide mapping of transcription start sites (TSSs)

Genome-wide mapping of transcription start sites in GBS strains A909 and BM110 was performed by using a differential RNA-sequencing protocol based on selective Tobacco Acid Pyrophosphatase (TAP) treatment and 5' adapter ligation as previously described [17,28]. Briefly, for each strain, mixes of RNA prepared at mid-exponential and stationary growth phases in a rich culture medium (TH) and at the beginning of the stationary phase in a poor culture medium (RPMI supplemented with glucose 1% and pH-buffered with 50 mM Hepes) were enriched in mRNA with the MICROBExpress Kit (Ambion). dRNA-seq libraries were prepared after or without prior TAP treatment by ligation with a 5' adapter (Illumina TruSeq Small RNA kit) and by reverse transcription using a random primer. After sequencing the reads were aligned to the reference genomes (A909: NC_007432.1; BM110: unpublished complete genome sequence obtained from PacBio sequencing) and analysed for statistical assignment of TSSs based on the differences between the number of reads originating at each sequence position under TAP+ and TAP- conditions. In addition, strand-specific RNA-seq libraries were prepared from the two strains grown at mid-exponential growth phase by using the Illumina primer ligation method. dRNA-seq and RNA-seq libraries were sequenced

on the Illumina GAIIX or HiSeq 2000. Alignments of the reads to the reference genomes were visualized with the IGV genome browser [29].

Construction of mutants

In frame deletion mutants of *spb1* in BM110 and A909 were constructed as previously described [13]. The deletion was confirmed by PCR and sequencing on the genomic DNA of the mutants.

Generation of polyclonal antibody

DNA fragment intragenic to *spb1* (from amino acids 26-460) was amplified by PCR using genomic DNA of GBS BM110 as a template and primers listed in Table S2. The amplification product was cloned into the pGEM®-T easy vector (TOPO Zero Blunt for Ap1). After verification by sequencing, the fragment was digested with the appropriate enzymes and cloned into pET28a. The resulting plasmid was introduced into *E. coli* BL21λDE3/pDIA17 for protein expression. Recombinant proteins were purified on Ni-NTA columns. Protein purity was checked on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and accurate protein concentration was determined using the Bradford protein assay. Rabbits polyclonal antibodies were produced and purchased from Covalab.

Immunofluorescence microscopy

GBS strains A909 and BM110 were grown overnight in Todd-Hewitt broth. Bacteria were washed twice in phosphate buffered saline (PBS), and incubated for 1h at 4°C with rabbit primary antibody anti-Spb1 diluted in PBS-BSA 1.5% (1/1000). After two washings with PBS, samples were incubated for 30 min with secondary DyLight$_{488}$-conjugated goat anti-rabbit immunoglobulin diluted in PBS-BSA 1.5% (1/1000 dilution; Thermo Scientific Pierce) and Hoechst 33342 (1/1000). Microscopic observations were done with a Nikon Eclipse Ni-U and images acquired with a Nikon Digital Camera DS-U3. Images taken at different color channels were treated and merged in Image J.

Flow-cytometry

To analyze PI-2b expression, exponentially growing bacteria (OD_{600}= 0.3-0.4), were collected and washed twice in PBS. Bacterial pellet was incubated with rabbit anti-Spb1 primary antibody (1/1000) diluted in PBS-BSA 1% for 1h at 4°C. After two washes with PBS, samples were incubated with the secondary antibody DyLight$_{488}$-conjugated goat anti-rabbit immunoglobulin (Thermo Scientific Pierce) diluted 1/1000 in PBS-BSA 1% for 30 min at 4°C. *S. agalactiae* and *L. lactis* strains carrying GFP reporter constructs were grown overnight at 30°C in 2 ml TH broth or M17-glucose, respectively, diluted 20x the following day, and grown at 37°C to exponential phase. One ml samples were harvested, and Hoechst 33342 was

directly added in the medium (1/1000), the tubes were left open and incubated 10 minutes under gentle agitation. After washing and resuspension in PBS, samples were acquired on a MACSQuant YGV Analyzer apparatus (Miltenyi Biotec) and data were analyzed using FlowJo X software.

Cell wall protein extraction and immunoblotting

Bacteria were grown in TH medium at 37°C and harvested for protein analysis during late exponential phase of culture. Bacteria were washed in PBS and then resuspended in mutanolysin digestion mix (100Uml^{-1} Mutanolysin (Sigma) resuspended in 50 mM Tris-HCl, pH 7.3, 20% sucrose, supplemented with 1x Proteinase Inhibitor Complex (Roche)). The digestion was performed for 2 h at 37°C under gentle agitation. After centrifuging at 13.000 g for 10 min at 4°C, supernatants corresponding to the cell wall fractions were analyzed on SDS-PAGE or kept frozen at −20°C.

Western immunoblots

For analysis of major- and ancillary- pilin expression, cell wall proteins were boiled in Laemmli sample buffer, separated by SDS-PAGE on 4–12% Tris-Acetate Midi Criterion XT Precast gels (Bio-Rad), and transferred to PVDF membrane using the Trans-Blot Turbo transfer pack (Bio-Rad). Immuno-detection was performed as follow: membrane was blocked in TBS–skimmed milk 5% and incubated for 1 hour with rabbit primary Bp and Ap1 antibodies and then with the secondary Dylight$_{800}$-coupled goat anti-rabbit antibody (Thermo Scientific Pierce). Between the two antibodies and before detection, membranes were extensively washed with TBS + 0.1% Tween 20. Bound pilins were detected using the LI-COR Odyssey Infrared Imaging System (LI-COR Biosciences).

Acknowledgments

We thank Asmaa Tazi and Arnaud Firon for the construction of BM110*ΔcovR* and BM110CovRD53A mutants respectively. We thank Philippe Glaser for fruitful discussion and critical reading of the manuscript. This work was supported by the DIM MAL INF (Grant DIM130065). N.S was recipient of a doctoral fellowship from the DIM MAL INF (Grant DIM130065).

References

1. Musser JM, Mattingly SJ, Quentin R, Goudeau A, Selander RK (1989) Identification of a high-virulence clone of type III *Streptococcus agalactiae* (group B streptococcus) causing invasive neonatal disease. Proc Natl Acad Sci USA 86: 4731-4735.

2. Phares CR, Lynfield R, Farley MM, Mohle-Boetani J, Harrison LH, et al. (2008) Epidemiology of invasive group B streptococcal disease in the United States, 1999-2005. JAMA 299: 2056-2065.

3. Poyart C, Réglier-Poupet H, Tazi A, Billoët A, Dmytruk N, et al. (2008) Invasive group B streptococcal infections in infants, France. Emerging Infectious Diseases 14: 1647-1649.

4. Tazi A, Disson O, Bellais S, Bouaboud A, Dmytruk N, et al. (2010) The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. J Exp Med 207: 2313-2322.

5. Six A, Bellais S, Bouaboud A, Fouet A, Gabriel C, et al. (2015) Srr2, a multifaceted adhesin expressed by ST-17 hypervirulent group B streptococcus involved in binding to both fibrinogen and plasminogen. Mol Microbiol 97: 1209-1222.

6. Madzivhandila M, Adrian PV, Cutland CL, Kuwanda L, Madhi SA, et al. (2013) Distribution of pilus islands of group B streptococcus associated with maternal colonization and invasive disease in South Africa. J Med Microbiol 62: 249-253.

7. Rinaudo CD, Rosini R, Galeotti CL, Berti F, Necchi F, et al. (2010) Specific involvement of pilus type 2a in biofilm formation in group B streptococcus. PLoS ONE 5: e9216.

8. Springman AC, Lacher DW, Waymire EA, Wengert SL, Singh P, et al. (2014) Pilus distribution among lineages of group B streptococcus: an evolutionary and clinical perspective. BMC Microbiol 14: 159.

9. Teatero S, Ramoutar E, McGeer A, Li A, Melano RG, et al. (2016) Clonal Complex 17 group B streptococcus strains causing invasive disease in neonates and adults originate from the same genetic pool. Sci Rep 6: 20047.

10. Brochet M, Couvé E, Zouine M, Vallaeys T, Rusniok C, et al. (2006) Genomic diversity and evolution within the species *Streptococcus agalactiae*. Microbes Infect 8: 1227-1243.

11. Lazzarin M, Cozzi R, Malito E, Martinelli M, D'Onofrio M, et al. (2015) Noncanonical sortase-mediated assembly of pilus type 2b in group B streptococcus. FASEB J 29: 4629-4640.

12. Cozzi R, Malito E, Lazzarin M, Nuccitelli A, Castagnetti A, et al. (2015) Structure and assembly of group B streptococcus pilus 2b backbone protein. PLoS One 10: e0125875.

13. Chattopadhyay D, Carey AJ, Caliot E, Webb RI, Layton JR, et al. (2011) Phylogenetic lineage and pilus protein Spb1/SAN1518 affect opsonin-independent phagocytosis and intracellular survival of group B streptococcus. Microbes Infect 13: 369-382.

14. Dramsi S, Caliot E, Bonne I, Guadagnini S, Prévost M-C, et al. (2006) Assembly and role of pili in group B streptococci. Mol Microbiol 60: 1401-1413.

15. Konto-Ghiorghi Y, Mairey E, Mallet A, Duménil G, Caliot E, et al. (2009) Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. PLoS Pathog 5: e1000422.

16. Sutcliffe IC, Black GW, Harrington DJ (2008) Bioinformatic insights into the biosynthesis of the Group B carbohydrate in *Streptococcus agalac*tiae. Microbiology (Reading, Engl) 154: 1354-1363.

17. Rosinski-Chupin I, Sauvage E, Sismeiro O, Villain A, Da Cunha V, et al. (2015) Single nucleotide resolution RNA-seq uncovers new regulatory mechanisms in the opportunistic pathogen *Streptococcus agalactiae*. BMC Genomics 16: 419.

18. Jiang S-M, Cieslewicz MJ, Kasper DL, Wessels MR (2005) Regulation of virulence by a two-component system in group B Streptococcus. J Bacteriol 187: 1105-1113.

19. Lamy M-C, Zouine M, Fert J, Vergassola M, Couvé E, et al. (2004) CovS/CovR of group B streptococcus: a two-component global regulatory system involved in virulence. Mol Microbiol 54: 1250-1268.

20. Lembo A, Gurney MA, Burnside K, Banerjee A, de los Reyes M, et al. (2010) Regulation of CovR expression in Group B streptococcus impacts blood-brain barrier penetration. Mol Microbiol 77: 431-443.

21. Jiang S, Park SE, Yadav P, Paoletti LC, Wessels MR (2012) Regulation and function of pilus island 1 in group B streptococcus. J Bacteriol 194: 2479-2490.

22. Rosini R, Rinaudo CD, Soriani M, Lauer P, Mora M, et al. (2006) Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. Mol Microbiol 61: 126-141.

23. Maione D, Margarit I, Rinaudo CD, Masignani V, Mora M, et al. (2005) Identification of a universal group B streptococcus vaccine by multiple genome screen. Science 309: 148-150.

24. Campisi E, Rosini R, Ji W, Guidotti S, Rojas-Lopez M, et al. (2016) Genomic analysis reveals multi-drug resistance clusters in group B streptococcus CC17 hypervirulent isolates causing neonatal invasive disease in Southern Mainland China. Front Microbiol 7: 1265.

25. Parker RE, Laut C, Gaddy JA, Zadoks RN, Davies HD, et al. (2016) Association between genotypic diversity and biofilm production in group B streptococcus. BMC Microbiol 16: 86.

26. Da Cunha V, Davies MR, Douarre PE, Rosinski-Chupin I, Margarit I, et al. (2014) *Streptococcus agalactiae* clones infecting humans were selected and fixed through the extensive use of tetracycline. Nat Commun 5: 4544.

27. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

28. Rosinski-Chupin I, Soutourina O, Martin-Verstraete I (2014) Riboswitch discovery by combining RNA-seq and genome-wide identification of transcriptional start sites. Methods Enzymol 549: 3-27.

29. Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14: 178-192.

30. Lindahl G, Stålhammar-Carlemalm M, Areschoug T (2005) Surface proteins of *Streptococcus agalactiae* and related proteins in other bacterial pathogens. Clin Microbiol Rev 18: 102-127.

31. Lancefield RC, McCarty M, Everly WN (1975) Multiple mouse-protective antibodies directed against group B streptococci. Special reference to antibodies effective against protein antigens. J Exp Med 142: 165-179.

32. Glaser P, Rusniok C, Buchrieser C, Chevalier F, Frangeul L, et al. (2002) Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. Mol Microbiol 45: 1499-1513.

33. Firon A, Tazi A, Da Cunha V, Brinster S, Sauvage E, et al. (2013) The Abi-domain protein Abx1 interacts with the CovS histidine kinase to control virulence gene expression in group B streptococcus. PLoS Pathog 9: e1003179

34. Kuipers OP, de Ruyter PG, Kleerebezem M, de Vos WM (1998) Quorum sensing-controlled gene expression in lactic acid bacteria. J Biotechnol 64: 15-21.

Figure Legends

Figure 1. Comparative expression of the gene (*spb1*) encoding the major PI-2b pilin in *S. agalactiae* strains BM110 (ST-17) and A909 (non ST-17). (A) Schematic representation of the PI-2b pilus operon. Arrows represent coding sequence and the direction of transcription. Genes encoding pilus structural proteins are shown in purple, those encoding sortases are in brown, lep encodes a putative signal peptidase and *orf* codes for a conserved hypothetical protein. Gene nomenclature is according to [11]. (B) Immunofluorescence microscopy of *S. agalactiae* strains using anti-Spb1 antibody. (C) Flow-cytometry analysis of Spb1 expression in WT and isogenic *Δspb1* strains. (D) Transcriptional analysis of *spb1* gene by quantitative RT-PCR in exponentially growing *S. agalactiae* cells using *gyrA* as an internal standard. Results are expressed as the n-fold change with respect to the WT strain BM110 whose value has been set arbitrarily to 1. Results are means +/- SD from at least two independent cultures in triplicates. Asterisks represent P values (*P<0.05) evaluated using a Student's t test.

Figure 2. Expression of Spb1 in BM110 and in the isogenic *Δ43* mutant (deletion of the 43-bp sequence forming a hairpin-like structure) (A) Schematic comparison of the PI-2b locus in BM110 (ST-17 strain) as compared to A909 (non-ST17 strain). (B) Flow-cytometry analysis of Spb1 expression in WT BM110 and isogenic *Δ43* mutant, back to wild-type *Δ43*bWT and *Δspb1* used as a negative control. (C) Quantitative RT-PCR analysis of RNAs extracted from exponentially growing *S. agalactiae*. The expression levels were normalized using *gyrA*. Results are means +/- SD from at least two independent cultures in triplicates. Asterisks represent P values (*P<0.05) evaluated using a Student's t test.

Figure 3. Role of the PI-2b hairpin-like sequence as a transcriptional terminator. Detection of a read-through transcript (arrowhead) encompassing the last gene of *AgB* operon and the first gene of PI-2b locus in strain A909 but not in strain BM110. RT-PCR was performed using the Gly3/Gly4 (lanes 1, 4, 7, 10, 13, 16), ORF1/ORF2 (lanes 3, 6, 9, 12, 15, 18) and Gly1/ORF2 (lanes 2, 5, 8, 11, 14, 17) oligonucleotides pairs (Table S2). Lanes 1 to 3 and 10 to 12, positive controls using chromosomal DNA (chr); lanes 4 to 6 and 13 to 15, read-through transcript of *san1522* (according to COH1 annotation) and *orf* (cDNA); Lanes 7 to 9 and 16 to 18 represent negative controls without reverse transcriptase (NRT).

Figure 4. Transcriptional start site of the PI-2b operon. (A) Characterization of PI-2b transcription start site in *S. agalactiae* strains A909 and BM110 by dRNA-seq. The sequence reads corresponding to transcript 5' ends generated after (TAP+) or without (TAP-) TAP treatment or obtained from strand-specific RNA-seq (RNA-seq) were aligned to the genomes of strains A909 and BM110. A significantly higher number of reads under TAP+ conditions as compared with TAP- is indicative of 5' -triphosphate ends of transcripts characteristic of

transcription start sites (TSS). Identical TSS upstream PI-2b locus are detected in strains A909 and BM110. In addition, coverage of the intergenic regions between *sak1445* (or *san1522*) and *orf* under conditions of RNA-seq experiments reveals a transcriptional read-through originating from *sak1445*, in A909 only, that could participate to the global level of PI-2b transcription in A909. (B) Primer extension analysis of the PI-2b mRNA. Primer elongation product obtained with oligonucleotide E3 and 15 µg of total RNA from A909 (lane1) or BM110 (lane 2). Lanes T, G, C, A, results of sequencing reactions performed with primer E3. Arrow indicated the transcriptional start site. (C) Schematic representation and genomic DNA sequence, from nucleotide positions -281 to +3 (numbering from the A of the ATG start codon of *orf*, negative in the -3'-to-5' direction and positive in the 5' to-3' direction) of the region upstream from the PI-2b locus. Transcription start site is indicated in boldface and arrow. Consensus -10 and -35 sequences are indicated by grey boxes. The 43-bp sequence, present in BM110 and absent in A909, is underlined. The ATG start codon of *orf* is indicated in boldface.

Figure 5. Characterization of PI-2b promoter region using transcriptional fusions with *gfp* as a reporter gene. (A) Schematic representation of the region upstream from the PI-2b locus. The putative +1 transcriptional start site detected by primer extension is indicated. Length of the intergenic region is indicated at the bottom. The four pTCV plasmids (pTCV1 to pTCV4) are drawn schematically. The size of the PI-2b fragment inserted upstream of the reporter gene is indicated on the left. (B) Flow-cytometry analyses of GFP expression in *S. agalactiae* NEM316 (ST-23) (left part) or in the heterologous *Lactococcus lactis* strain NZ9000 harboring the various pTCV1-4 plasmids (right part). Left, Black NEM316 + pTCV4; blue, NEM316 + pTCV2; green, NEM316 + pTCV1; orange, NEM316 + pTCV3; Right, Black NZ9000 + pTCV4; blue, NZ9000 + pTCV2; green, NZ9000 + pTCV1; orange, NZ9000 + pTCV3. (C) Role of CovR on PI-2b pilus expression in *S. agalactiae*. Left panel, Flow-cytometry analysis of GFP expression in NEM316 WT (black) and isogenic *ΔcovR* (orange) and CovRD53A (blue) strains harboring pTCV3. NEM316 carrying pTCV2 (grey) was used as a negative control. Right panel, transcriptional analysis of *san1522* (AgB locus, according to COH1 annotation) and PI2b-*orf* by quantitative RT-PCR in exponentially growing *S. agalactiae* cells using *gyrA* as an internal standard. Results are expressed as the n-fold change with respect to the WT strain BM110 whose value has been set arbitrarily to 1. Results are means +/- SD from at least two independent cultures in triplicates.

Figure 1.

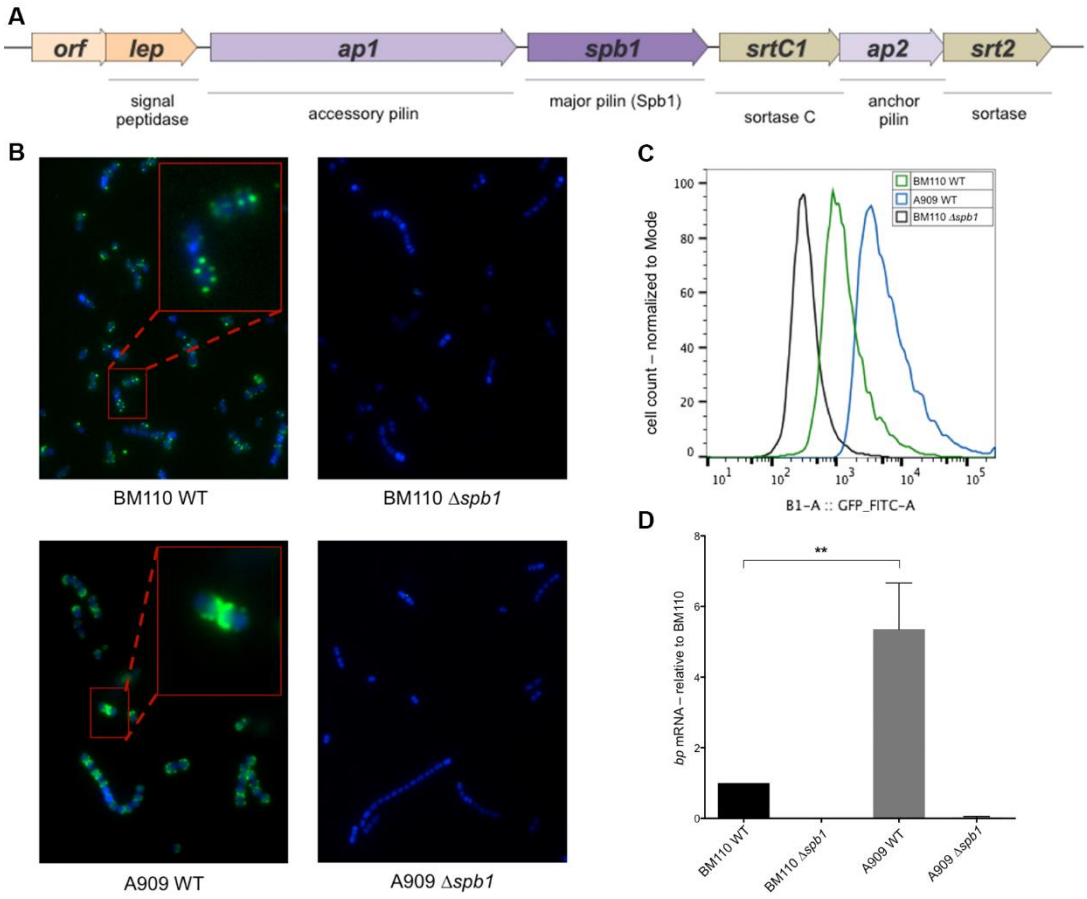

A
orf
lep
ap1
spb1
srtC1
ap2
srt2
signal peptidase
accessory pilin
major pilin (Spb1)
sortase C
anchor pilin
sortase
B
BM110 WT
BM110 Δspb1
A909 WT
A909 Δspb1
C
BM110 WT
A909 WT
BM110 Δspb1
cell count – normalized to Mode
B1-A :: GFP_FITC-A
D
**
bp mRNA – relative to BM110
BM110 WT
BM110 Δspb1
A909 WT
A909 Δspb1

Figure 2.

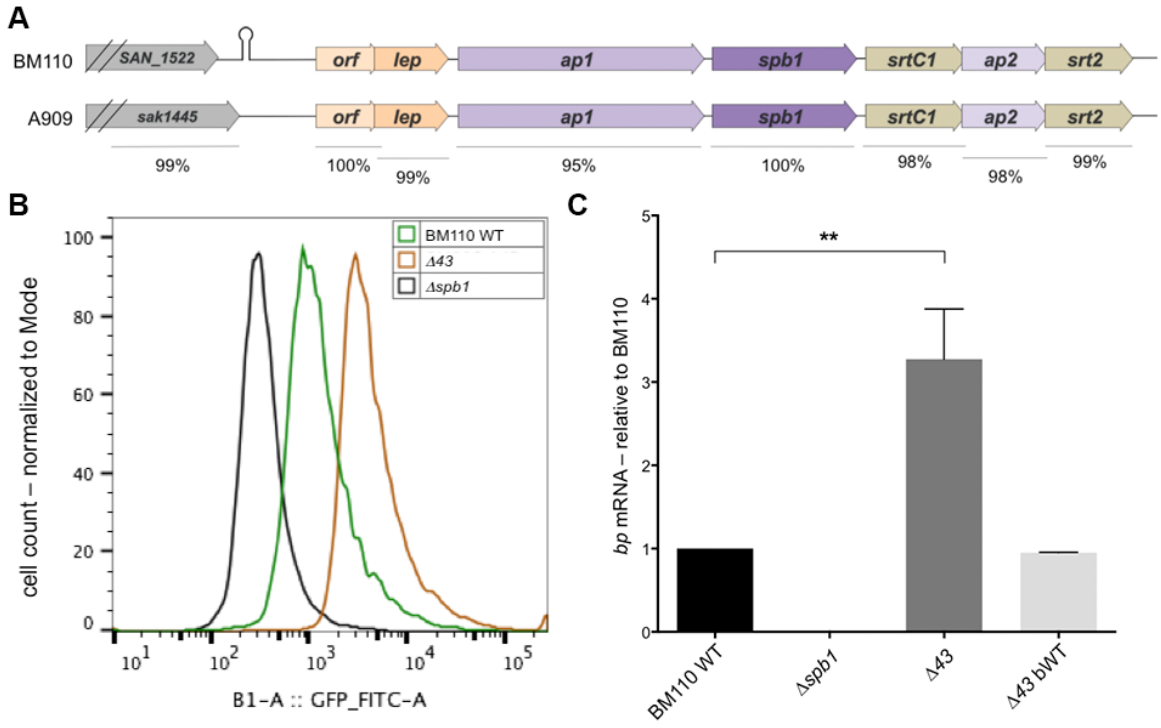


Figure 3.

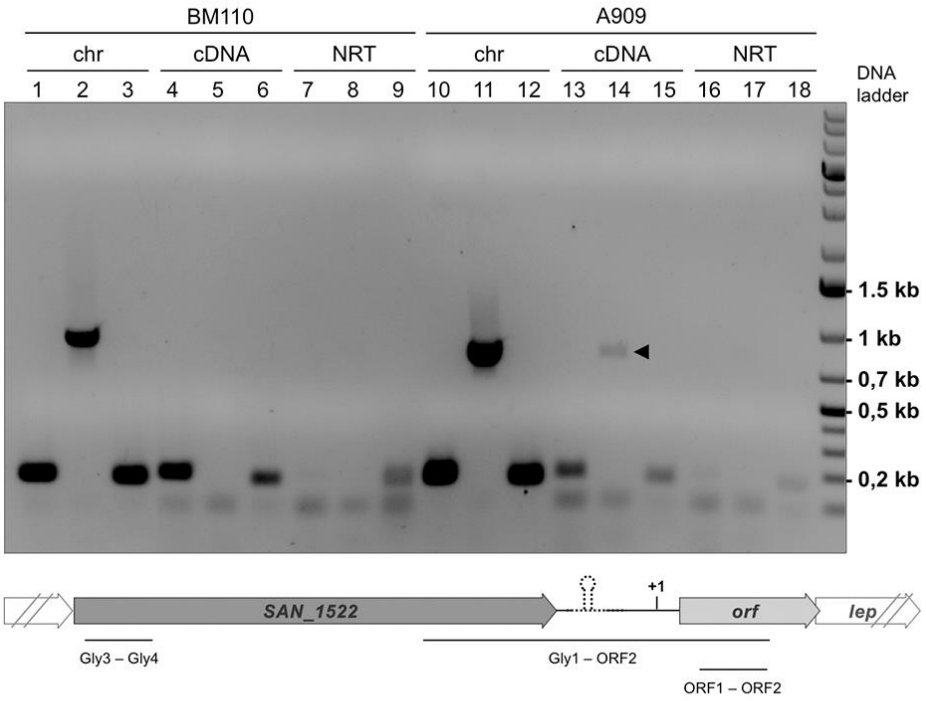


Figure 4.

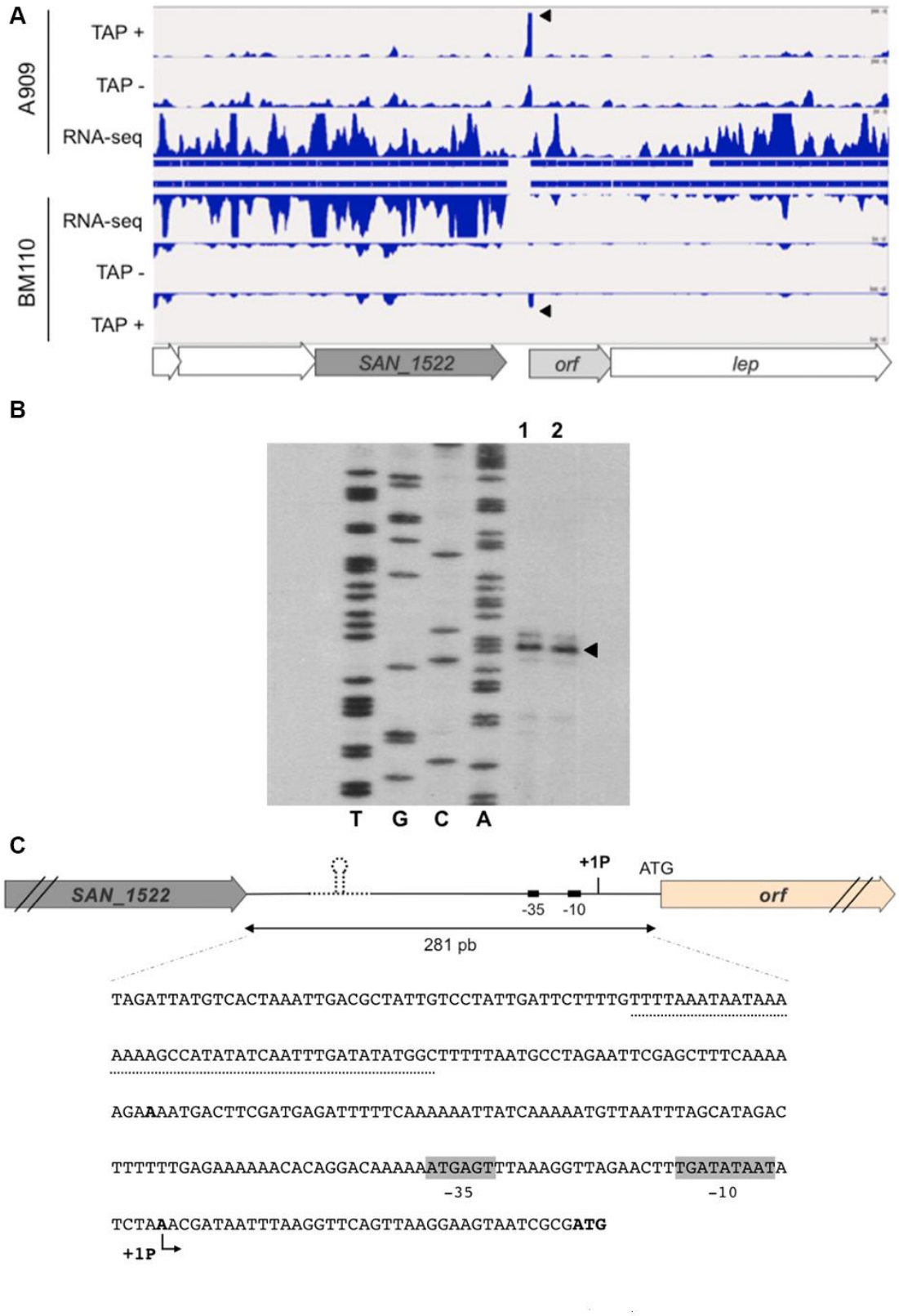
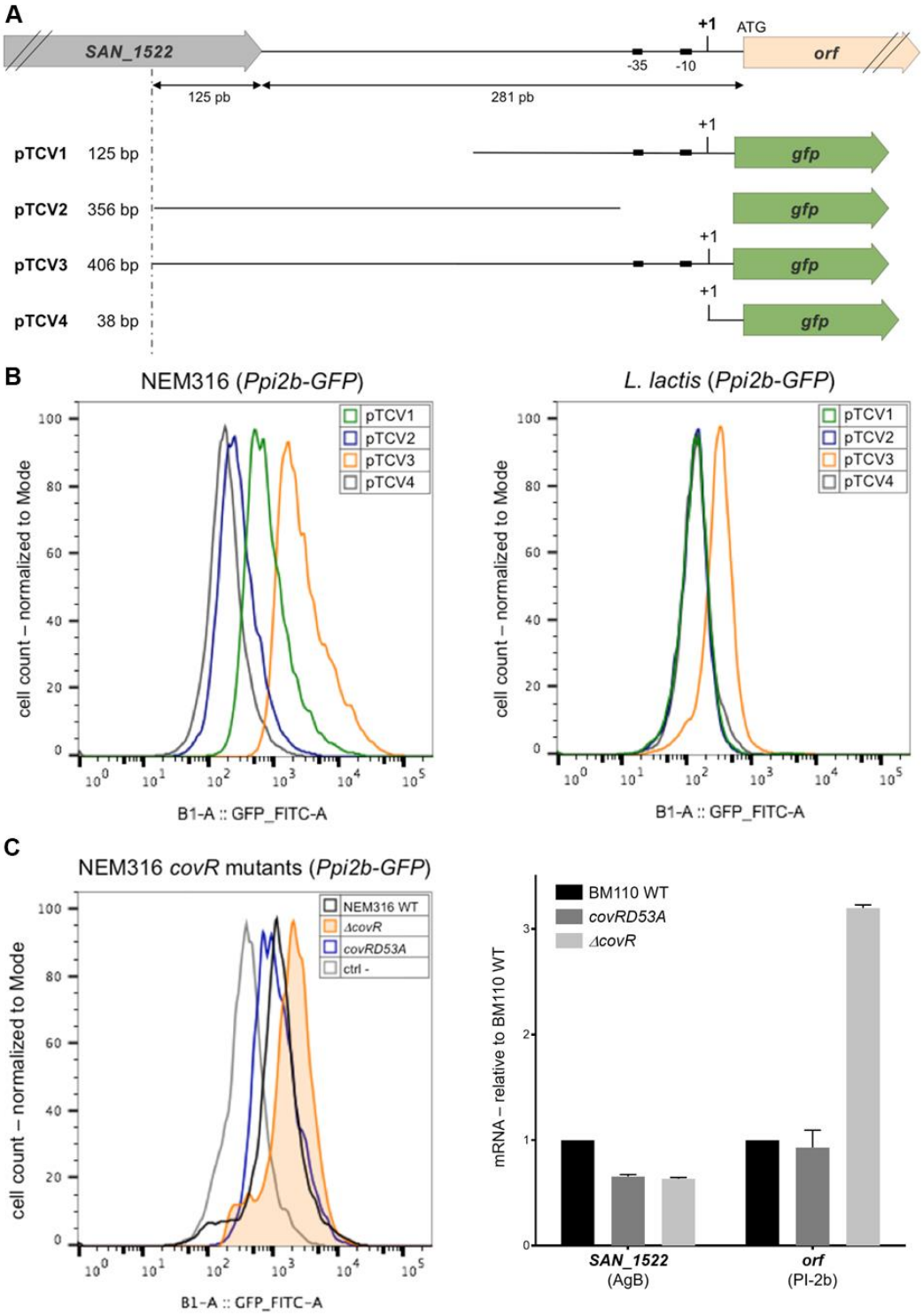

A
A909
TAP +
TAP -
RNA-seq
BM110
RNA-seq
TAP -
TAP +
SAN_1522
orf
lep
B
1 2
T G C A
C
SAN_1522
+1P
ATG
orf
-35 -10
281 pb
TAGATTATGTCACTAAATTGACGCTATTGTCCTATTGATTCTTTTGTTTTAAATAATAAA
AAAAGCCATATATCAATTTGATATATGGCTTTTTAATGCCTAGAATTCGAGCTTTCAAAA
AGAAAATGACTTCGATGAGATTTTTCAAAAAATTATCAAAAATGTTAATTTAGCATAGAC
TTTTTTGAGAAAAAACACAGGACAAAAAATGAGTTTAAAGGTTAGAACTTTGATATAATA
-35
-10
TCTAAACGATAATTTAAGGTTCAGTTAAGGAAGTAATCGCGATG
+1P

Figure 5.


A
SAN_1522
+1
ATG
orf
-35
-10
125 pb
281 pb
pTCV1 125 bp
pTCV2 356 bp
pTCV3 406 bp
pTCV4 38 bp
gfp
B
NEM316 (Ppi2b-GFP)
L. lactis (Ppi2b-GFP)
pTCV1
pTCV2
pTCV3
pTCV4
cell count – normalized to Mode
B1-A :: GFP_FITC-A
C
NEM316 covR mutants (Ppi2b-GFP)
NEM316 WT
ΔcovR
covRD53A
ctrl -
BM110 WT
covRD53A
ΔcovR
mRNA – relative to BM110 WT
SAN_1522
(AgB)
orf
(PI-2b)

Supplementary tables

Table S1. Strains and plasmids

Strain or plasmid	Relevant characteristics	Reference
Strains		
E. coli		
DH5α™	F– Φ80Δ*lacZ*Δ*M15* Δ(*lacZYA-argF*) U169 *recA1 endA1 hsdR17* (rK–, mK+) *phoA* supE44 λ– *thi-1 gyrA96 relA1*	Invitrogen™
S. agalactiae		
BM110	Serotype III, ST-17, human clinical isolate	(30)
BM110*ΔcovR*	In frame deletion of *covR* in BM110	Asma Tazi
BM110CovRD53A	BM110 mutant expressing CovR with a D53A substitution	Arnaud Firon
BM110*Δbp*	In frame deletion of *sbp1* (*san1519 according to COH1*)	This study
BM110*Δbp* bWT	Back to the WT strain obtained during the construction of BM110*Δbp*	This study
BM110*Δ43*	Deletion of the 43-bp sequence (5' GTTTTAAATAATAAAAAAAGCCATATATCAATTTGATATATGGC)	This study
BM110*Δ43* bWT	Back to the WT strain obtained during the construction of BM110*Δ43*	This study
A909	Serotype Ia, ST-7, human clinical isolate	(31)
A909*Δbp*	In frame deletion of *bp* (*sak1439*)	This study
A909*Δbp* bWT	Back to the WT strain obtained during the construction of A909*Δbp*	This study
NEM316	Serotype III, ST-23, human clinical isolate	(32)
NEM316*ΔcovR*	In frame deletion of *covR* in NEM316	(33)
NEM316CovRD53A	NEM316 mutant expressing CovR with a D53A substitution	(33)
L. lactis		
NZ9000	*L. lactis* subsp. *cremoris* MG1363 containing the *nisRK* gene in the genome	(34)
Plasmids		
pTCVΩ*gfp*	*EGFP* expression vector	
pTCV1	pTCVΩ*gfp* with 125-bp PCR product (2bUp14[a] / 2bUp12) from A909[b]	This study
pTCV2	pTCVΩ*gfp* with 356-bp PCR product (2bUp6 / 2bUp20) from A909	This study
pTCV3	pTCVΩ*gfp* with 406-bp PCR product with 2bUp6 / 2bUp12 from A909	This study
pTCV4	pTCVΩ*gfp* with 38-bp DNA fragment (2bUp1/2bUp2) (AAACGATAATTTAAGGTTCAGTTAAGGAAGTAATCGCG)	This study

[a] Primer sequences are listed in Table S2.

[b] Genomic DNA of the corresponding strain was used as template for sequence amplification

Table S2. Primers used in this study

Primer	Sequence (5'-3')
For plasmids construction	
2bUp1	TCTC*GAGCTCCCCGGG*AAACGATAATTTAAGGTTCAGTTAAGGAAGTAATCGCG*GGATCC*TCTC TC[a]
2bUp2	GAGAGA*GGATCC*CGCGATTACTTCCTTAACTGAACCTTAAATTATCGTTT*CCCGGGGAGCTC*GA GA[a]
2bUp6	TC*GAGCTCCCCGGG*TGATAGCCCTGAGTTAGGA[b]
2bUp12	TCTC*GGATCC*ATTATCGTTTAGATATTATATC[c]
2bUp14	TCTC*CCCGGG*GACTTCGATGAGATTTTTC[d]
2bUp20	TCTC*GGATCC*GTCCTGTGTTTTTTCTC[c]
For qRT-PCR	
GyrAgal-1	GAGCGTCAGAGTCAAGCTAT
GyrAgal-2	GCTTAACCTCATCCATCTCT
Spb1	CCTGGGTCATCATTGCTAGT
Spb2	CGATTACTATTCCGTGGGCA
ORF1	GAACAGCAACTTCTTATTG
ORF2	AATAGGTATCTGCAGCTAA
Gly1	ATGGCACAGACAGATATCT
Gly3	CTTCCTCGGAGGTATTGAAC
Gly4	AGGATAGCGCTGCTTGACAA
SAN0698-1	CCAGTAGCTCAGTTTGCGAC
SAN0698-2	CATTGTCACCAAGTTTAGCATAG
For primer extension	
2bEA1	CTGTAATAGATTATGTCAC
2bEA2	TTCAATAAGAAGTTGCTG
2bEA3	CATAATTTCTAGCAAGTC

[a] In italics, *Sac*I, *Sma*I and *Bam*HI sites

[b] In italics, *Sac*I and *Sma*I sites

[c] In italics, *Bam*HI site

[d] In italics, *Sma*I site

Supplementary figure legends

Figure S1. Schematic comparison of the PI-2a genomic locus in strain NEM316 versus the allelic PI-2b locus in strain A909. The gene annotated *gbs1479* [30], also known as *rogB*, encoding a RofA-like transcriptional regulator, is shown in green. No regulatory gene can be found in the vicinity of PI-2b operon.

Figure S2. Expression of the PI-2b major (Spb1) and minor (Ap1) pilins in *S. agalactiae* strains BM110 (ST-17) and A909 (non-ST-17). Western blot analysis of cell wall anchored proteins isolated from *S. agalactiae* BM110 and A909, separated on 4%-12% gradient Criterion XT SDS-PAGE, and detected by immunoblotting with specific polyclonal anti-Spb1 and anti-Ap1 antibodies. Equivalent amounts (15 µg) of total protein was loaded in each well. The monomers (m) and high-molecular weight species corresponding to pili polymers (p) of Spb1 and Ap1 are indicated.

Figure S3. Role of CovR on PI-I pilus expression in *S. agalactiae* BM110. Transcriptional analysis of *san0698* (according to COH1) encoding the major pilin of the PI-1 pilus by quantitative RT-PCR in exponentially growing *S. agalactiae* cells using *gyrA* as an internal standard. Results are expressed as the n-fold change with respect to the WT strain BM110 whose value has been set arbitrarily to 1. Results are means +/- SD from at least two independent cultures in triplicates.

Figure S1.

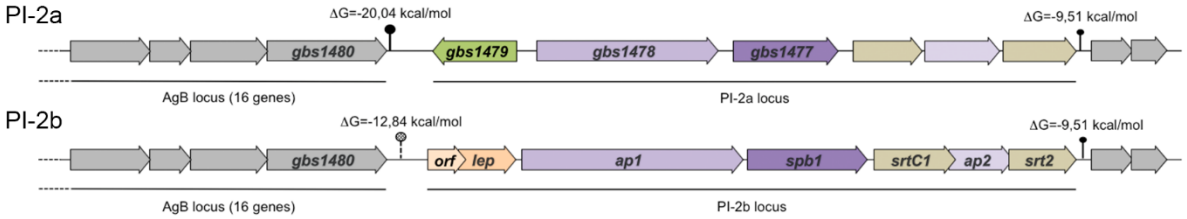


Figure S2.

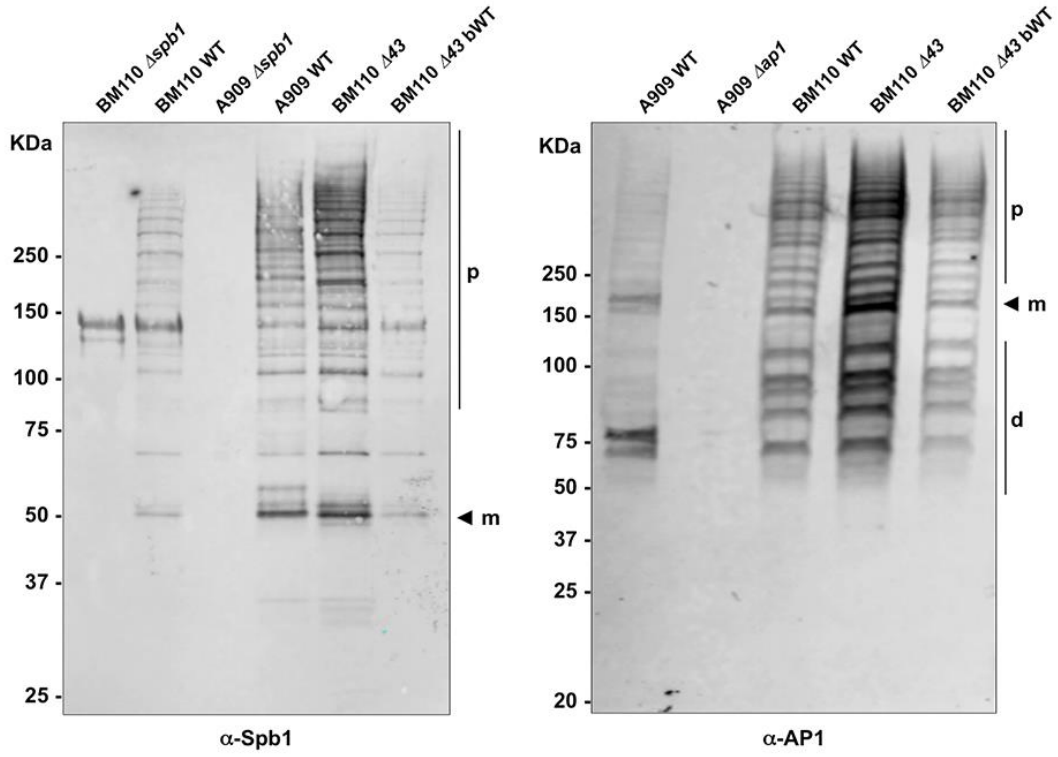
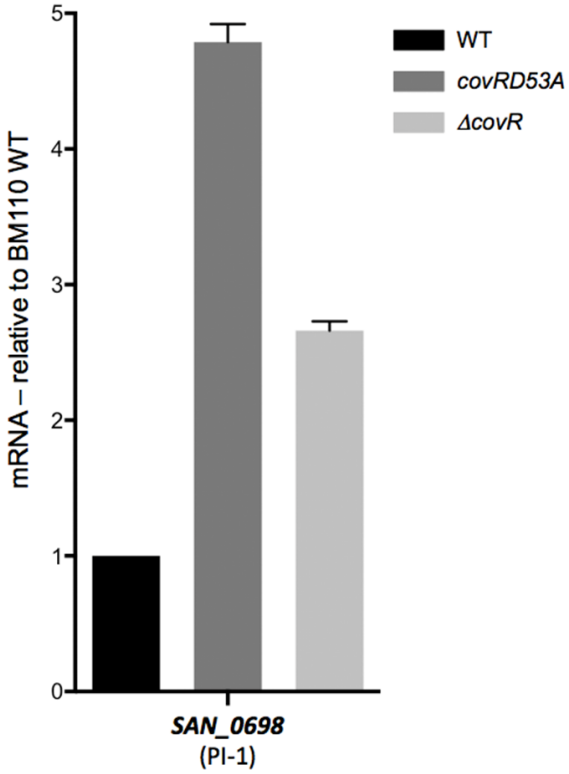


Figure 3.


mRNA – relative to BM110 WT
5
4
3
2
1
0
WT
covRD53A
ΔcovR
SAN_0698
(PI-1)

Additional results

The hairpin affects the transcription of its environment

To assess the effect of the 43bp hairpin-forming sequence on the expression of downstream genes, fragments encompassing different domains of PI-2b intergenic region (Figure 1A) were cloned upstream of the reporter gene encoding *gfp* using the vector pTCV-GFP, a low-copy number plasmid and assayed for GFP expression in the non-PI-2b GBS strain NEM316 (ST-23). The various constructs containing or not the 43-bp hairpin sequence to mimic BM110 or A909 context, respectively, are drawn schematically on Figure. Plasmid pTCV4 was used as a negative control. Quantification of GFP fluorescence was performed using flow-cytometry in exponentially growing bacteria in TH broth supplemented with erythromycin for plasmid maintenance.

Among the constructs without the hairpin sequence (pTCV1, pTCV2, and pTCV3), highest level of GFP expression was observed when the promoter region contained the full intergenic sequence (publication – Figure 5B). Similar analysis with fragments containing the hairpin sequence (pTCV5 and pTCV6) showed decrease in GFP expression with pTCV6 and pTCV7 as compared to pTCV2 and 3, respectively (Figure 5A). This was also verified at the transcriptional level using *lacZ* plasmid fusions with the same Ppi-2b promoter fragments and monitoring *lacZ* gene transcription by qRT-PCR (results not shown). These results suggest that the presence of the hairpin modulate PI-2b expression independently of its transcriptional terminator activity.

RT-PCR analysis of *san1552*, the last gene of AgB operon, and of the first gene (*orf*) of PI-2b in GBS strains A909 and BM110 confirms the role of the hairpin as a transcriptional terminator (publication – Figure 3). Furthermore, using a set of primer pairs to cover the entire *san1552*, we also observed the decrease of mRNA detectability by qRT-PCR in the proximity of the hairpin in BM110, compared to the levels measured for the corresponding region in BM110*Δ43* mutant in which the hairpin was deleted (Figure 5B). This result suggests that the effect of the hairpin on the transcription of its environment is bidirectional, and affects both the upstream and downstream genes by an unknown mechanism.

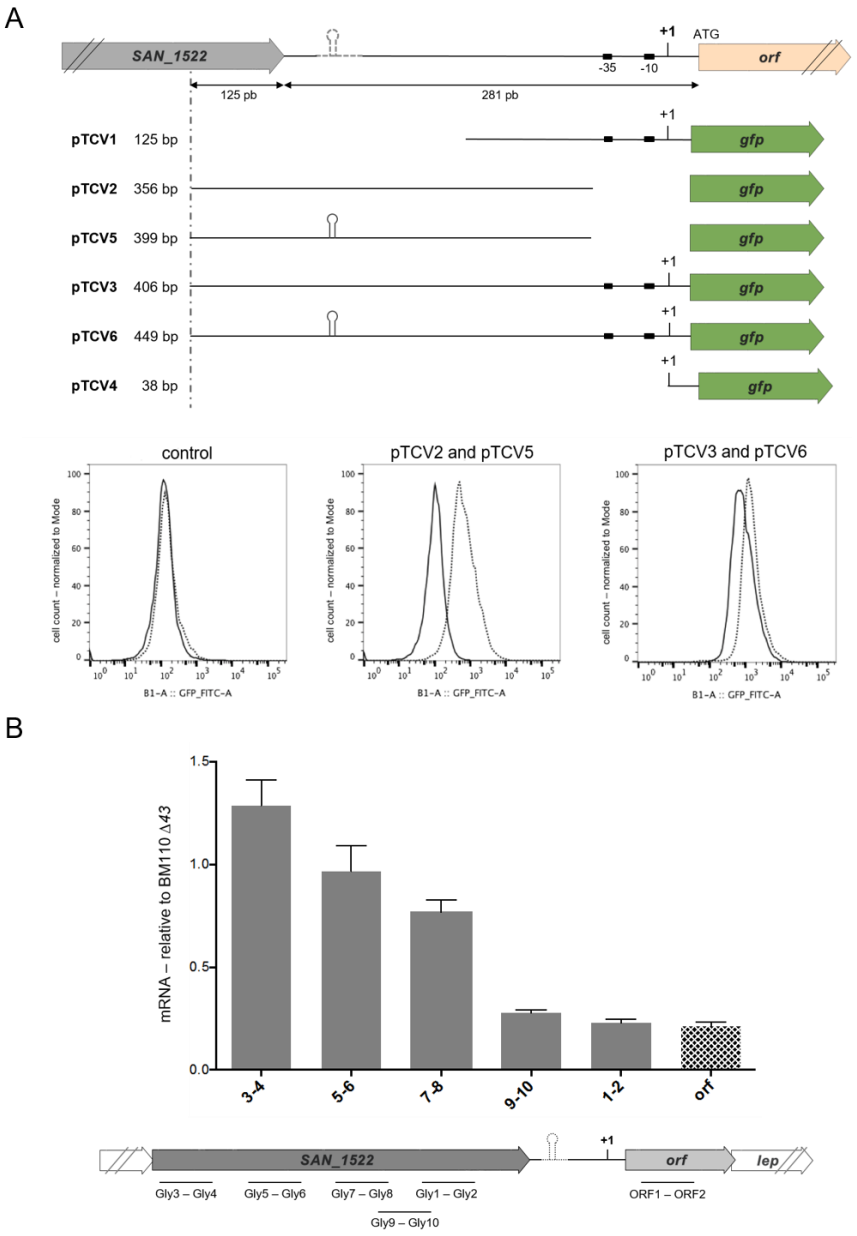


Figure 5 | Effect of the 43bp hairpin forming sequence in the PI-2b promoter region on GFP expression. *(A) Schematic representation of the various pTCV gfp-containing $P_{PI\text{-}2b}$ reporter constructs that were introduced in GBS strain NEM316 and GFP expression analyzed by flow-cytometry. The first panel, negative control where no promoter region is cloned upstream of gfp (pTCV2, solid line) or pTCV4 (dotted line). Middle panel, GFP expression for pTCV2 (solid line) and its hairpin-containing version, pTCV5 (dotted line). Right panel, GFP expression in pTCV6 (solid line) and pTCV3 (dotted line). (B) Quantification of mRNA transcripts upstream and downstream of the hairpin structure. The five primer pairs corresponding to different regions of san1522 coding for the last glycosyltransferase of the AgB operon, and to orf of PI-2b are schematically represented. Transcription levels measured by qRT-PCR were normalized to the housekeeping gene gyrA, the chart represents the relative amount of each mRNA product in BM110 wild type relative to mRNA levels in the isogenic Δ43 hairpin deletion mutant. Results are means +/- SD from at least two independent cultures in triplicate.*

Close inspection of the 43 bp hairpin sequence revealed a putative binding site for CovR, the response regulator protein of the CovR/S two-component system. To assess the possible role of this system in PI-2b regulation, we tested the GFP expression of our promoter constructs in the covRD53A and *ΔcovR* mutants of NEM316. In NEM316 *ΔcovR*, we observed a 2-3-fold increase in GFP expression driven from the PI-2b promoter containing the transcription start site (publication – Figure 5C), and this potentiating effect was not altered by the presence of the 43 bp hairpin. The sequence identified in the hairpin is therefore not participating to the regulation by CovR under these circumstances, and the decrease of transcription in the presence of this 43bp sequence is due to other mechanisms.

Temperature-dependent activity of the PI-2b promoter

Highest GFP expression from PI-2b promoter construct observed at 37°C

Secondary nucleic acid structures such as hairpins usually show temperature-dependent structural stability, and destabilization at higher temperatures. As the 43bp hairpin-forming sequence identified in the genome of ST-17 strains upstream of the PI-2b promoter reduces the expression of a *gfp* reporter gene from a plasmid fusion, we sought to identify if this effect is temperature-dependent. First, in order to see if the PI-2b promoter has a temperature-dependent activity even independent of the presence of the hairpin, we grew the NEM316 GBS strain transformed with the plasmid containing the entire PI-2b upstream intergenic region without the hairpin (pTCV3 – Figure 5A) at 30°C, 37°C and 41°C. The highest GFP expression was observed for the bacteria grown at 37°C, while surprisingly, GFP expression was reduced 2-fold both at 30°C and 41°C (Figure 6). This suggests that expression form the PI-2b promoter might be important in the human host. Furthermore, since we showed that GBS-specific factors are necessary for transcription from this promoter, our results also indicate that some of these regulatory factors have temperature-dependent activity.

The hairpin only affects expression from Ppi-2b at 37°C

In order to investigate possible additional effects of the presence of the hairpin on the temperature-dependence of the activity of the PI-2b promoter, we compared GFP expression of the reporter constructs with and without the 43bp sequence (plasmids pTCV6 and pTCV3) at 30°C, 37°C and 41°C. Reduction of GFP expression in the presence of the hairpin was already shown at 37°C, but surprisingly no similar effect was observed at 30°C and 41°C, where GFP-expression was not further reduced (Figure 6). Therefore, the hairpin has no contribution to the temperature-dependence of PI-2b promoter activity, and only

reduces the expression of its environment at higher levels of transcription, suggesting that it might serve as a target of other regulatory elements.

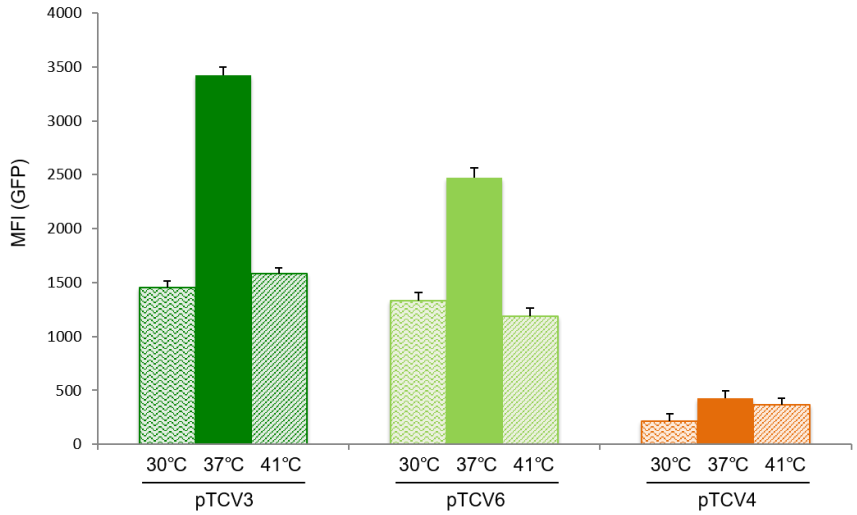


Figure 6 | Expression of PI-2b promoter is maximal at 37°C. *Results obtained at the noted temperatures with pTCV3 in dark green, pTCV6 in light green, pTCV4 (negative control) in orange, the plasmid constructs are schematically represented in Figure 5. Results are mean fluorescence intensities +/- SD from at least three independent cultures.*

Temperature-dependent Spb1 expression was not observed in A909 and BM110

We showed a temperature-dependent activity of the PI-2b promoter via a reporter gene construct, where expression at 30°C and 41°C was reduced 2-fold. However, comparing Spb1 surface expression in BM110 and A909 grown at 30°C, 37°C and 41°C, no such difference could be observed neither by whole cell blotting nor by flow cytometry. As the ensemble or our results suggests a complex regulation of PI-2b pilus expression by several factors at multiple levels, the possibility of the temperature-dependent activity of the promoter being masked by other regulatory events under these conditions is highly possible.

Spb1 expression in GBS ST-17 clinical isolates

Comparison of the PI-2b locus in 32 publicly available ST-17 genomes

Upon alignment of partially assembled genomes of 32 publicly available ST-17 clinical isolates, representing different clusters of ST-17 strains (Da Cunha *et al.*, 2014), the PI-2b locus was found to be present in all strains. For most genes, DNA sequence alignments

showed 100% identity, while we identified 2 alleles for the *srt2* gene, found at the same frequency, differing only in one base-pair at the 451 bp position, resulting in the substitution of a Glycine residue for a Serine at the 150th amino acid position of the Srt2 protein. This difference does not affect the catalytic domain of Srt2 the protein and has no predicted effect on its structure. The encompassing genomic region (ie. 1000 bp upstream and downstream) was also identical in most strains, containing the upstream 43 bp hairpin-forming sequence identified in the ST-17 reference strain BM110 and missing from the non-ST-17 reference strain A909. In only one strain, CCH208800031 (designated here as strain 25), the 43 bp sequence was absent. This strain also contained a frameshift mutation in *spb1*, leading to a truncated form of Spb1 protein. Accordingly, no PI-2b pilus was detected in ST-17 strain 25 (Figure 7B).

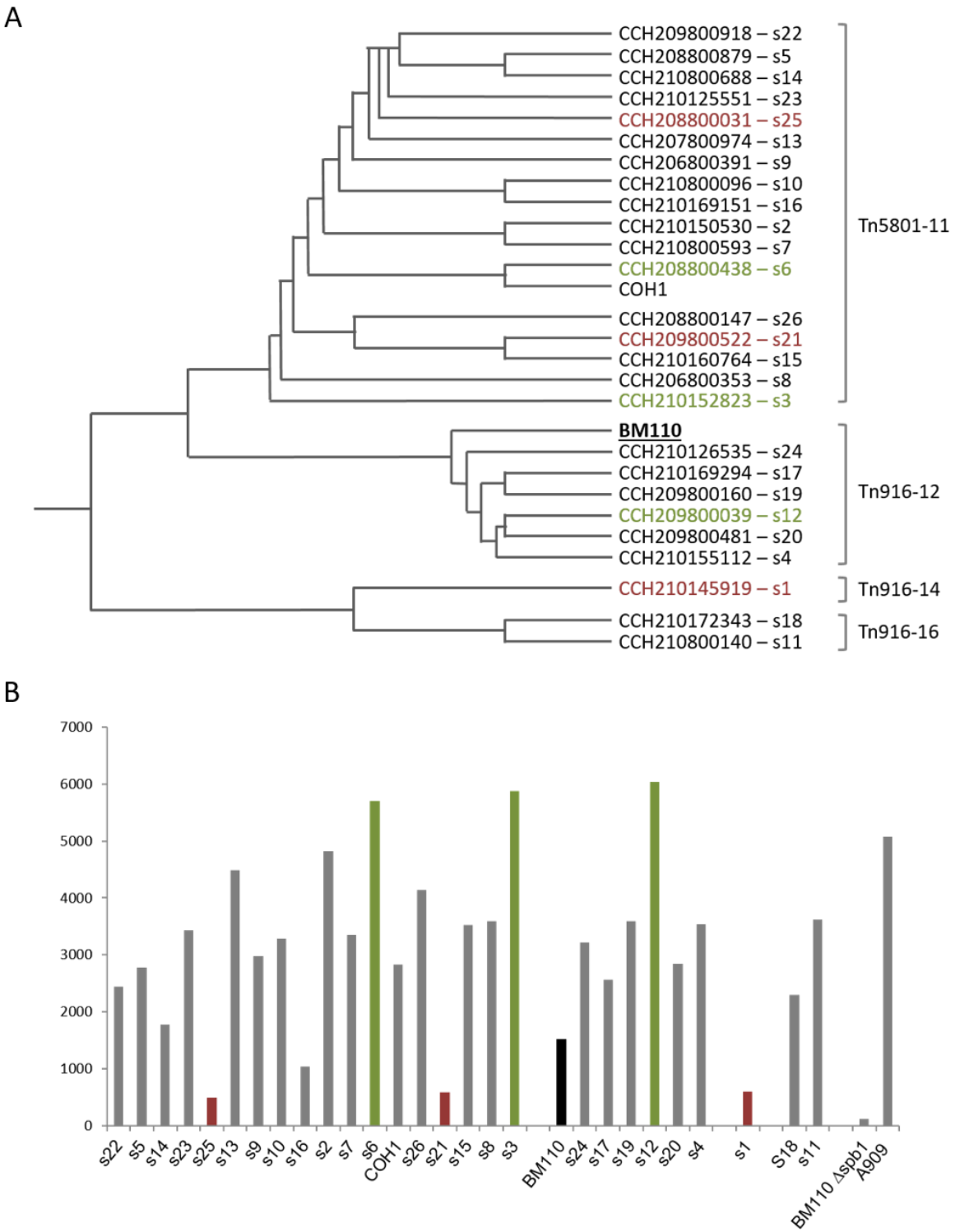


Figure 7 | Comparison of Spb1 levels in 26 ST-17 clinical isolates. *A) Phylogenetic representation of the 26 ST-17 strains obtained from the CNR Streptococci with the indication of the corresponding ST-17 clusters as mentioned in Da Cunha et al, 2014. BM110 and COH1 are generally used as ST-17 reference strains. B) Mean fluorescence intensity measured by flow-cytometry using Spb1-specific primary, and* $Alexa_{488}$*-conjugated secondary antibodies in exponentially growing bacteria (*OD_{600}*=0,5). Our ST-17 reference strain, BM110 is indicated in black, the three highest expressing strains in green, the three lowest in red. Results are mean fluorescence intensities obtained from two independent cultures.*

Heterogeneous expression of Spb1 among ST-17 strains

A set of 26 sequenced ST-17 clinical isolates obtained from the French National Reference Center for Streptococci (Figure 7A) were compared for the expression of PI-2b by quantification of the major pilin Spb1 by dot blot and flow-cytometry analyses. A preliminary analysis with whole bacteria spotted on nitrocellulose membrane revealed highly heterogeneous levels of Spb1 in the various clinical isolates, with BM110 being similar to the lower-average expressing strains. As internal loading control, i.e. to check the quantity of bacteria loaded on the membrane, we used a polyclonal antibody directed against BM110. However, we found that this antibody did not recognize equivalently all ST-17 isolates and thus could be a source a bias, and we thus decided to analyze Spb1 expression by flow cytometry as well. Flow-cytometry results also demonstrated an important heterogeneity of Spb1 levels detected at the surface of ST-17 clinical isolates (Figure 7B). The highest and lowest expressing strains were shown to be the same by the two approaches, notably s3, s6 and s12 as the highest expressing and s1, s21 and s25 as the lowest or not expressing (Figure 7B). As s25 is the isolate that contains a frameshift mutation in the *spb1* gene, this result confirms that no detectable Spb1 protein is produced in this strain.

Spb1 detectability by specific antibodies is not masked by the capsule

In order to determine whether the differences in Spb1 detectability in the different ST-17 isolates were a result of the possible masking of the pilus structures by the capsule, we measured capsule expression in all isolates, by whole cell blot and flow cytometry, using polyclonal antibodies against the serotype-III capsule. A non-capsulated mutant of BM110, where *cpsE*, a gene essential for capsule biosynthesis was deleted, was used as control. The capsule expression of the ST-17 strains was shown to be also variable, but no correlation between Spb1 and capsule expression levels could be established. Spb1 detection levels were shown to be similar in BM110 and its isogenic non-capsulated *ΔcpsE* mutant, suggesting that the capsule is not responsible for the differences in Spb1 detection among ST-17 strains. Although all ST-17 isolates were described as belonging to the capsular serotype III, there was no signal detected for the s1 isolate with specific antibodies, which questions the classification of this strain as ST-17.

Genome comparison of high and low expressing strains

In order to identify potential regulators of PI-2b, expression, we compared the draft genomes of the three highest expressing strains, s3, s6 and s12 and the three lowest expressing strains, s1, s21 and s25 to that of BM110 and searched for mutations in known regulatory genes, using the RAST software for sequence alignment and annotation. In the

s25 strain, a punctual mutation was identified in the *spxA2* gene, which is involved in oxidative stress response regulation, but a frameshift mutation in *spb1* was also subsequently identified in this strain. The receptor-kinase CovS of the CovR/S two component system involved in virulence regulation was found to have an amino acid substitution of a Glycine to a Serine at the 133th position in s21; however, the effect of this point mutation in CovS activity has not been tested.

Variable levels of spb1 transcription in high expressing strains

Transcription levels of PI2b*orf*, the first gene of the PI-2b locus, were also analyzed by qRT-PCR in the high expressing strains s3, s6 and s12 (Figure 7B and 8A), as compared to BM110 in order to determine if the 3-4-fold higher level of Spb1 expression (Figure 7B) is a direct result of higher transcription levels. As shown in figure 8, *orf*-specific transcripts were variable in these isolates. Only GBS s12 showed a 3-fold increase in PI-2b transcription that correlates with the cytometry results (Figure 8A). For the other two strains s3 and s6, no or only partial increase in PI-2b transcripts was observed as compared to BM110 suggesting post-transcriptional levels of regulation (Figure 8B). As control, we checked for the expression of the glycosyl-transferase-coding gene *san1522* located upstream from PI-2b operon and encoding the last gene of antigen B operon in these various clinical isolates. With two different primer pairs, *san1522* (GlycT) levels were shown to be similar in s3, s6, s12 and BM110 strains (Figure 8B). These results suggest a complex and specific regulation of PI-2b expression in ST-17 clinical isolates, both at transcriptional and post-transcriptional levels.

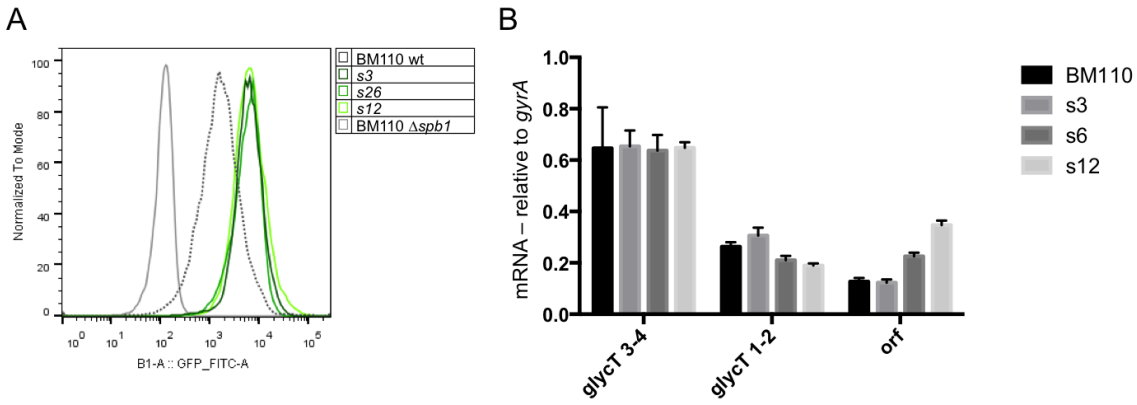


Figure 8 | Quantification of PI-2b transcription in the three isolates expressing high levels of PI-2b *A) Spb1 levels were measured by flow-cytometry using Spb1-specific primary, and Alexa$_{488}$-conjugated secondary antibodies in exponentially growing bacteria (OD$_{600}$=0,5). CNR collection isolates are represented in green as indicated, reference strain BM110 wild type is shown in grey dotted line, the negative control BM110 spb1 deletion mutant in solid grey line. B) Gene expression of the upstream glycosyl transferase (glycT, san1522) and orf as measured by qRT-PCR using specific primer couples. mRNA detectability values are shown relative to gyrA as mean ΔCq +/- SD from at least two independent cultures in triplicates.*

Part 2. - The role of PI-2b genes in pilus biosynthesis

As previously described, the PI-2b operon, similarly to other pilus loci in Gram-positive bacteria, is essentially composed of genes coding for the LPxTG pilus subunits: a pilus backbone protein Spb1, an AP2-2b ancillary pilin responsible for anchoring, and an AP1-2b putative adhesin; as well as two sortase enzymes shown to be necessary for pilus polymerization (SrtC1) and anchoring (Srt2) (Lazzarin *et al.,* 2015). However, in contrast with most other Gram-positive pilus loci, no regulatory gene could be identified in the vicinity of the locus. PI-2b also contains two overlapping additional genes, *orf* and *lep*, where *lep* codes for putative signal peptidase and *orf* is a small gene of unknown function (Figure 9).

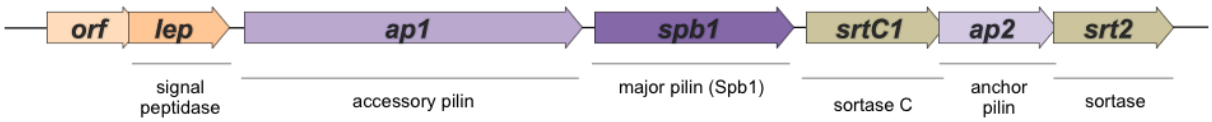


Figure 9 | Schematic representation of the PI-2b locus.

Signal peptidases in the expression of Gram-positive pili

One of the genes whose role is still entirely undetermined in the PI-2b locus is *san1520*, designated as *lep* based on its sequence similarity with a known S26A family of signal peptidases (Rosini *et al.*, 2006). The role of Lep-type signal peptidases in Gram-negative pilus formation is well-established, and a growing number of studies also associated them to pilus biosynthesis in a handful of Gram-positive bacteria. To date, three different organization have been described for *lep*-type genes relative to the locus coding for the pili whose expression they contribute to.

The first such setup is a *lep* gene coded among the structural elements of the pilus, as observed in certain pili of *Streptococcus pyogenes* and *Streptococcus pneumoniae* (in purple on Figure 3). The *lepA*-type gene of a *S. pyogenes* M49 type strain was the first Gram-positive signal peptidase to be proven essential for the polymerization and surface-anchoring of pili. In the absence of *lepA*, no pilus filaments were observed on the bacterial surface, only the monomers of the FctA major pilin. This deficient phenotype was successfully complemented with plasmid-expressed LepA, confirming the specific role of this molecule in the polymerization process (Nakata *et al.*, 2009). A homologue of LepA, SipA was subsequently described to have the same role in the M1T1 clinical strain. This suggests that

LepA/SipA function is likely to be conserved in all *S. pyogenes* strains with a pilus locus containing a such gene (Young *et al.*, 2014).

These proteins structurally belong to the Type I Signal peptidase family, which regroups enzymes that are usually responsible for the liberation of secreted proteins from the cell-wall by specifically cleaving the signal-peptide sequence of their target. Such a signal peptide has been identified in the majority of Gram-positive pilins as it is by definition necessary for their display on the cell surface which allows their processing by sortase enzymes.

In the case of *S. pyogenes* LepA/SipA, however, the Ser-Lys catalytic dyad that was shown to be essential for signal-peptide cleavage in Gram-positive bacteria, is replaced by an Asp and a Gly residue, resulting in loss of peptidase activity (Young *et al.*, 2014). Similar organization, function and characteristics were described for the SipA homologue protein of the PI-2 pneumococcal pilus locus. The authors, similarly to Young and colleagues for SipA, proposed a chaperone-like function for Sip proteins in pilus biosynthesis (Bagnoli *et al.*, 2008; Young *et al.*, 2014). Studies in a *S. suis* M6 strain revealed that SipA affected the stability of a T3 backbone pilin protein mutated at the first position of the signal peptide cleavage site, suggesting direct interaction between SipA and the pilus structure (Zahner and Scott, 2007). In these bacteria, the signal peptidase function necessary for targeting pilus subunits to the cell surface is most likely carried out by a more universal peptidase protein coded elsewhere in the genome.

A second, different case has been recently described for *Actinomyces oris* Type 1 fimbirae, where the signal peptidase that contributes to pilus surface-expression and polymerization is found coded outside of the pilus loci. A locus coding for two Type I signal peptidases have been identified in a suppressor mutant experiment, the two genes, designated as *lepB1* and *lepB2*, were revealed to be responsible for processing specific sets of surface factors. In the studied strain, LepB2 was specifically responsible for the polymerization of both FimA and FimP major pilins, coded by different loci (and LPXTG proteins in general), while the two peptidases show functional redundancy in the case of other substrates (Siegel *et al.*, 2016).

LepB2 contains the Type I signal peptidase catalytic dyad, its signal peptidase activity was confirmed and a conserved AxA cleavage site in its target protein sequences has been identified. As deletion of *lepB2* or *lepB1* did not entirely abolish the surface detectability of their targets, but not other Type I SPs have been identified in *A. oris*, it was suggested that signal peptides are still cut by proteolytic degradation activity of unidentified peptidases, resulting in ragged N-terminal target sequences. The abnormal N-terminals could perturb surface expression and polymerization of pili, explaining the decreased quantity but not entire absence of pilus polymers on the surface in a *lepB2* mutant (Siegel *et al.*, 2016).

The third type of organization is observed in the case of the GBS PI-2b locus (Rosini *et al.*, 2006) in the Pil3 pilus island of *Streptococcus gallolyticus* (UCN34 strain) (Martins *et al.*, 2015) and two predicted pilus loci of *Streptococcus suis*, FctE and FctF (Takamatsu *et al.*, 2009), where the *lep* gene is at the upstream end of the pilus operon, preceded only by an open reading frame (*orf*) of unknown function. These signal peptidase sequences show 60-70% sequence identity among each other, suggesting a conserved role for this organization in these pilus loci. The individual *orf* sequences, however, are highly dissimilar and contain no identifiable functional domains.

The SipF signal peptidase coded in the FctF *S. suis* pilus locus was confirmed to contain the Ser-Lys catalytic dyad necessary for enzymatic activity. A decrease in molecular weight pilin polymers were observed in the *sipF* mutant, suggesting that its deletion reduces polymerization efficiency. However, the export of the major pilin Sfp1 was not prevented, suggesting that other enzymes might compensate the loss of SipF function in pilin peptide-signal removal. Whether SipF acts during the biosynthesis of FctF pili as a signal peptidase specific for pilus subunits or has a different role remains to be elucidated (Fittipaldi *et al.*, 2010). In the case of the similarly organized *Streptococcus gallolyticus lep* gene, possible roles in pilus formation have not yet been addressed, although it also contains the type-I peptidase catalytic diad. To date, no other Type I signal peptidase have been identified in *S. suis*, suggesting the possible involvement of a ragged-end N-terminal proteolysis process taking over in the mutant, as it was suggested in *A. oris* (Siegel *et al.*, 2016).

In this study, we confirmed a similar role for the Lep protein coded in the PI-2b locus of the *S. agalactiae* hypervirulent strain BM110, the only known Type I signal peptidase found in this strain.

Bioinformatic analysis of *lep* and *orf*

The Lep protein coded in the BM110 PI-2b locus also contains the catalytic dyad predicted for signal peptidase functionality. The canonic bacterial Type I signal peptidase recognition and cleavage site, an AxA amino acid sequence motif (Auclair, Bahnu and Kendall, 2012), was identified in PI-2b pilins as well. The major pilin Spb1, the cleavage site falls between the two Alanines of an AFA-AE-motif at the 29-30th position, as predicted by SignalP 4.1, Signal-BLAST and PrediSi software (Petersen *et al.*, 2011; Frank and Sippl, 2008; Hiller *et al.*, 2004). For AP2-2b (San1516), the prediction was more ambiguous, as no AxA sequence is present, but all three algorithms predicted a signal peptide, with a cleavage site at the 24th position, in a VSA-DT motif, with also relatively high probability. Interestingly, no such peptide signal has been found in the sequence of AP1-2b (San1519), the putative adhesin subunit.

As this genomic organization with a *lep*-like gene overlapping with a short *orf* at the upstream end of the pilus locus is conserved in other streptococcal pilus loci, namely the Pil3 locus of *S. gallolyticus* and the SrtE and SrtF loci of *S. suis* (Figure 3), it is convenient to hypothesize that the Type 1 signal peptidase coded in these loci also plays a similar role in pilus biosynthesis as we observed for Lep in GBS. These proteins show significantly higher similarity in sequence and size with GBS Lep than *A. oris* LepB2, the sequence identity between GBS lep shares 51% identity with *S. gallolyticus* LepB (72% similarity), 58% identity (82% similarity) with *S. suis* SipF, supporting a conserved role for these genes in streptococcal pili.

Unlike in *A. oris*, where only two type I signal peptidases are coded in the genome, of which only LepB2 is specific for LPXTG proteins (Siegel *et al.*, 2016), GBS strains code for other putative Type 1 SP-s outside of the PI-2b locus. Based on structural prediction, two different type I signal peptidases are coded in the non-PI-2b carrying GBS strains, while PI-2b carrying strains code for the additional Lep coded in the pilus island. In A909, the peptidase coded by *SAK_1038* shows 27% amino acid identity with Lep (47% similarity), the other predicted SP, coded by *SAK_1731*, shows no significant sequence identity with Lep, but both are predicted to contain the consensus catalytic dyad for enzyme activity.

The role of *lep* and *orf* in PI-2b biosynthesis

In order to investigate the contribution of these genes to the expression of PI-2b, independent in-frame deletion mutants for *orf*, *lep*, *ap1-2b* and *spb1* were constructed in both BM110 and A909 and studied excessively. In the case of the overlapping *orf* and *lep* genes, integrity of the overlapping gene was verified by sequencing in the single deletion mutants. During the construction of each mutant, a so-called "back to wild-type" (bWT) clone was also preserved, which was obtained during the construction of the deletion but contained a wild type copy of the target gene. Our attempts to delete the entire PI-2b locus in order to test it role in GBS virulence *in vivo* were unsuccessful in either of our representative strains. We managed to construct an *ap1-2b-spb1* double mutant but it was heavily affected in growth and morphology on both A909 and BM110 backgrounds.

Phenotypic profiling of PI-2b gene deletion mutants

Morphology and Gram-staining

GBS usually forms chains of several attached coccid cells during growth, the length of which varies between strains and is usually correlated to growth rate, at faster growth the cells have less time to separate. By confocal microscopy, a slight reduction of chain formation was seen for the *Δspb1* mutant deleted for the major pilin and *Δlep* mutant

deleted for the putative signal peptidase in A909, compared to the wild type that forms long chains. BM110 naturally forms shorter chains; a similar effect was also noted (Figure 10). Deletion of the putative adhesin *ap1-2b*, or *orf* did not considerably alter cell shape or chain formation. None of the single mutants showed changes in Gram-staining profiles compared to the corresponding wild type, cell shape and size was identical. Surprisingly, the *Δap1-2bΔspb1* double mutants of both BM110 and A909 were shown to form aggregates and longer chains, as observed first by flow-cytometry and subsequently by confocal microscopy (Figure 10). The bWT strains obtained during the construction of either of the deletion mutants looked like the corresponding wild type strain (data not shown).

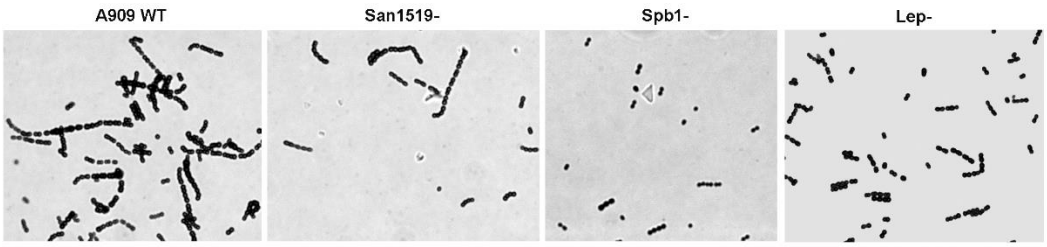


Figure 10 | Morphology of PI-2b gene deletion mutants in A909. *Bacteria from early stationary phase liquid cultures (OD_{600}=1) were stained using the conventional Gram method, and photographed under white light at 60x magnification.*

Growth in liquid medium

For mutants deleted for *lep*, the second gene of the PI-2b locus, coding for a putative signal peptidase, a slight growth defect was observed at 37°C in TH broth in BM110. In this mutant, prolonged lag- and exponential growth phases and lower culture OD at stationary phase (16h) were measured (Figure 11A). Interestingly, the same growth defect was observed in the "back to wild-type" (bWT) clones (data not shown), suggesting the presence of secondary mutations in these bacteria. Concomitantly, upon overexpression of *lep* from the pTCV*erm*P_{tet}::*lep*-StrepTag plasmid construct in the BM110*Δlep* mutant, the growth defect was not restored (Figure 11B). Under the same conditions, the *Δorf*, *Δspb1* and *Δap1-2b* mutants of both BM110 and A909 behaved exactly like the corresponding WT (and bWT) strains.

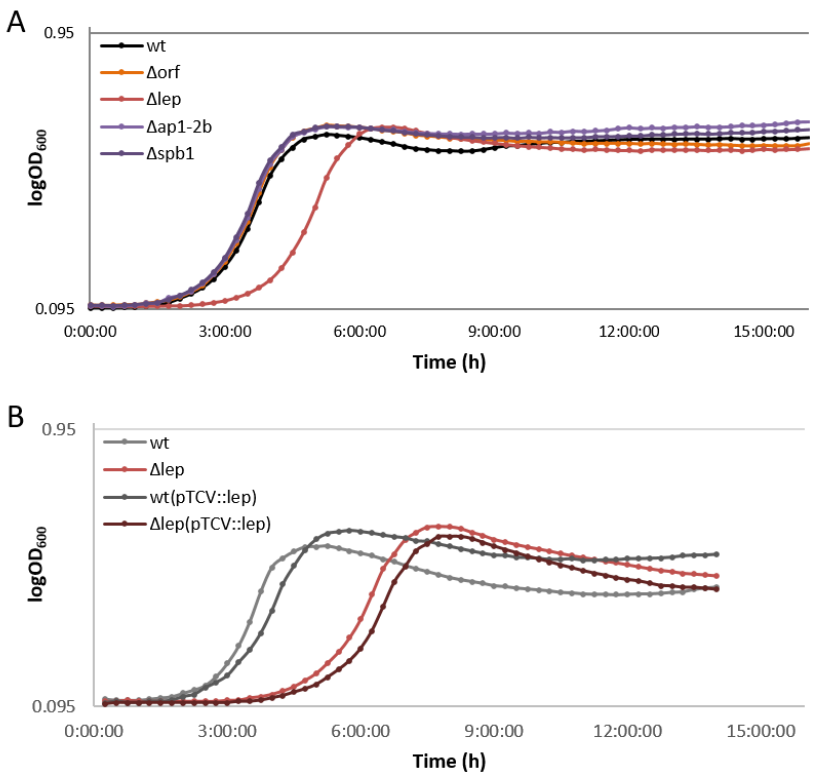


Figure 11 | Growth curves of PI-2b gene deletion mutants of BM110. *All cultures were grown from a preculture diluted to OD_{600}=0,15 in TH at 37C on a 96-well-plate, and culture optical density was measured every 15 minutes. A) Comparison of BM110 WT (black) and the different PI-2b gene deletion mutants (in corresponding colors). B) Comparison of WT (black and grew) and Δlep (red and orange) strains with or without the pTCV plasmid construct overexpressing lep from the constitutive P_{tet} promoter.*

Pigmentation and hemolysis

Slightly increased hemolytic activity for both *Δlep* and *Δorf* mutants but also the corresponding bWT strains was observed on 5% horse blood agar plates, and increased pigmentation on Granada plates. Complemented strains *Δlep(*pTCV*erm*P_{tet}::*lep*-StrepTag) and *Δorf*(pTCV*erm*P_{tet}::*orf*-HAtag) showed the same increased level of pigmentation and hemolysis as the original *Δlep* or *Δorf* mutants. These phenotypes remained unchanged in the wild type strains upon the introduction of either overexpression construct, reinforcing the hypothesis that these effects are unrelated to the role of the individual *orf* or *lep* genes, and are products of secondary mutations. Mutants deleted for *ap1-2b*, *spb1* or both were unaffected in both their pigmentation and hemolysis phenotypes compared to the corresponding wild types.

Contribution of PI-2b genes to pilus expression

The minor pilin AP1-2b is not necessary for pilus polymerization

Beside the major pilin Spb1, the PI-2b locus codes for two accessory pilins, AP1-2b and AP2-2b. AP2-2b, coded by *san1516* in BM110, has been shown to be responsible for the anchoring of pilus polymers to the cell surface (Lazzarin *et al.*, 2015), while the role in pilus assembly of the putative adhesin AP1-2b, coded by *san1519* in BM110, is unknown. We therefore analyzed expression of Spb1 in in-frame deletion mutants of *ap1-2b* available in the laboratory in both BM110 and A909 genetic backgrounds. By whole cell blotting, Spb1 expression was shown to be reduced 8-10-fold in the *Δap1-2b* mutant of both BM110 and A909 (Figure 12A).

The reduction of Spb1 surface expression in the absence of *ap1-2b* was also confirmed by flow cytometry, with a 4-6-fold decrease in the mutant strain compared to the wild type in both BM110 and A909 (Figure 12B). As PI-2b expression is generally higher in A909, subsequent experiments were carried out in this context to enable a more sensitive detection of expression changes. Western blotting of cell wall extracts shows decreased pilus backbone polymerization in the *Δap1-2b* strain; the highest molecular weight polymers detected in the wild type are absent in the mutant (Figure 12C). Blotting of concentrated supernatants from stationary phase cultures also shows an increased release of Spb1 in the supernatant of the *Δap1-2b* mutant (Figure 12C).

The transcription of *spb1* was also affected in the *Δap1-2b* strains. We detected a 2-fold decrease in *spb1* transcription by qRT-PCR in the A909*Δap1-2b* mutant (Figure 12D), which only partially explains the effect observed at protein level. As *spb1* is coded directly downstream of *ap1-2b* (Figure 5), the transcription of the rest of the PI-2b operon is most likely affected as well. Insufficient expression of the other PI-2b genes, especially that of *srtC1,* the sortase responsible for pilin polymerization, and the anchoring subunit *ap2-2b* would correlate with the presence of shorter pili, and increased Spb1 release in the supernatant in the *Δap1-2b* mutant.

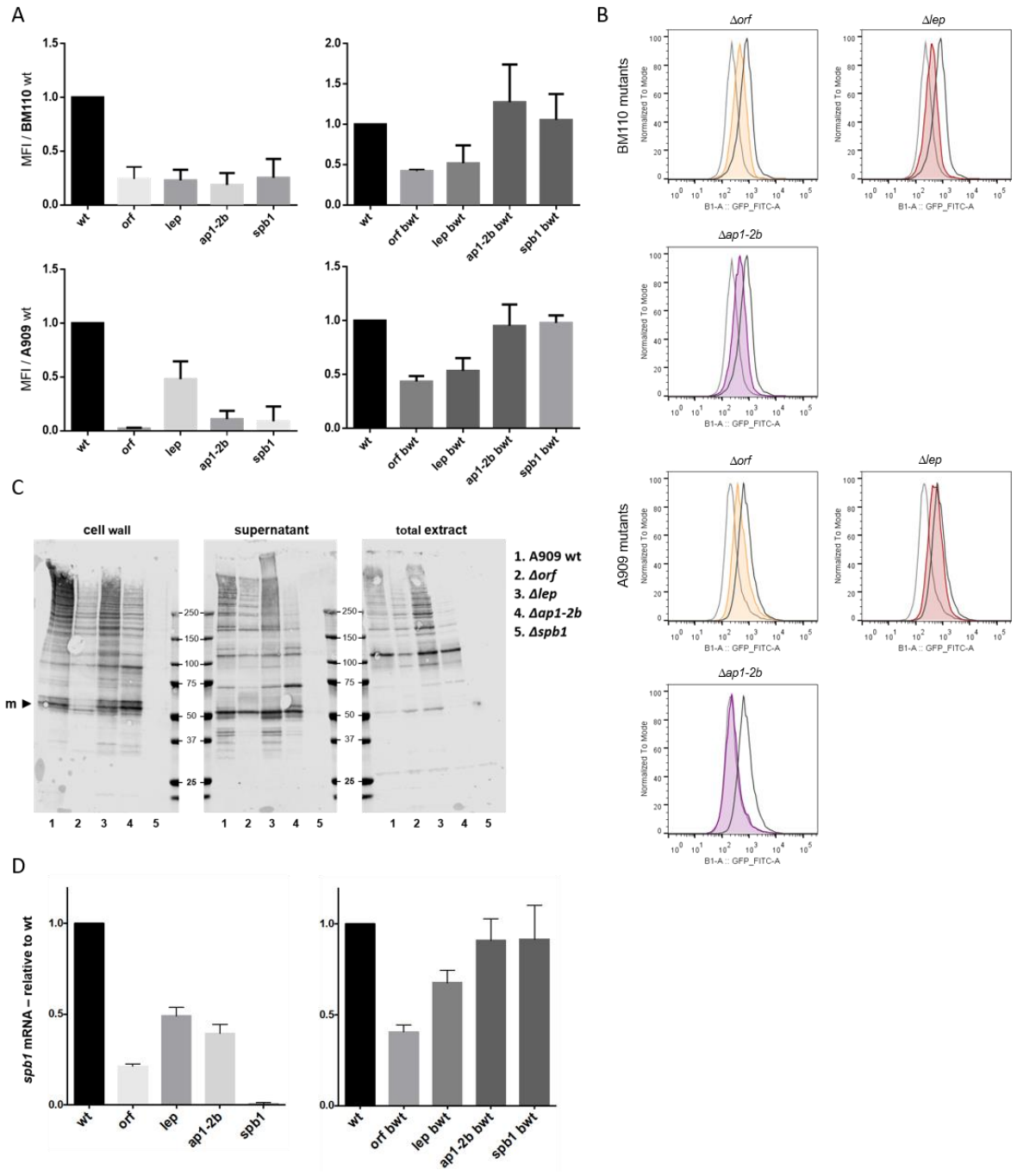


Figure 12 | Spb1 expression in the different PI-2b gene mutants of A909. *A) Detection of Spb1 on whole bacteria and quantification of the fluorescent signal. Results are expressed as MFI relative to the wild type strain (BM110 upper line, A909 bottom line). Each value is obtained from at least three independent cultures in triplicates. B) Spb1 detection by flow cytometry in BM110 and its mutants (upper group) and A909 and its mutants (bottom group). The X axis indicates fluorescence signal normalized to mode. On each graph, dark grey line indicates the wild type, light grey line the negative control, and filled peak denotes the mutant. C) Western blot analysis of Spb1 expression in cell wall extracts, culture supernatant and total protein extract from S. agalactiae A909 and its mutants. Samples were separated on 4-12% gradient Criterion XT SDS-PAGE, and detected by immunoblotting with specific polyclonal anti-Spb1. Equivalent amounts (15 µg) of total protein was loaded in each well. Monomers (m) of Spb1 are indicated. D) Transcription level of spb1 as measured by qRT-PCR using specific primers. mRNA detectability values are shown relative to gyrA as reference gene and the wild type strain (BM110 or A909) as reference sample, as mean ΔΔcq +/- SD from at least three independent cultures in triplicates.*

In order to test if the effects of *ap1-2b* deletion on Spb1 expression are due to the absence of the AP1-2b protein or not, we overexpressed Spb1 under a constitutive promoter from a plasmid construct (pTCV*erm*P_{tet}::*spb1*) in the A909*Δap1-2b* strain. The A909*Δap1-2b*(pTCV*erm*P_{tet}::*spb1*) mutant expressed higher Spb1 surface levels than wild type A909 (Figure 13A and B), and complementation of the polymerization defect could also be observed (Figure 13C). These results suggest that the deletion of the entire *ap1-2b* gene, that represents the biggest gene of the PI-2b operon, destabilizes the PI-2b mRNA, leading to suboptimal pilus polymerization and anchoring, and that AP1-2b protein is not indispensable for PI-2b pilus polymerization.

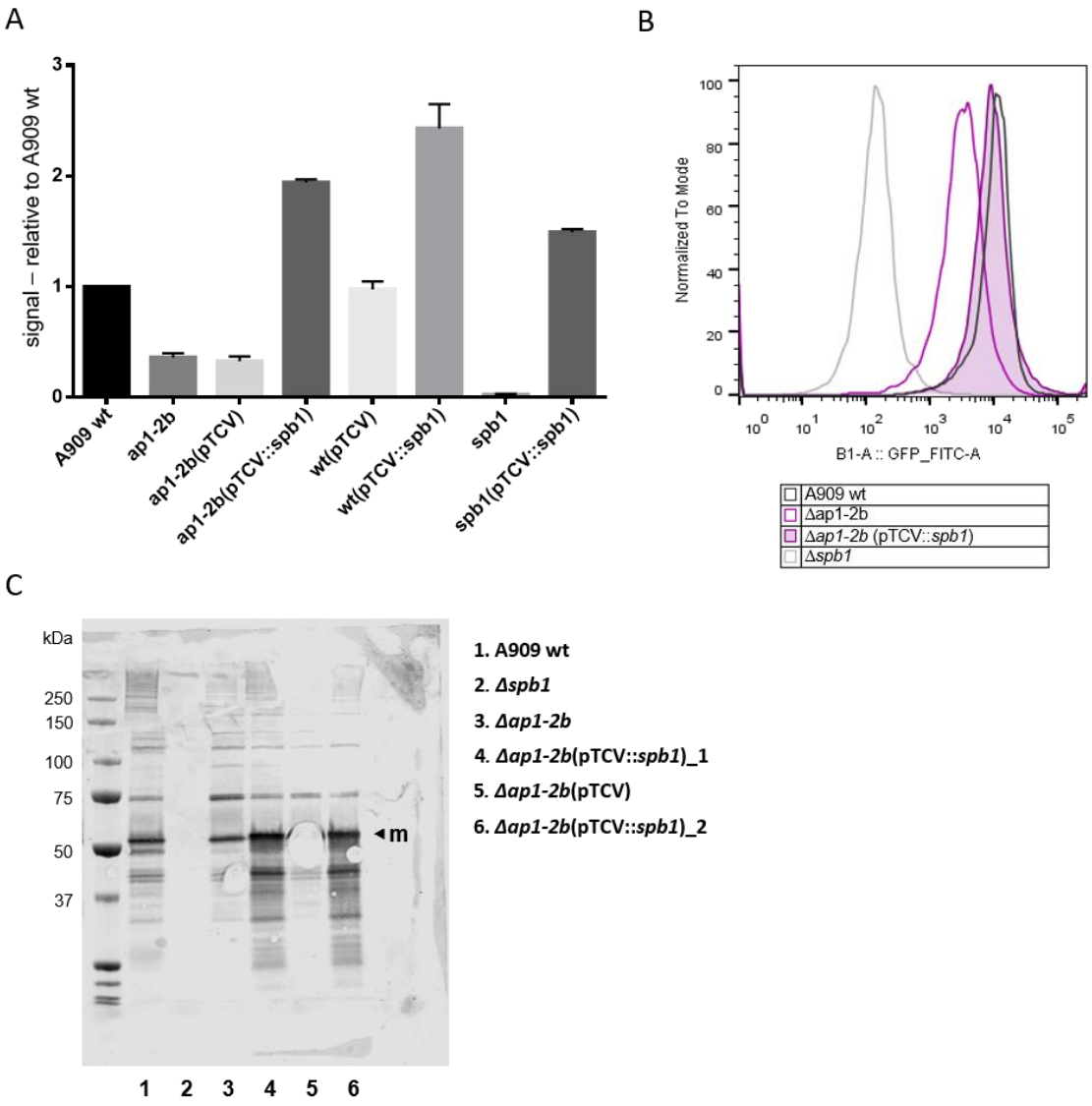


Figure 13 | Effect of spb1 overexpression in Δap1-2b deletion mutant. *A) Detection of Spb1 expression by whole cell blotting. 3µl of overnight bacterial culture were spotted on nitrocellulose membrane and blotted with polyclonal anti-Spb1 primary followed by fluorescent secondary antibodies. Results are expressed as MFI relative to wild type strain. Each value is obtained from at least three independent cultures in triplicates. B) Spb1 detection by flow cytometry in the A909 ap1-b2 mutant overexpressing (filled purple) or not (purple outline) spb1 expressed from pTCV under P_{tet}. Dark grey line indicates the wild type, light gray line the negative control. The X axis indicates fluorescence signal normalized to mode. C) Western blot analysis of Spb1 expression on cell wall extracts. Samples were separated on 4-12% gradient Criterion XT SDS-PAGE, and detected by immunoblotting with specific polyclonal anti-Spb1. Equivalent amounts (15 µg) of protein extracts were loaded in each well. Monomers (m) of Spb1 are indicated.*

PI-2b surface expression strongly affected by deletion of orf

The first gene of the PI-2b locus is an open reading frame coding for a protein of unknown function. Upon deletion of *orf*, surface detectability of both Spb1 and AP2-2b was decreased 8-10-fold in both BM110 and A909 backgrounds, compared to the corresponding wild type strains as observed by both whole cell blotting (Figure 14A) and flow-cytometry (Figure 14B). The bWT strains obtained during the construction of the orf deletion mutants also showed reduced pilus protein expression compared to the wild type, but unlike in the case of *Δlep*, expression levels in the bWT are significantly higher than in the mutants in both BM110 and A909 backgrounds (Figure 14A). Correspondingly, a 4-5-fold decrease in *spb1* transcription was detected in the *Δorf* mutants by qRT-PCR (Figure 14D), which was only partially restored in the bWT, although the integrity of the overlapping lep gene was confirmed by sequencing. The deletion of *orf* clearly affects the expression of the downstream genes, to such an extent that Spb1 surface expression was not restored in these mutants upon the introduction of the pTCV*erm*P_{tet}::*spb1* plasmid construct (Figure 14).

In order to investigate if an eventual Orf protein has an effect on pilus polymerization or PI-2b gene expression; we also constructed a pTCV*erm*P_{tet}::*orf*-HAtag plasmid vector carrying the *orf* gene under a strong constitutive promoter. The transcription of *orf* from the plasmid construct was verified by qRT-PCR, using *orf*-specific primers, but our attempts at the detection of the HA-tagged ORF protein using specific antibodies against the tag were unsuccessful. In accord with the overall decrease in PI-2b gene expression, overexpression of *orf* alone in the *Δorf* mutants did not restore Spb1 surface expression (Figure 14). Transformation of the pTCV*erm*P_{tet}::*orf*-HAtag plasmid in the wild type BM110 and A909 strains had no considerable effect on pilus protein expression. Altogether, we found no evidence of the production of an Orf protein, but our results suggest that the orf gene has an important role in the structural integrity of PI-2b transcripts and therefore in pilus expression.

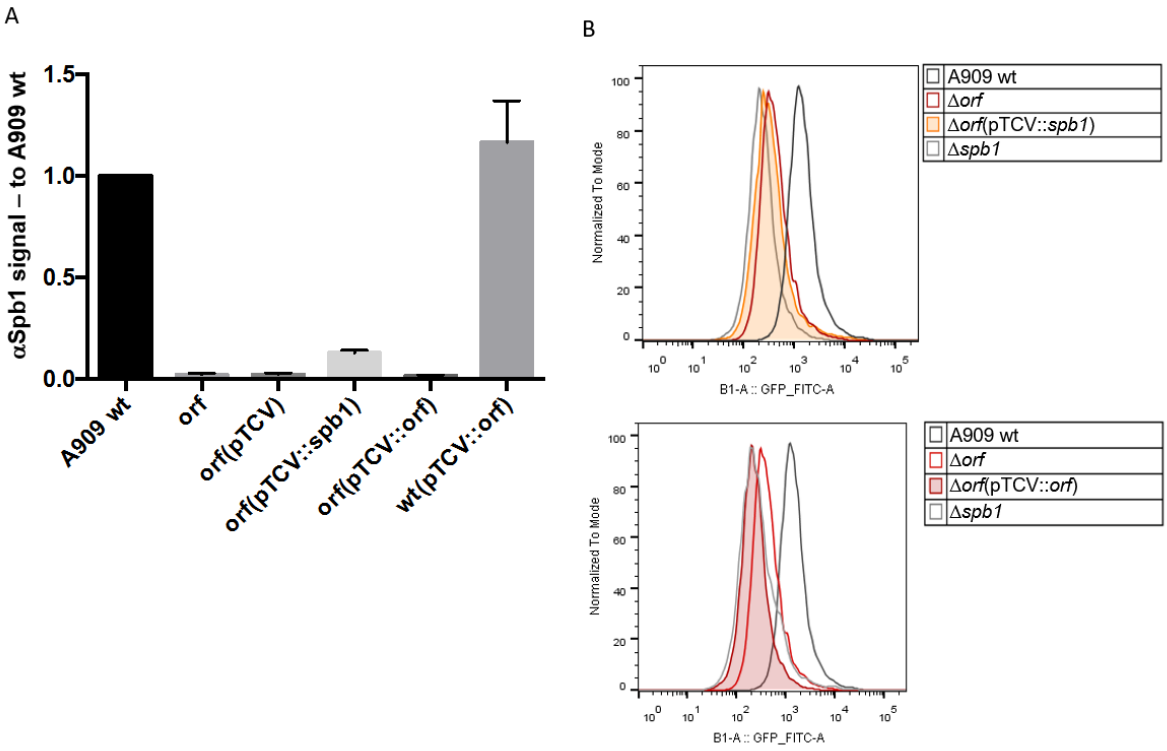


Figure 14 | Spb1 expression in the orf deletion mutant of A909. *A) Detection of Spb1 expression by whole cell blotting. 3µl of overnight bacterial culture spotted on nitrocellulose membrane, blotting with polyclonal anti-Spb1 primary, and fluorescent secondary antibodies. The values are indicated relative to the wild type strain. Each value is obtained from at least three independent cultures in triplicates. B) Spb1 detection by flow cytometry in the A909 orf mutant with (filled peak) or without (colored outline) complementation with a pTCV plasmid construct expressing either spb1 (orange) or orf (red) from the P_{tet} constitutive promoter. Dark grey line indicates the wild type, light grey the negative control in both graphs. The X axis indicates fluorescence signal normalized to mode.*

Lep is involved in the surface expression of Spb1

Lep, the second gene of the PI-2b operon, was predicted to code for a signal peptidase of the by *in silico* analysis, and contains the catalytic amino acids defined to be necessary for enzymatic activity of this protein family. Since all pilin subunits carry a peptide signal that is essential for their translocation to the cell surface, we investigated a potential role of Lep in the cleavage of the signal peptide. By whole cell blot, Spb1 surface expression in the *Δlep* mutant was shown to be 2-fold lower than wild type in both BM110 and A909 contexts (Figure 15A). This result was confirmed using flow cytometry (Figure 15B). Using qRT-PCR, we found a 2-fold decrease in *spb1* transcription in A909 *Δlep* mutant but also in the control bWT strain (data not shown). In contrast, overexpression of *spb1* from the pTCV*erm*P_{tet}::*spb1* plasmid did not restore wild type surface expression levels of Spb1 in the *Δlep* mutants indicating a role of Lep in Spb1 assembly (Figure 15A and B). Western blotting of cell wall and total protein extracts confirmed the overproduction of Spb1 in the A909*Δlep*(pTCV*erm*P_{tet}::*spb1*) strains, with an increase in Spb1 levels mainly detectable in the total extract (Figure 15C). This indicates that the absence of the Lep protein also affects PI-2b expression, beyond the possible general destabilizing effects of the *Δlep* mutation on PI-2b mRNA.

To confirm the specific role of Lep in the surface expression of Spb1, we overexpressed *lep* under a constitutive promoter from the pTCV*erm*P_{tet}::*lep*-StrepTag plasmid. Using antibodies against the Strep-tag of the Lep protein transcribed from pTCV*erm*P_{tet}::*lep*-StrepTag, overexpression of the protein was confirmed in all strains containing the plasmid. Spb1 surface expression was restored upon the introduction of pTCV*erm*P_{tet}::*lep*-StrepTag in the A909 *Δlep* strain to the levels observed in the wild type, as detected by dot blot and flow cytometry (Figure 16). Overexpression of *lep* in the wild type strain did not significantly increase pilus expression on the cell surface. Similar results were obtained with specific antibodies against the accessory pilin AP1-2b (data not shown). Although transcription and surface expression of Spb1 was not restored to wild type levels in the bWT strains obtained during the deletion of *lep*, the successful complementation of piliation in the *Δlep* mutant by a plasmid copy confirms that Lep protein activity is necessary for the optimal surface expression of PI-2b pili.

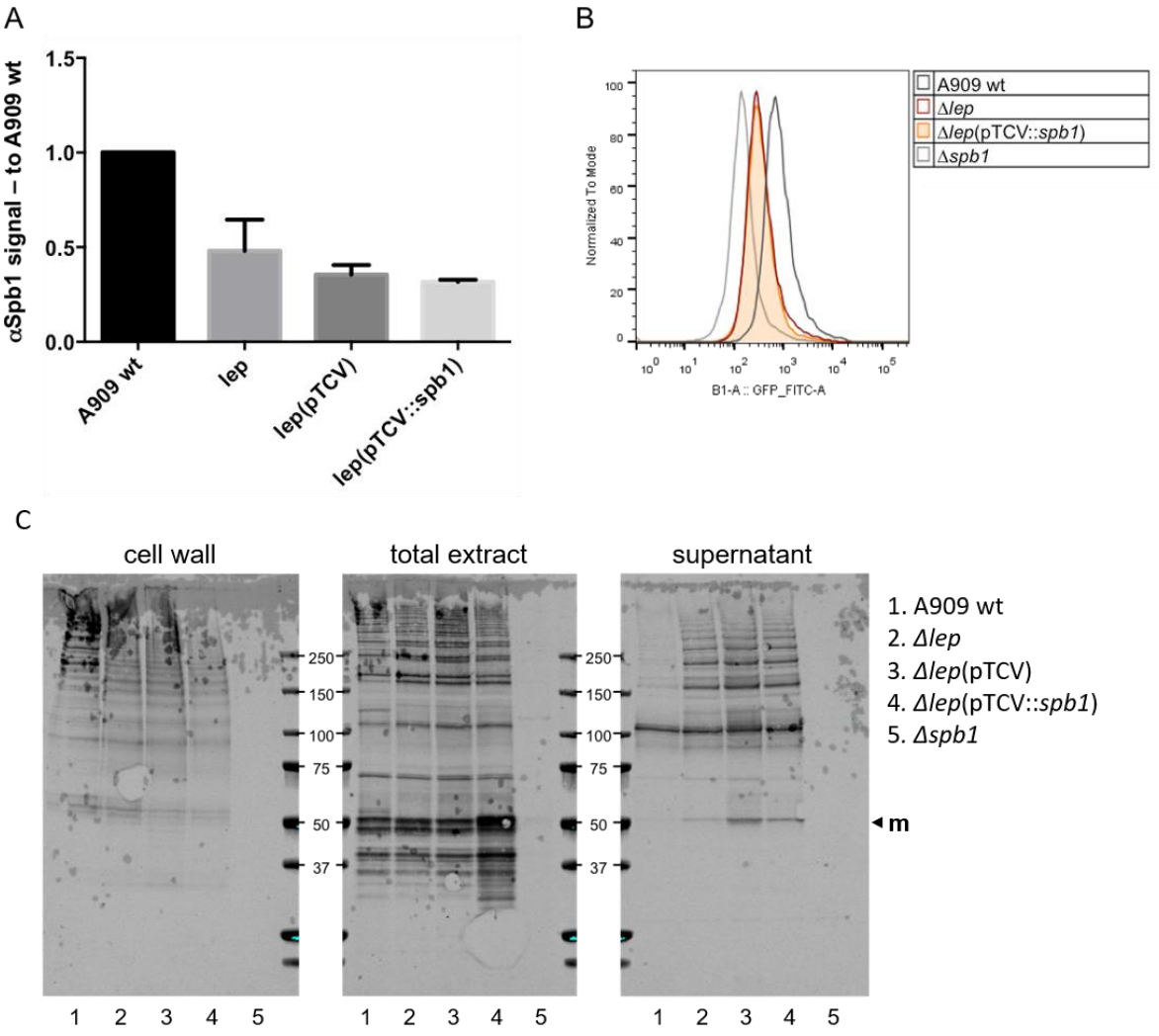


Figure 15 | Spb1 expression by the lep deletion mutant of A909. *A) Detection of Spb1 expression by whole cell blotting. 3µl of overnight bacterial culture spotted on nitrocellulose membrane, blotting with polyclonal anti-Spb1 primary, and fluorescent secondary antibodies. The values are indicated relative to the wild type strain. Each value is obtained from at least three independent cultures in triplicates. B) Spb1 detection by flow cytometry in the A909 lep mutant with (filled peak) or without (colored outline) complementation with a pTCV plasmid construct expressing spb1 from the P_{tet} constitutive promoter. Dark grey line indicates the wild type, light grey line the negative control. The X axis indicates fluorescence signal normalized to mode. C) Western blot analysis of Spb1 expression in a cell wall preparation, culture supernatant and total proteins. Samples were separated on 4-12% gradient Criterion XT SDS-PAGE, and detected by immunoblotting with specific polyclonal anti-Spb1. Equivalent amounts (15 µg) of total protein was loaded in each well. The monomers (m) of Spb1 are indicated.*

The absence of Lep, however, was not entirely deleterious for pilus expression, suggesting that the general signal peptidase is probably able to carry out this reaction in the absence of Lep. This phenotype is reminiscent of *A. oris lepB2* mutants (Siegel et al., 2016), suggesting that it has a similar role in pilus polymerization, despite the relatively low sequence identity between the two proteins. The GBS Lep and A. oris LepB2 amino acid sequences share only

29% identity (and 42% similarity) and also greatly differ in size, despite the two proteins belonging to the same protein family. Taken together, these results confirm the importance of a Type I signal peptidase in the processing of GBS PI-2b subunits, and this role is likely to be carried out at least partially by Lep.

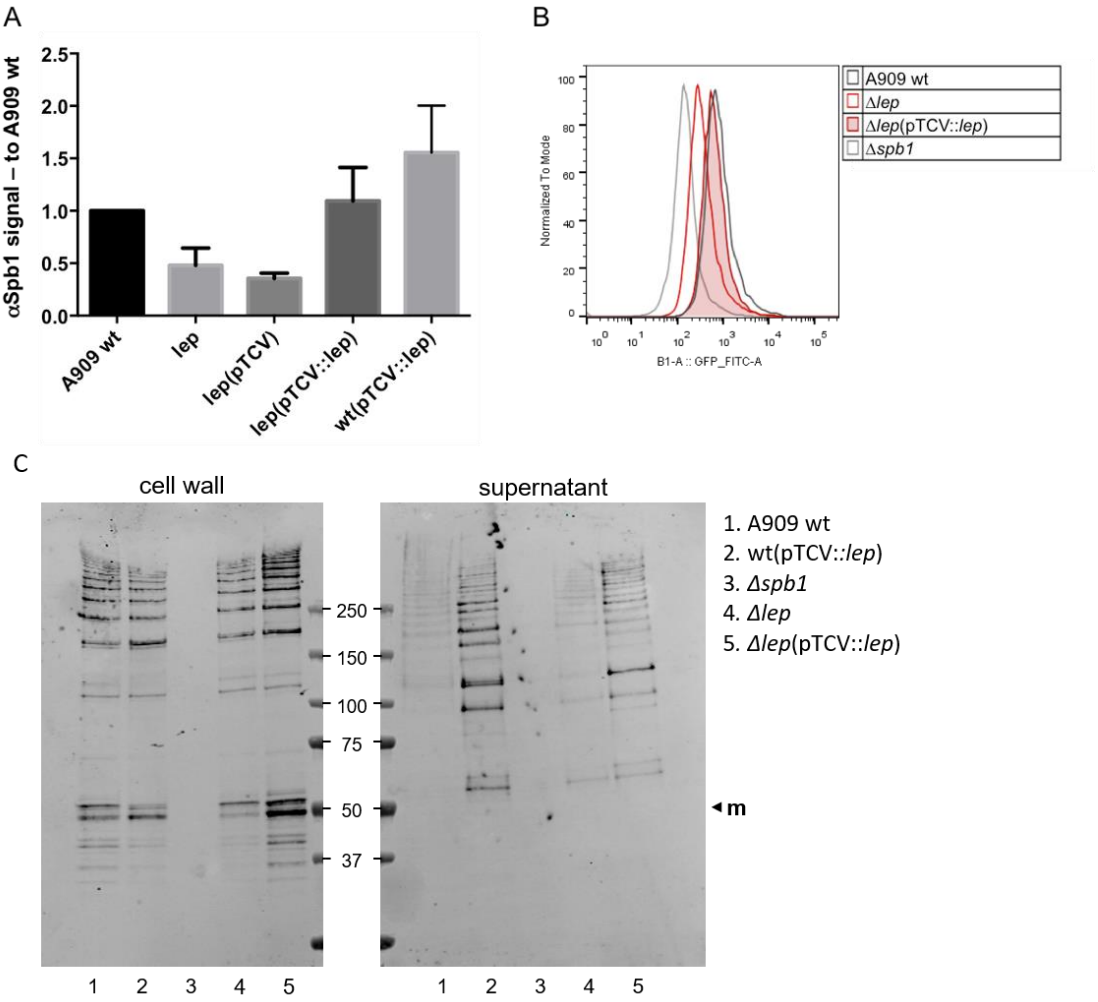


Figure 16 | Complementation of Spb1 expression in the lep deletion mutant with a plasmid construct expressing the lep gene in A909. *A) Detection of Spb1 expression by whole cell blotting. 3µl of overnight bacterial culture spotted on nitrocellulose membrane, blotting with polyclonal anti-Spb1 primary, and fluorescent secondary antibodies. The values are indicated relative to the wild type. Each value is obtained from at least three independent cultures in triplicates. B) Spb1 detection by flow cytometry in the A909 lep mutant with (filled peak) or without (colored outline) complementation with a pTCV plasmid construct expressing lep from the* P_{tet} *constitutive promoter. Dark grey line indicates the wild type, light grey line the negative control. The X axis indicates fluorescence signal normalized to mode. C) Western blot analysis of Spb1 expression in a cell wall preparation, culture supernatant and whole protein extract. Samples were separated on 4%-12% gradient Criterion XT SDS-PAGE, and detected by immunoblotting with specific polyclonal anti-Spb. Equivalent amounts (15 µg) of total protein was loaded in each well. The monomers (m) of Spb1 are indicated.*

Experimental procedures

Confocal microscopy

Bacteria were diluted in TH medium from an overnight culture to an OD_{600}=0,15 and incubated at 37°C until an OD_{600}=1 was obtained. 1ml culture was resuspended in PBS after one cycle of washing, then 10µl was spotted on glass sheets. Gram staining was carried out with a Gram-staining kit (Sigma-Aldrich) according to the manufacturer protocol. Microscopic observations were done with a Nikon Eclipse Ni-U and images acquired with a Nikon Digital Camera DS-U3. Images were treated in Image J.

Dot blots

Overnight cultures grown in TH broth at 37°C were washed twice in PBS, and after the measurement of OD_{600}, each sample was diluted in PBS to obtain an OD_{600}=1. 3µl of such dilutions were spotted on nitrocellulose membrane in triplicates, and air-dried for 15 minutes. Membranes were blocked in TBS–skimmed milk 5% and incubated for 1 hour with rabbit primary Spb1 antibodies and then with the secondary $Dylight_{800}$-coupled goat anti-rabbit antibody (Thermo Scientific Pierce). Between the two antibodies and before detection, membranes were extensively washed with TBS + 0.1% Tween 20. Bound antibodies were detected and corresponding signals were quantified using the LI-COR Odyssey Infrared Imaging System (LI-COR Biosciences).

Growth curves

Bacteria were diluted from an overnight culture grown at 37°C in TH broth to fresh TH to obtain an OD_{600} of 0,2. Triplicate sample of 200µl of each diluted culture were loaded in a sterile 96-well plate and covered with spectroscopy-compatible adhesive sheets. Incubation at 37°C and measurement of culture OD were both carried out in the Synergy HT Spectrophotometer system (BioTEK), OD_{600} was registered at 15 minute intervals for 16 hours, with 30s of slow agitation before each read.

Total protein extraction

For total protein extraction, cultures grown in TH broth at 37°C were harvested at late exponential phase (OD_{600}=0,6). 10ml of such culture was concentrated in 1ml PBS and after two cycles of washing, the bacterial pellet was resuspended in 1ml PBS and transferred into FastPrep™ tubes containing Lysing Matrix B by MP Biomedicals. Samples were subsequently processed using the *Staphylococcus aureus* protocol of the FastPrep-24™ 5G Instrument (2x40s of processing at maximum intensity). Tubes were centrifuged at 13400rpm at 4°C for 5 minutes, and the supernatant was collected and stored at -20°C for analysis.

Construction of overexpression plasmids

The selected genes, *orf* and *lep*, were amplified by PCR from the BM110 genome, using reverse primers synthesized to contain the desired tag sequence, HA-tag for orf_2_HA and the Strep-tag for lep_2_Strep, and forward primers orf_1 and lep_1, respectively. The thus obtained DNA products were subcloned in the TOPO ZeroBlunt plasmid (Thermo Scientific), and subsequently digested by the corresponding FastDigest restriction enzymes (Thermo Scientific), and ligated into the pre-digested pTCV*erm*P_{tet} constitutive expression plasmid with the T4 DNA ligase according to factory protocol (NEB). The obtained plasmids were verified by sequencing. The pTCV*erm*P_{tet}::*spb1* plasmid was constructed prior to my arrival using the same protocol.

Table 1. Strains and plasmids

Strain or plasmid	Relevant characteristics	Reference
Strains		
BM110*Δorf*	In frame deletion of *orf*	This study
BM110*Δorf* bWT	Back to the WT strain obtained during the construction of BM110*Δorf*	This study
BM110*Δlep*	In frame deletion of *lep*	This study
BM110*Δlep* bWT	Back to the WT strain obtained during the construction of BM110*Δlep*	This study
BM110*Δap1-2b*	In frame deletion of *ap1-2b (san1519)*	This study
BM110*Δap1-2b* bWT	Back to the WT strain obtained during the construction of BM110*Δap1-2b*	This study
A909*Δorf*	In frame deletion of *bp* (*sak1439*)	This study
A909*Δorf* bWT	Back to the WT strain obtained during the construction of BM110*Δorf*	This study
A909*Δlep*	In frame deletion of *lep*	This study
A909*Δlep* bWT	Back to the WT strain obtained during the construction of BM110*Δlep*	This study
A909*Δap1-2b*	In frame deletion of *ap1-2b (san1519)*	This study
A909*Δap1-2b* bWT	Back to the WT strain obtained during the construction of A909*Δap1-2b*	This study
Plasmids		
pTCV5	pTCVΩ*gfp* with 399-bp PCR product (2bUp6 / 2bUp20)[a] from BM110[b]	This study
pTCV6	pTCVΩ*gfp* with 449-bp PCR product with (2bUp6 / 2bUp12)[a] from BM110[b]	This study
pTCV*erm*P_{tet}	pTCVerm plasmid containing the P_{tet} constitutive promoter	
pTCV*erm*P_{tet}::spb1	pTCV*erm*P_{tet} with *spb1* cloned under the P_{tet} promoter	This study
pTCV*erm*P_{tet}::orf-HAtag	pTCV*erm*P_{tet} with *orf* from BM110[b] with a HA-tag cloned under P_{tet}	This study
pTCV*erm*P_{tet}::lep-StrepTag	pTCVermPtet with *lep* from BM110[b] with a Strep-tag cloned under P_{tet}	This study

[a] Primer sequences are listed in Table S2 of the publication Périchon and Szili *et al.*, 2017.

[b] Genomic DNA of the corresponding strain was used as template for sequence amplification.

Table 2. Primers

Primer	Sequence (5'-3')
For plasmids construction	
Orf_1	CCC*GGATCC*GGTTCAGTTAAGGAAGTAATCGCG[a]
Orf_2_HA	GGG*CTGCAG*CTA**GGCGTAGTGGGGGACGTCGTAGGGGTA**ATCTGTCTTTTCAT[b]
Lep_1	CCC*GGATCC*GTAGATAAGAGAGGCGTAGAC[a]
Lep_2_Strep	CC*CTGCAG*TTA**CTTTTCAAACTGTGGATGACTCCA**ATTGATACTACTAATCTTACC[c]

[a] In italics, *Bam*HI sites

[b] In italics, *Pst*I site, in bold, HA-tag sequence

[c] In italics, *Pst*I site, in bold, Strep-tag sequence

Discussion

Pili are bacterial surface appendages found in many Gram-positive bacteria. In *Streptococcus agalactiae*, two different pilus loci have been described, PI-1, which is located on a mobile genetic element, and PI-2 encoded in the core genome. This latter can be found in GBS strains as two mutually exclusive allelic variants, PI-2a and PI-2b, with PI-2a being significantly more widespread (Rosini *et al.*, 2006; Dramsi *et al.*, 2006). PI-2b is mainly present in the hypervirulent ST-17 isolates that are responsible for the majority of neonatal meningitis cases, but it is also often found in ST-61 and ST-63 bovine strains and certain fish pathogenic isolates (Brochet *et al.*, 2006; Parker *et al.*, 2016; Almeida *et al.*, 2017 accepted). While all our laboratory GBS ST-17 strains encode the PI-1 pilus, in an increasing number of recent ST-17 human isolates, this island has been replaced with antibiotic resistance genes (Teatero *et al.*, 2016; Campisi *et al.*, 2016).

The presence of a predicted hairpin structure in the promoter region of PI-2b was shown to contribute to lower PI-2b pilus expression in the BM110 ST-17 clinical strain in comparison with the A909 non-ST-17 isolate in which this sequence element is absent. This putative hairpin can be found in all of the available ST-17 isolate genomes in the NCBI database, but also in some non-ST-17 non-human pathogenic strains, such as fish isolates. Among the complete GBS genomes in the NCBI database, the three ST-17 trains that contained the PI-2b locus, namely NGBS128, NGBS572, and COH1 (the assembled BM110 chromosome is not publicly available yet), all contained the hairpin, while the only human isolate on the list that has PI-2b but is not an ST-17 was A909, which has no hairpin sequence as indicated. Other hairpin-containing isolates were mainly fish pathogens.

The presence of the hairpin has been studied in a subset of four bovine strains isolated from mastitis cases in Portugal, sequences kindly provided by A. Almeida (Institut Pasteur), where both the PI-2b locus and the upstream hairpin structure have been detected. In a pilot experiment, we compared Spb1 expression levels on the cell surface by flow cytometry and whole cell blotting, and although these isolates were generally found to express considerably higher levels of the major pilin Spb1, the expression levels were variable among isolates (data not shown). It should be noted that most bovine strains display mutations in the capsular locus and thus are non-capsulated.

Expression of pili has been shown to be regulated by different environmental conditions, especially temperature. Maximal PI-2b promoter activity at the transcriptional level was observed at 37˚C, which is the physiological optimum for GBS growth during infection.

However, when BM110 and A909 wild type strains were compared for PI-2b display on the cell surface at various temperatures, no much differences could be observed by flow cytometry, probably because of the remarkably high stability of pilus structures. We also tested Spb1 expression by flow cytometry in acidic medium (pH 5) versus neutral TH broth (pH 7.8). Overall, Spb1 expression was found to be 2-4 times lower under acidic conditions (data not shown). However, it is difficult to rely on these results since strong bacterial aggregation was observed at pH 5 which could interfere with correct analysis by flow-cytometry. In addition, GBS strains grew at a considerably slower rate at lower pH.

Among all the other environmental conditions that were tested for their effect on Spb1 (eg. minimal medium, addition of human plasma or serum, growth phase), oxidative stress was shown to have a small inhibitory effect. SpxA2 is a major regulator of oxidative stress-response in Streptococci (Zheng *et al.*, 2014; Kajfasz *et al.*, 2015; Port *et al.*, 2017), correspondingly, upon its deletion, mutants were shown to lose their resistance to oxidative stress and were impaired for growth at higher temperatures. We showed that *spb1* transcription was reduced by 50% in an A909 strain deleted for *spxA2* (data not shown). As SpxA2 cannot directly bind to DNA, its regulatory effects are indirect and rely on partner regulators (Zheng *et al.*, 2014; Kajfasz *et al.*, 2015).

Interestingly, Srt2, the PI-2b-specific sortase which anchors this pilus to the cell-wall contains two additional cysteine side-chains in addition to the catalytic Cysteine residue. We speculate that oxidative conditions could lead to the formation of intramolecular disulfide bond, which in turn could potentially inactivate this enzyme, resulting in pilus shedding in the extracellular medium.

Previous studies have indicated a role for Spb1 and PI-2b pili in the promotion of phagocytosis and prolonged intracellular survival in murine macrophages (Chattopadhyay *et al.*, 2011), using mainly GBS isolates displaying various levels of Spb1. These experiments should be repeated using the isogenic mutants that were constructed during this project (ongoing experiments). Another study suggested that deletion of *spb1* in an ST-17 clinical isolate led to a slightly decreased adhesion and invasion to several epithelial cell lines (human pulmonary A549, human colonic C2Bbe1, and cervical HeLa cells) (Adderson *et al.*, 2013). Our preliminary experiments testing adhesion to Caco-2 cells did not show significant differences (data not shown). We also tested the role of PI-2b in binding to different extracellular matrix components, including fibrinogen, fibrin, plasminogen, collagen and colonic mucus, but again we did not find significant differences when comparing WT versus pilin or adhesin mutants. However, all these experiments were carried out in standard Todd-Hewitt medium, and it would be important to test if PI-2b expression is induced *in vivo* in a murine model.

A mutant deleted for both major pilin and the putative adhesin has been successfully constructed by Bruno Périchon both in A909 and BM110 backgrounds. Despite numerous attempts, we have been unable to delete the entire PI-2b locus. We hypothesize that either PI-2b pili play a role in the cell surface organization needed for proper division or that this deletion impacts the expression of the upstream locus encoding the essential group B carbohydrate (AgB). Prior to my arrival, wild type BM110 and its *Δspb1* mutant have been compared in *vivo in* an intraperitoneal (IP) mouse infection model, but no difference in virulence was observed (data not shown). The simple and double mutants in the BM110 background will be tested for gut colonization using oral infection in mice (collaboration with the team of Pr. Claire Poyart – Institut Cochin). In vivo induction of PI-2b transcription will be addressed in this murine model using qRT-PCR in bacteria harvested from intestines or blood.

However, as ST-17 strains mainly cause CNS infections during late-onset disease and since the physiopathology of these infections is still unclear, new efforts should be directed towards a better understanding of ST-17 specific features during infection. And, in particular, how these strains are transmitted from the mother to the newborn? Finally, it would be important to test the role of PI-2b in a bovine context for the carriage status.

As the sequence of the putative pilus-associated adhesin (AP1-2b), shows conserved host-adapted sequence elements in bovine and human pathogenic strains (Springman *et al.*, 2014; Parker *et al.*, 2016), it is convenient to hypothesize host-specific roles for PI-2b pili during infection. Additionally, an increasing number of reports address the possibility of GBS transmission via breastmilk as a potential mechanism behind late onset GBS infections. Thus, it would be interesting to test the role of PI-2b for GBS growth in breastmilk.

In conclusion, PI-2b is an immunogenic proteinaceous appendage present on the bacterial surface of hypervirulent GBS ST-17 strains and thus constitute a candidate antigen for vaccine development and may be a useful molecular tool for the diagnostic of ST-17 strains in hospital settings.

References

Abbot, E.L., et al., *Pili mediate specific adhesion of Streptococcus pyogenes to human tonsil and skin.* Cell Microbiol, 2007. **9**(7): p. 1822-33.

Adderson, E.E., et al., *Subtractive hybridization identifies a novel predicted protein mediating epithelial cell invasion by virulent serotype III group B Streptococcus agalactiae.* Infect Immun, 2003. **71**(12): p. 6857-63.

Ahmed, M.S., et al., *Immune responses to pneumococcal pilus RrgA and RrgB antigens and their relationship with pneumococcal carriage in humans.* J Infect, 2014. **68**(6): p. 562-71.

Al Safadi, R., et al., *Two-component system RgfA/C activates the fbsB gene encoding major fibrinogen-binding protein in highly virulent CC17 clone group B Streptococcus.* PLoS One, 2011. **6**(2): p. e14658.

Almeida, A., et al., *Whole-Genome Comparison Uncovers Genomic Mutations between Group B Streptococci Sampled from Infected Newborns and Their Mothers.* J Bacteriol, 2015. **197**(20): p. 3354-66.

Amerighi, F., et al., *Identification of a Monoclonal Antibody Against Pneumococcal Pilus 1 Ancillary Protein Impairing Bacterial Adhesion to Human Epithelial Cells.* J Infect Dis, 2016. **213**(4): p. 516-22.

Andrade, W.A., et al., *Group B Streptococcus Degrades Cyclic-di-AMP to Modulate STING-Dependent Type I Interferon Production.* Cell Host Microbe, 2016. **20**(1): p. 49-59.

Auclair, S.M., M.K. Bhanu, and D.A. Kendall, *Signal peptidase I: cleaving the way to mature proteins.* Protein Sci, 2012. **21**(1): p. 13-25.

Bae, T. and O. Schneewind, *The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing.* J Bacteriol, 2003. **185**(9): p. 2910-9.

Bagnoli, F., et al., *A second pilus type in Streptococcus pneumoniae is prevalent in emerging serotypes and mediates adhesion to host cells.* J Bacteriol, 2008. **190**(15): p. 5480-92.

Ballard, M.S., et al., *The changing epidemiology of group B streptococcus bloodstream infection: a multi-national population-based assessment.* Infect Dis (Lond), 2016. **48**(5): p. 386-91.

Banerjee, A., et al., *Bacterial Pili exploit integrin machinery to promote immune activation and efficient blood-brain barrier penetration.* Nat Commun, 2011. **2**: p. 462.

Barnham, M., *The gut as a source of the haemolytic streptococci causing infection in surgery of the intestinal and biliary tracts.* J Infect, 1983. **6**(2): p. 129-39.

Barocchi, M.A., et al., *A pneumococcal pilus influences virulence and host inflammatory responses.* Proc Natl Acad Sci U S A, 2006. **103**(8): p. 2857-62.

Basset, A., et al., *Expression of the type 1 pneumococcal pilus is bistable and negatively regulated by the structural component RrgA.* Infect Immun, 2011. **79**(8): p. 2974-83.

Basset, A., et al., *An epigenetic switch mediates bistable expression of the type 1 pilus genes in Streptococcus pneumoniae.* J Bacteriol, 2012. **194**(5): p. 1088-91.

Basset, A., et al., *Toll-like receptor (TLR) 2 mediates inflammatory responses to oligomerized RrgA pneumococcal pilus type 1 protein.* J Biol Chem, 2013. **288**(4): p. 2665-75.

Beaussart, A., et al., *Molecular mapping of the cell wall polysaccharides of the human pathogen Streptococcus agalactiae.* Nanoscale, 2014. **6**(24): p. 14820-7.

Beckert, S., B. Kreikemeyer, and A. Podbielski, *Group A streptococcal rofA gene is involved in the control of several virulence genes and eukaryotic cell attachment and internalization.* Infect Immun, 2001. **69**(1): p. 534-7.

Bellais, S., et al., *Capsular switching in group B Streptococcus CC17 hypervirulent clone: a future challenge for polysaccharide vaccine development.* J Infect Dis, 2012. **206**(11): p. 1745-52.

Beloin, C., A. Roux, and J.M. Ghigo, *Escherichia coli biofilms.* Curr Top Microbiol Immunol, 2008. **322**: p. 249-89.

Berardi, A., et al., *Group B streptococcus late-onset disease and milk transmission.* Acta Paediatr, 2013. **102**(3): p. e95.

Bierne, H. and S. Dramsi, *Spatial positioning of cell wall-anchored virulence factors in Gram-positive bacteria.* Curr Opin Microbiol, 2012. **15**(6): p. 715-23.

Bisharat, N., et al., *Hyperinvasive neonatal group B streptococcus has arisen from a bovine ancestor.* J Clin Microbiol, 2004. **42**(5): p. 2161-7.

Blumberg, H.M., et al., *Invasive group B streptococcal disease: the emergence of serotype V.* J Infect Dis, 1996. **173**(2): p. 365-73.

Bohnsack, J.F., et al., *Serotype III Streptococcus agalactiae from bovine milk and human neonatal infections.* Emerg Infect Dis, 2004. **10**(8): p. 1412-9.

Bolduc, G.R. and L.C. Madoff, *The group B streptococcal alpha C protein binds alpha1beta1-integrin through a novel KTD motif that promotes internalization of GBS within human epithelial cells.* Microbiology, 2007. **153**(Pt 12): p. 4039-49.

Bourgogne, A., L.C. Thomson, and B.E. Murray, *Bicarbonate enhances expression of the endocarditis and biofilm associated pilus locus, ebpR-ebpABC, in Enterococcus faecalis.* BMC Microbiol, 2010. **10**: p. 17.

Brega, S., et al., *SecA localization and SecA-dependent secretion occurs at new division septa in group B Streptococcus.* PLoS One, 2013. **8**(6): p. e65832.

Brittan, J.L. and A.H. Nobbs, *Group B Streptococcus pili mediate adherence to salivary glycoproteins.* Microbes Infect, 2015. **17**(5): p. 360-8.

Brochet, M., et al., *Genomic diversity and evolution within the species Streptococcus agalactiae.* Microbes Infect, 2006. **8**(5): p. 1227-43.

Brown, W.V. and E.B. Collins, *End products and fermentation balances for lactic streptococci grown aerobically on low concentrations of glucose.* Appl Environ Microbiol, 1977. **33**(1): p. 38-42.

Budzik, J.M., et al., *Intramolecular amide bonds stabilize pili on the surface of bacilli.* Proc Natl Acad Sci U S A, 2009. **106**(47): p. 19992-7.

Buscetta, M., et al., *PbsP, a cell wall-anchored protein that binds plasminogen to promote hematogenous dissemination of group B Streptococcus.* Mol Microbiol, 2016. **101**(1): p. 27-41.

Buscetta, M., et al., *FbsC, a novel fibrinogen-binding protein, promotes Streptococcus agalactiae-host cell interactions.* J Biol Chem, 2014. **289**(30): p. 21003-21015.

Caliot, É., et al., *Role of the Group B antigen of Streptococcus agalactiae: a peptidoglycan-anchored polysaccharide involved in cell wall biogenesis.* PLoS Pathog, 2012. **8**(6): p. e1002756.

Campisi, E., et al., *Genomic Analysis Reveals Multi-Drug Resistance Clusters in Group B Streptococcus CC17 Hypervirulent Isolates Causing Neonatal Invasive Disease in Southern Mainland China.* Front Microbiol, 2016. **7**: p. 1265.

Cantey, J.B., et al., *Late and very late onset group B Streptococcus sepsis: one and the same?* World J Pediatr, 2014. **10**(1): p. 24-8.

Carlin, A.F., et al., *Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes.* J Bacteriol, 2007. **189**(4): p. 1231-7.

Chaffin, D.O., et al., *The serotype of type Ia and III group B streptococci is determined by the polymerase gene within the polycistronic capsule operon.* J Bacteriol, 2000. **182**(16): p. 4466-77.

Chang, C., et al., *Cell surface display of minor pilin adhesins in the form of a simple heterodimeric assembly in Corynebacterium diphtheriae.* Mol Microbiol, 2011. **79**(5): p.

1236-47.

Chang, Y.C. and V. Nizet, *The interplay between Siglecs and sialylated pathogens.* Glycobiology, 2014. **24**(9): p. 818-25.

Chang, Y.C., et al., *Group B Streptococcus engages an inhibitory Siglec through sialic acid mimicry to blunt innate immune and inflammatory responses in vivo.* PLoS Pathog, 2014. **10**(1): p. e1003846.

Charrel-Dennis, M., et al., *TLR-independent type I interferon induction in response to an extracellular bacterial pathogen via intracellular recognition of its DNA.* Cell Host Microbe, 2008. **4**(6): p. 543-54.

Chattopadhyay, D., et al., *Phylogenetic lineage and pilus protein Spb1/SAN1518 affect opsonin-independent phagocytosis and intracellular survival of Group B Streptococcus.* Microbes Infect, 2011. **13**(4): p. 369-82.

Chen, M., et al., *Streptococcus agalactiae isolates of serotypes Ia, III and V from human and cow are able to infect tilapia.* Vet Microbiol, 2015. **180**(1-2): p. 129-35.

R., C., A.N. E., and M.-P. E., *A note on a lytic phenomenon shown by group B streptococci.* The Australian journal of experimental biology and medical science, 1944. **22**: p. 4.

Cieslewicz, M.J., et al., *Structural and genetic diversity of group B streptococcus capsular polysaccharides.* Infect Immun, 2005. **73**(5): p. 3096-103.

Cisar, J.O., et al., *Mutants of Actinomyces viscosus T14V lacking type 1, type 2, or both types of fimbriae.* Infect Immun, 1988. **56**(11): p. 2984-9.

Phyllis, C.M., *Streptococcal Infection in Childbirth and Septic Abortion: Source of Infection and Grouping of Haemolytic Strains.* The Lancet, 1935. **226**(5858): p. 2.

Cozzi, R., et al., *New insights into the role of the glutamic acid of the E-box motif in group B Streptococcus pilus 2a assembly.* FASEB J, 2012. **26**(5): p. 2008-18.

Da Cunha, V., et al., *Streptococcus agalactiae clones infecting humans were selected and fixed through the extensive use of tetracycline.* Nat Commun, 2014. **5**: p. 4544.

Dangor, Z., et al., *Review on the association of Group B Streptococcus capsular antibody and protection against invasive disease in infants.* Expert Rev Vaccines, 2015. **14**(1): p. 135-49.

Dangwetngam, M., et al., *Serotype distribution and antimicrobial susceptibilities of Streptococcus agalactiae isolated from infected cultured tilapia (Oreochromis niloticus) in Thailand: Nine-year perspective.* J Med Microbiol, 2016. **65**(3): p. 247-54.

Danne, C. and S. Dramsi, *Pili of gram-positive bacteria: roles in host colonization.* Res Microbiol, 2012. **163**(9-10): p. 645-58.

Danne, C., et al., *Single cell stochastic regulation of pilus phase variation by an attenuation-like mechanism.* PLoS Pathog, 2014. **10**(1): p. e1003860.

Danne, C., et al., *Molecular characterization of a Streptococcus gallolyticus genomic island encoding a pilus involved in endocarditis.* J Infect Dis, 2011. **204**(12): p. 1960-70.

Davies, H.D., et al., *Multilocus sequence typing of serotype III group B streptococcus and correlation with pathogenic potential.* J Infect Dis, 2004. **189**(6): p. 1097-102.

De Angelis, G., et al., *The Streptococcus pneumoniae pilus-1 displays a biphasic expression pattern.* PLoS One, 2011. **6**(6): p. e21269.

De La Rosa, M., et al., *Granada medium for detection and identification of group B streptococci.* J Clin Microbiol, 1983. **18**(4): p. 779-85.

De Mouy, D., et al., *Antibiotic susceptibility and mechanisms of erythromycin resistance in clinical isolates of Streptococcus agalactiae: French multicenter study.* Antimicrob Agents Chemother, 2001. **45**(8): p. 2400-2.

DebRoy, S., et al., *Riboswitches. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator.* Science, 2014. **345**(6199): p. 937-40.

Delannoy, C.M., et al., *Genomic comparison of virulent and non-virulent Streptococcus agalactiae in fish.* J Fish Dis, 2016. **39**(1): p. 13-29.

Deng, L., et al., *Characterization of the linkage between the type III capsular polysaccharide and the bacterial cell wall of group B Streptococcus.* J Biol Chem, 2000. **275**(11): p. 7497-504.

Devi, A.S. and K. Ponnuraj, *Cloning, expression, purification and ligand binding studies of novel fibrinogen-binding protein FbsB of Streptococcus agalactiae.* Protein Expr Purif, 2010. **74**(2): p. 148-55.

Dillon, H.C., et al., *Anorectal and vaginal carriage of group B streptococci during pregnancy.* J Infect Dis, 1982. **145**(6): p. 794-9.

Dramsi, S., et al., *Assembly and role of pili in group B streptococci.* Mol Microbiol, 2006. **60**(6): p. 1401-13.

Dramsi, S., et al., *Rga, a RofA-like regulator, is the major transcriptional activator of the PI-2a pilus in Streptococcus agalactiae.* Microb Drug Resist, 2012. **18**(3): p. 286-97.

Dramsi, S., et al., *Epidemiologically and clinically relevant Group B Streptococcus isolates do not bind collagen but display enhanced binding to human fibrinogen.* Microbes Infect, 2012. **14**(12): p. 1044-8.

Dramsi, S., P. Trieu-Cuot, and H. Bierne, *Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria.* Res Microbiol, 2005. **156**(3): p. 289-97.

Easmon, C.S., *Group B streptococcus.* Infect Control, 1986. **7**(2 Suppl): p. 135-7.

Edwards, M.S., et al., *Capsular sialic acid prevents activation of the alternative complement pathway by type III, group B streptococci.* J Immunol, 1982. **128**(3): p. 1278-83.

Ekelund, K., et al., *Emergence of invasive serotype VIII group B streptococcal infections in Denmark.* J Clin Microbiol, 2003. **41**(9): p. 4442-4.

Emaneini, M., et al., *Comparison of virulence factors and capsular types of Streptococcus agalactiae isolated from human and bovine infections.* Microb Pathog, 2016. **91**: p. 1-4.

Farley, M.M., *Group B streptococcal disease in nonpregnant adults.* Clin Infect Dis, 2001. **33**(4): p. 556-61.

Farley, M.M., et al., *A population-based assessment of invasive disease due to group B Streptococcus in nonpregnant adults.* N Engl J Med, 1993. **328**(25): p. 1807-11.

Faro, S., *Group B beta-hemolytic streptococci and puerperal infections.* Am J Obstet Gynecol, 1981. **139**(6): p. 686-9.

Filleron, A., et al., *Group B streptococci in milk and late neonatal infections: an analysis of cases in the literature.* Arch Dis Child Fetal Neonatal Ed, 2014. **99**(1): p. F41-7.

Finch, L.A. and D.R. Martin, *Human and bovine group B streptococci: two distinct populations.* J Appl Bacteriol, 1984. **57**(2): p. 273-8.

Firon, A., et al., *Extracellular nucleotide catabolism by the Group B Streptococcus ectonucleotidase NudP increases bacterial survival in blood.* J Biol Chem, 2014. **289**(9): p. 5479-89.

Firon, A., et al., *The Abi-domain protein Abx1 interacts with the CovS histidine kinase to control virulence gene expression in group B Streptococcus.* PLoS Pathog, 2013. **9**(2): p. e1003179.

Fittipaldi, N., et al., *Mutations in the gene encoding the ancillary pilin subunit of the Streptococcus suis srtF cluster result in pili formed by the major subunit only.* PLoS One, 2010. **5**(1): p. e8426.

Flores-Mireles, A.L., et al., *Urinary tract infections: epidemiology, mechanisms of infection and treatment options.* Nat Rev Microbiol, 2015. **13**(5): p. 269-84.

Frank, K. and M.J. Sippl, *High-performance signal peptide prediction based on sequence alignment techniques.* Bioinformatics, 2008. **24**(19): p. 2172-6.

M., F.R., *Fatal Infections by Haemolytic Streptococcus Group B.* The Lancet, 1938. **231**(5969): p. 3.

Gaspar, A.H. and H. Ton-That, *Assembly of distinct pilus structures on the surface of*

Corynebacterium diphtheriae. J Bacteriol, 2006. **188**(4): p. 1526-33.

Gianfaldoni, C., et al., *Streptococcus pneumoniae pilus subunits protect mice against lethal challenge.* Infect Immun, 2007. **75**(2): p. 1059-62.

Glaser, P., et al., *Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease.* Mol Microbiol, 2002. **45**(6): p. 1499-513.

Granok, A.B., et al., *The RofA binding site in Streptococcus pyogenes is utilized in multiple transcriptional pathways.* J Bacteriol, 2000. **182**(6): p. 1529-40.

Grifantini, R., et al., *Peroxide stimulon and role of PerR in group A Streptococcus.* J Bacteriol, 2011. **193**(23): p. 6539-51.

Gutekunst, H., B.J. Eikmanns, and D.J. Reinscheid, *Analysis of RogB-controlled virulence mechanisms and gene repression in Streptococcus agalactiae.* Infect Immun, 2003. **71**(9): p. 5056-64.

Gutekunst, H., B.J. Eikmanns, and D.J. Reinscheid, *The novel fibrinogen-binding protein FbsB promotes Streptococcus agalactiae invasion into epithelial cells.* Infect Immun, 2004. **72**(6): p. 3495-504.

Handley, P.S., P.L. Carter, and J. Fielding, *Streptococcus salivarius strains carry either fibrils or fimbriae on the cell surface.* J Bacteriol, 1984. **157**(1): p. 64-72.

Harrison, L.H., D.M. Dwyer, and J.A. Johnson, *Emergence of serotype V group B streptococcal infection among infants and adults.* J Infect Dis, 1995. **171**(2): p. 513.

Hemsley, C., et al., *MgrA, an orthologue of Mga, Acts as a transcriptional repressor of the genes within the rlrA pathogenicity islet in Streptococcus pneumoniae.* J Bacteriol, 2003. **185**(22): p. 6640-7.

Hendrickx, A.P., et al., *Expression of two distinct types of pili by a hospital-acquired Enterococcus faecium isolate.* Microbiology, 2008. **154**(Pt 10): p. 3212-23.

Hendrickx, A.P., et al., *Isopeptide bonds of the major pilin protein BcpA influence pilus structure and bundle formation on the surface of Bacillus cereus.* Mol Microbiol, 2012. **85**(1): p. 152-63.

Hensler, M.E., S. Miyamoto, and V. Nizet, *Group B streptococcal beta-hemolysin/cytolysin directly impairs cardiomyocyte viability and function.* PLoS One, 2008. **3**(6): p. e2446.

Hensler, M.E., et al., *CAMP factor is not essential for systemic virulence of Group B Streptococcus.* Microb Pathog, 2008. **44**(1): p. 84-8.

Hiller, K., et al., *PrediSi: prediction of signal peptides and their cleavage positions.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W375-9.

Hilleringmann, M., et al., *Pneumococcal pili are composed of protofilaments exposing adhesive clusters of Rrg A.* PLoS Pathog, 2008. **4**(3): p. e1000026.

Hiramune, T., et al., *Phase variation of pili of Corynebacterium pilosum.* Zentralbl Veterinarmed B, 1991. **38**(4): p. 303-5.

Honda, E. and R. Yanagawa, *Attachment of Corynebacterium renale to tissue culture cells by the pili.* Am J Vet Res, 1975. **36**(11): p. 1663-6.

Honda, E. and R. Yanagawa, *Pili-mediated attachment of Corynebacterium renale to mucous membrane of urinary bladder of mice.* Am J Vet Res, 1978. **39**(1): p. 155-8.

Hooven, T.A., et al., *The essential genome of Streptococcus agalactiae.* BMC Genomics, 2016. **17**: p. 406.

Isaka, M., et al., *The YvqE two-component system controls biofilm formation and acid production in Streptococcus pyogenes.* APMIS, 2016. **124**(7): p. 574-85.

Ito, H., E. Ono, and R. Yanagawa, *Comparison of surface hydrophobicity of piliated and non-piliated clones of Corynebacterium renale and Corynebacterium pilosum.* Vet Microbiol, 1987. **14**(2): p. 165-71.

Jakubovics, N.S., S.A. Yassin, and A.H. Rickard, *Community interactions of oral streptococci.* Adv Appl Microbiol, 2014. **87**: p. 43-110.

Jiang, S.M., et al., *Regulation of virulence by a two-component system in group B streptococcus.* J Bacteriol, 2005. **187**(3): p. 1105-13.

Jiang, S., et al., *Regulation and function of pilus island 1 in group B streptococcus.* J Bacteriol, 2012. **194**(10): p. 2479-90.

Jones, C.H., et al., *FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae.* Proc Natl Acad Sci U S A, 1995. **92**(6): p. 2081-5.

Jones, N., et al., *Multilocus sequence typing system for group B streptococcus.* J Clin Microbiol, 2003. **41**(6): p. 2530-6.

Jones, N., et al., *Enhanced invasiveness of bovine-derived neonatal sequence type 17 group B streptococcus is independent of capsular serotype.* Clin Infect Dis, 2006. **42**(7): p. 915-24.

Jørgensen, H.J., et al., *Streptococcus agalactiae in the environment of bovine dairy herds-rewriting the textbooks?* Vet Microbiol, 2016. **184**: p. 64-72.

Kajfasz, J.K., et al., *Transcription of Oxidative Stress Genes Is Directly Activated by SpxA1 and, to a Lesser Extent, by SpxA2 in Streptococcus mutans.* J Bacteriol, 2015. **197**(13): p. 2160-70.

Kang, H.J. and E.N. Baker, *Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of Streptococcus pyogenes.* J Biol Chem, 2009.

284(31): p. 20729-37.

Kang, H.J. and E.N. Baker, *Structure and assembly of Gram-positive bacterial pili: unique covalent polymers.* Curr Opin Struct Biol, 2012. **22**(2): p. 200-7.

Kang, H.J., et al., *Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure.* Science, 2007. **318**(5856): p. 1625-8.

Kang, H.J., et al., *Crystal structure of Spy0129, a Streptococcus pyogenes class B sortase involved in pilus assembly.* PLoS One, 2011. **6**(1): p. e15969.

Kang, H.J., et al., *The Corynebacterium diphtheriae shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds.* Proc Natl Acad Sci U S A, 2009. **106**(40): p. 16967-71.

Kasper, D.L., et al., *Immunodeterminant specificity of human immunity to type III group B streptococcus.* J Exp Med, 1979. **149**(2): p. 327-39.

Keefe, G.P., *Streptococcus agalactiae mastitis: a review.* Can Vet J, 1997. **38**(7): p. 429-37.

Kimura, K.R., et al., *Involvement of T6 pili in biofilm formation by serotype M6 Streptococcus pyogenes.* J Bacteriol, 2012. **194**(4): p. 804-12.

Kollmann, T.R., et al., *Innate immune function by Toll-like receptors: distinct responses in newborns and the elderly.* Immunity, 2012. **37**(5): p. 771-83.

Konto-Ghiorghi, Y., et al., *Dual role for pilus in adherence to epithelial cells and biofilm formation in Streptococcus agalactiae.* PLoS Pathog, 2009. **5**(5): p. e1000422.

Kreikemeyer, B., et al., *Group A streptococcal RofA-type global regulators exhibit a strain-specific genomic presence and regulation pattern.* Microbiology, 2002. **148**(Pt 5): p. 1501-11.

Kreikemeyer, B., et al., *Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule.* Mol Microbiol, 2001. **39**(2): p. 392-406.

Kreikemeyer, B., et al., *Genomic organization, structure, regulation and pathogenic role of pilus constituents in major pathogenic Streptococci and Enterococci.* Int J Med Microbiol, 2011. **301**(3): p. 240-51.

Krishnan, V., et al., *An IgG-like domain in the minor pilin GBS52 of Streptococcus agalactiae mediates lung epithelial cell adhesion.* Structure, 2007. **15**(8): p. 893-903.

Lachenauer, C.S., et al., *Serotypes VI and VIII predominate among group B streptococci isolated from pregnant Japanese women.* J Infect Dis, 1999. **179**(4): p. 1030-3.

Lalioui, L., et al., *The SrtA Sortase of Streptococcus agalactiae is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for*

colonization of the mouse intestine. Infect Immun, 2005. **73**(6): p. 3342-50.

Lamy, M.C., et al., *Rapid detection of the "highly virulent" group B Streptococcus ST-17 clone.* Microbes Infect, 2006. **8**(7): p. 1714-22.

Lamy, M.C., et al., *CovS/CovR of group B streptococcus: a two-component global regulatory system involved in virulence.* Mol Microbiol, 2004. **54**(5): p. 1250-68.

Lancefield, R.C., *A Serological Differentiation of Human and Other Groups of Hemolytic Streptococci.* J Exp Med, 1933. **57**(4): p. 571-95.

Lancefield, R.C. and E.H. Freimer, *Type-specific polysaccharide antigens of group B streptococci.* J Hyg (Lond), 1966. **64**(2): p. 191-203.

Lancefield, R.C. and R. Hare, *The Serological Differentiation of Pathogenic and Non-Pathogenic Strains of Hemolytic Streptococci from Parturient Women.* J Exp Med, 1935. **61**(3): p. 335-49.

Landwehr-Kenzel, S. and P. Henneke, *Interaction of Streptococcus agalactiae and Cellular Innate Immunity in Colonization and Disease.* Front Immunol, 2014. **5**: p. 519.

Lartigue, M.F., et al., *Variability of neuD transcription levels and capsular sialic acid expression among serotype III group B Streptococcus strains.* Microbiology, 2011. **157**(Pt 12): p. 3282-91.

Lauer, P., et al., *Genome analysis reveals pili in Group B Streptococcus.* Science, 2005. **309**(5731): p. 105.

Lazzarin, M., et al., *Noncanonical sortase-mediated assembly of pilus type 2b in group B Streptococcus.* FASEB J, 2015. **29**(11): p. 4629-40.

Lembke, C., et al., *Characterization of biofilm formation by clinically relevant serotypes of group A streptococci.* Appl Environ Microbiol, 2006. **72**(4): p. 2864-75.

Lembo, A., et al., *Regulation of CovR expression in Group B Streptococcus impacts blood-brain barrier penetration.* Mol Microbiol, 2010. **77**(2): p. 431-43.

Lemire, P., et al., *Role of capsular polysaccharide in Group B Streptococccus interactions with dendritic cells.* Microbes Infect, 2012. **14**(12): p. 1064-76.

Lemire, P., M. Houde, and M. Segura, *Encapsulated group B Streptococcus modulates dendritic cell functions via lipid rafts and clathrin-mediated endocytosis.* Cell Microbiol, 2012. **14**(11): p. 1707-19.

Levent, F., et al., *Early outcomes of group B streptococcal meningitis in the 21st century.* Pediatr Infect Dis J, 2010. **29**(11): p. 1009-12.

Lier, C., et al., *Analysis of the type II-A CRISPR-Cas system of Streptococcus agalactiae reveals*

distinctive features according to genetic lineages. Front Genet, 2015. **6**: p. 214.

Lin, W.J., et al., *Threonine phosphorylation prevents promoter DNA binding of the Group B Streptococcus response regulator CovR.* Mol Microbiol, 2009. **71**(6): p. 1477-95.

Lindahl, G., M. Stålhammar-Carlemalm, and T. Areschoug, *Surface proteins of Streptococcus agalactiae and related proteins in other bacterial pathogens.* Clin Microbiol Rev, 2005. **18**(1): p. 102-27.

Lindberg, B., J. Lönngren, and D.A. Powell, *Structural studies on the specific type-14 pneumococcal polysaccharide.* Carbohydr Res, 1977. **58**(1): p. 177-86.

Linke-Winnebeck, C., et al., *Structural model for covalent adhesion of the Streptococcus pyogenes pilus through a thioester bond.* J Biol Chem, 2014. **289**(1): p. 177-89.

Liu, Z., et al., *The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen group A Streptococcus.* Mol Microbiol, 2012. **86**(1): p. 140-54.

Luan, S.L., et al., *Multilocus sequence typing of Swedish invasive group B streptococcus isolates indicates a neonatally associated genetic lineage and capsule switching.* J Clin Microbiol, 2005. **43**(8): p. 3727-33.

Luo, F., S. Lizano, and D.E. Bessen, *Heterogeneity in the polarity of Nra regulatory effects on streptococcal pilus gene transcription and virulence.* Infect Immun, 2008. **76**(6): p. 2490-7.

Madoff, L.C., et al., *Phenotypic diversity in the alpha C protein of group B streptococci.* Infect Immun, 1991. **59**(8): p. 2638-44.

Maeland, J.A., et al., *Survey of immunological features of the alpha-like proteins of Streptococcus agalactiae.* Clin Vaccine Immunol, 2015. **22**(2): p. 153-9.

Magalhães, V., et al., *Group B Streptococcus hijacks the host plasminogen system to promote brain endothelial cell invasion.* PLoS One, 2013. **8**(5): p. e63244.

Maisey, H.C., et al., *A group B streptococcal pilus protein promotes phagocyte resistance and systemic virulence.* FASEB J, 2008. **22**(6): p. 1715-24.

Mandlik, A., et al., *Corynebacterium diphtheriae employs specific minor pilins to target human pharyngeal epithelial cells.* Mol Microbiol, 2007. **64**(1): p. 111-24.

Mandlik, A., et al., *Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development.* Trends Microbiol, 2008. **16**(1): p. 33-40.

Manetti, A.G., et al., *Environmental acidification drives S. pyogenes pilus expression and microcolony formation on epithelial cells in a FCT-dependent manner.* PLoS One, 2010. **5**(11): p. e13864.

Manetti, A.G., et al., *Streptococcus pyogenes pili promote pharyngeal cell adhesion and*

biofilm formation. Mol Microbiol, 2007. **64**(4): p. 968-83.

Manning, S.D., et al., *Multilocus sequence types associated with neonatal group B streptococcal sepsis and meningitis in Canada.* J Clin Microbiol, 2009. **47**(4): p. 1143-8.

Marques, M.B., et al., *Prevention of C3 deposition by capsular polysaccharide is a virulence mechanism of type III group B streptococci.* Infect Immun, 1992. **60**(10): p. 3986-93.

Martins, E.R., et al., *Distribution of pilus islands in Streptococcus agalactiae that cause human infections: insights into evolution and implication for vaccine development.* Clin Vaccine Immunol, 2013. **20**(2): p. 313-6.

Martins, E.R., J. Melo-Cristino, and M. Ramirez, *Evidence for rare capsular switching in Streptococcus agalactiae.* J Bacteriol, 2010. **192**(5): p. 1361-9.

Martins, M., et al., *Streptococcus gallolyticus Pil3 Pilus Is Required for Adhesion to Colonic Mucus and for Colonization of Mouse Distal Colon.* J Infect Dis, 2015. **212**(10): p. 1646-55.

Martins, M., et al., *The Pil3 pilus of Streptococcus gallolyticus binds to intestinal mucins and to fibrinogen.* Gut Microbes, 2016. **7**(6): p. 526-532.

Mazmanian, S.K., H. Ton-That, and O. Schneewind, *Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus.* Mol Microbiol, 2001. **40**(5): p. 1049-57.

McBride, S.M. and A.L. Sonenshein, *The dlt operon confers resistance to cationic antimicrobial peptides in Clostridium difficile.* Microbiology, 2011. **157**(Pt 5): p. 1457-65.

Meyn, L.A., et al., *Association of sexual activity with colonization and vaginal acquisition of group B Streptococcus in nonpregnant women.* Am J Epidemiol, 2002. **155**(10): p. 949-57.

Mishra, A., et al., *Sortase-catalyzed assembly of distinct heteromeric fimbriae in Actinomyces naeslundii.* J Bacteriol, 2007. **189**(8): p. 3156-65.

Mishra, A., et al., *The Actinomyces oris type 2 fimbrial shaft FimA mediates co-aggregation with oral streptococci, adherence to red blood cells and biofilm development.* Mol Microbiol, 2010. **77**(4): p. 841-54.

Mistou, M.Y., et al., *Molecular dissection of the secA2 locus of group B Streptococcus reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence.* J Bacteriol, 2009. **191**(13): p. 4195-206.

Mora, M., et al., *Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15641-6.

Morello, E., et al., *Evidence for the Sialylation of PilA, the PI-2a Pilus-Associated Adhesin of Streptococcus agalactiae Strain NEM316.* PLoS One, 2015. **10**(9): p. e0138103.

Moschioni, M., et al., *The two variants of the Streptococcus pneumoniae pilus 1 RrgA adhesin*

retain the same function and elicit cross-protection in vivo. Infect Immun, 2010. **78**(12): p. 5033-42.

Moulin, P., et al., *The Adc/Lmb System Mediates Zinc Acquisition in Streptococcus agalactiae and Contributes to Bacterial Growth and Survival.* J Bacteriol, 2016. **198**(24): p. 3265-3277.

Mu, R., et al., *Identification of CiaR Regulated Genes That Promote Group B Streptococcal Virulence and Interaction with Brain Endothelial Cells.* PLoS One, 2016. **11**(4): p. e0153891.

Mu, R., et al., *Identification of a group B streptococcal fibronectin binding protein, SfbA, that contributes to invasion of brain endothelium and development of meningitis.* Infect Immun, 2014. **82**(6): p. 2276-86.

Nakata, M., et al., *Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in Streptococcus pyogenes serotype M49.* Infect Immun, 2009. **77**(1): p. 32-44.

Nakata, M., A. Podbielski, and B. Kreikemeyer, *MsmR, a specific positive regulator of the Streptococcus pyogenes FCT pathogenicity region and cytolysin-mediated translocation system genes.* Mol Microbiol, 2005. **57**(3): p. 786-803.

Nallapareddy, S.R., et al., *Conservation of Ebp-type pilus genes among Enterococci and demonstration of their role in adherence of Enterococcus faecalis to human platelets.* Infect Immun, 2011. **79**(7): p. 2911-20.

Nelson, A.L., et al., *RrgA is a pilus-associated adhesin in Streptococcus pneumoniae.* Mol Microbiol, 2007. **66**(2): p. 329-40.

Nobbs, A.H., et al., *Sortase A utilizes an ancillary protein anchor for efficient cell wall anchoring of pili in Streptococcus agalactiae.* Infect Immun, 2008. **76**(8): p. 3550-60.

N., N. and M. R., *Sur une mammite contagieuse des vaches laitieres.* Annales de l'Institut Pasteur, 1887. **1**(1): p. 17.

O'Higgins, A.C., et al., *A clinical review of maternal bacteremia.* Int J Gynaecol Obstet, 2014. **124**(3): p. 226-9.

Okahashi, N., et al., *Pili of oral Streptococcus sanguinis bind to salivary amylase and promote the biofilm formation.* Microb Pathog, 2011. **50**(3-4): p. 148-54.

Orrskog, S., et al., *Pilus adhesin RrgA interacts with complement receptor 3, thereby affecting macrophage function and systemic pneumococcal disease.* MBio, 2012. **4**(1): p. e00535-12.

Pailhories, H., R. Quentin, and M.F. Lartigue, *The transcription of the neuD gene is stronger in serotype III group B streptococci strains isolated from cerebrospinal fluid than in strains isolated from vagina.* FEMS Microbiol Lett, 2013. **349**(1): p. 71-5.

Palmer, R.J., et al., *Coaggregation-mediated interactions of streptococci and actinomyces detected in initial human dental plaque.* J Bacteriol, 2003. **185**(11): p. 3400-9.

Paoletti, L.C., R.A. Ross, and K.D. Johnson, *Cell growth rate regulates expression of group B Streptococcus type III capsular polysaccharide.* Infect Immun, 1996. **64**(4): p. 1220-6.

Paoletti, L.J., J. Bradford, and L.C. Paoletti, *A serotype VIII strain among colonizing group B streptococcal isolates in Boston, Massachusetts.* J Clin Microbiol, 1999. **37**(11): p. 3759-60.

Papasergi, S., et al., *The GBS PI-2a pilus is required for virulence in mice neonates.* PLoS One, 2011. **6**(4): p. e18747.

Papasergi, S., et al., *Analysis of the Streptococcus agalactiae exoproteome.* J Proteomics, 2013. **89**: p. 154-64.

Park, S.E., S. Jiang, and M.R. Wessels, *CsrRS and environmental pH regulate group B streptococcus adherence to human epithelial cells and extracellular matrix.* Infect Immun, 2012. **80**(11): p. 3975-84.

Parker, R.E., et al., *Association between genotypic diversity and biofilm production in group B Streptococcus.* BMC Microbiol, 2016. **16**: p. 86.

Patti, J.M. and M. Höök, *Microbial adhesins recognizing extracellular matrix macromolecules.* Curr Opin Cell Biol, 1994. **6**(5): p. 752-8.

Pérez-Pascual, D., et al., *RovS and its associated signaling peptide form a cell-to-cell communication system required for Streptococcus agalactiae pathogenesis.* MBio, 2015. **6**(1).

Périchon, B., et al., *Regulation of PI-2b Pilus Expression in Hypervirulent Streptococcus agalactiae ST-17 BM110.* PLoS One, 2017. **12**(1): p. e0169840.

Petersen, T.N., et al., *SignalP 4.0: discriminating signal peptides from transmembrane regions.* Nat Methods, 2011. **8**(10): p. 785-6.

Pezzicoli, A., et al., *Pilus backbone contributes to group B Streptococcus paracellular translocation through epithelial cells.* J Infect Dis, 2008. **198**(6): p. 890-8.

Pierno, M., et al., *FbsA-driven fibrinogen polymerization: a bacterial "deceiving strategy".* Phys Rev Lett, 2006. **96**(2): p. 028108.

Pietrocola, G., et al., *FbsA, a fibrinogen-binding protein from Streptococcus agalactiae, mediates platelet aggregation.* Blood, 2005. **105**(3): p. 1052-9.

Port, G.C., et al., *SpxA1 and SpxA2 Act Coordinately To Fine-Tune Stress Responses and Virulence in Streptococcus pyogenes.* MBio, 2017. **8**(2).

Poyart, C., et al., *Regulation of D-alanyl-lipoteichoic acid biosynthesis in Streptococcus*

agalactiae involves a novel two-component regulatory system. J Bacteriol, 2001. **183**(21): p. 6324-34.

Poyart, C., et al., *Contribution of Mn-cofactored superoxide dismutase (SodA) to the virulence of Streptococcus agalactiae.* Infect Immun, 2001. **69**(8): p. 5098-106.

Poyart, C., et al., *Attenuated virulence of Streptococcus agalactiae deficient in D-alanyl-lipoteichoic acid is due to an increased susceptibility to defensins and phagocytic cells.* Mol Microbiol, 2003. **49**(6): p. 1615-25.

Poyart, C., et al., *Characterization of the Tn916-like transposon Tn3872 in a strain of abiotrophia defectiva (Streptococcus defectivus) causing sequential episodes of endocarditis in a child.* Antimicrob Agents Chemother, 2000. **44**(3): p. 790-3.

Pritchard, D.G., B.M. Gray, and H.C. Dillon, *Characterization of the group-specific polysaccharide of group B Streptococcus.* Arch Biochem Biophys, 1984. **235**(2): p. 385-92.

Pritzlaff, C.A., et al., *Genetic basis for the beta-haemolytic/cytolytic activity of group B Streptococcus.* Mol Microbiol, 2001. **39**(2): p. 236-47.

Quach, D., et al., *The CiaR response regulator in group B Streptococcus promotes intracellular survival and resistance to innate immune defenses.* J Bacteriol, 2009. **191**(7): p. 2023-32.

Ragunathan, P., et al., *Metal binding is critical for the folding and function of laminin binding protein, Lmb of Streptococcus agalactiae.* PLoS One, 2013. **8**(6): p. e67517.

Rajagopal, L., *Understanding the regulation of Group B Streptococcal virulence factors.* Future Microbiol, 2009. **4**(2): p. 201-21.

Rajagopal, L., et al., *Regulation of cytotoxin expression by converging eukaryotic-type and two-component signalling mechanisms in Streptococcus agalactiae.* Mol Microbiol, 2006. **62**(4): p. 941-57.

Reiss, A., et al., *Bacterial pore-forming cytolysins induce neuronal damage in a rat model of neonatal meningitis.* J Infect Dis, 2011. **203**(3): p. 393-400.

Richards, V.P., et al., *Comparative genomics and the role of lateral gene transfer in the evolution of bovine adapted Streptococcus agalactiae.* Infect Genet Evol, 2011. **11**(6): p. 1263-75.

Rinaudo, C.D., et al., *Specific involvement of pilus type 2a in biofilm formation in group B Streptococcus.* PLoS One, 2010. **5**(2): p. e9216.

Rodriguez-Granger, J., et al., *Non-haemolytic and non-pigmented group b streptococcus, an infrequent cause of early onset neonatal sepsis.* Pathog Dis, 2015. **73**(9): p. ftv089.

Rosa-Fraile, M., S. Dramsi, and B. Spellerberg, *Group B streptococcal haemolysin and*

pigment, a tale of twins. FEMS Microbiol Rev, 2014. **38**(5): p. 932-46.

Rosa-Fraile, M., et al., *Identification of serum and urine proteins responsible for enhanced pigment production by group B streptococci as amylases.* Clin Diagn Lab Immunol, 1996. **3**(5): p. 594-6.

Rosch, J.W., et al., *Convergence of regulatory networks on the pilus locus of Streptococcus pneumoniae.* Infect Immun, 2008. **76**(7): p. 3187-96.

Rosini, R., et al., *Identification of novel genomic islands coding for antigenic pilus-like structures in Streptococcus agalactiae.* Mol Microbiol, 2006. **61**(1): p. 126-41.

Rosinski-Chupin, I., et al., *Reductive evolution in Streptococcus agalactiae and the emergence of a host adapted lineage.* BMC Genomics, 2013. **14**: p. 252.

Ross, R.A., L.C. Madoff, and L.C. Paoletti, *Regulation of cell component production by growth rate in the group B Streptococcus.* J Bacteriol, 1999. **181**(17): p. 5389-94.

Rubens, C.E., et al., *Identification of cpsD, a gene essential for type III capsule expression in group B streptococci.* Mol Microbiol, 1993. **8**(5): p. 843-55.

Rubens, C.E., et al., *Transposon mutagenesis of type III group B Streptococcus: correlation of capsule expression with virulence.* Proc Natl Acad Sci U S A, 1987. **84**(20): p. 7208-12.

Saar-Dover, R., et al., *D-alanylation of lipoteichoic acids confers resistance to cationic peptides in group B streptococcus by increasing the cell wall density.* PLoS Pathog, 2012. **8**(9): p. e1002891.

Salloum, M., et al., *Diversity of prophage DNA regions of Streptococcus agalactiae clonal lineages from adults and neonates with invasive infectious disease.* PLoS One, 2011. **6**(5): p. e20256.

Samen, U., et al., *Rga is a regulator of adherence and pilus formation in Streptococcus agalactiae.* Microbiology, 2011. **157**(Pt 8): p. 2319-27.

Santi, I., et al., *CsrRS regulates group B Streptococcus virulence gene expression in response to environmental pH: a new perspective on vaccine development.* J Bacteriol, 2009. **191**(17): p. 5387-97.

Santi, I., et al., *BibA induces opsonizing antibodies conferring in vivo protection against group B Streptococcus.* J Infect Dis, 2009. **200**(4): p. 564-70.

Santi, I., et al., *BibA: a novel immunogenic bacterial adhesin contributing to group B Streptococcus survival in human blood.* Mol Microbiol, 2007. **63**(3): p. 754-67.

Schubert, A., et al., *The fibrinogen receptor FbsA promotes adherence of Streptococcus agalactiae to human epithelial cells.* Infect Immun, 2004. **72**(11): p. 6197-205.

Schubert, A., et al., *A fibrinogen receptor from group B Streptococcus interacts with fibrinogen by repetitive units with novel ligand binding sites.* Mol Microbiol, 2002. **46**(2): p. 557-69.

Schuchat, A., *Group B streptococcus.* Lancet, 1999. **353**(9146): p. 51-6.

Scott, J.R. and D. Zähner, *Pili with strong attachments: Gram-positive bacteria do it differently.* Mol Microbiol, 2006. **62**(2): p. 320-30.

Seo, H.S., et al., *Characterization of fibrinogen binding by glycoproteins Srr1 and Srr2 of Streptococcus agalactiae.* J Biol Chem, 2013. **288**(50): p. 35982-96.

Seo, H.S., et al., *Binding of glycoprotein Srr1 of Streptococcus agalactiae to fibrinogen promotes attachment to brain endothelium and the development of meningitis.* PLoS Pathog, 2012. **8**(10): p. e1002947.

Seo, H.S., Y.Q. Xiong, and P.M. Sullam, *Role of the serine-rich surface glycoprotein Srr1 of Streptococcus agalactiae in the pathogenesis of infective endocarditis.* PLoS One, 2013. **8**(5): p. e64204.

Sharma, P., et al., *Role of pilus proteins in adherence and invasion of Streptococcus agalactiae to the lung and cervical epithelial cells.* J Biol Chem, 2013. **288**(6): p. 4023-34.

Sheen, T.R., et al., *Serine-rich repeat proteins and pili promote Streptococcus agalactiae colonization of the vaginal tract.* J Bacteriol, 2011. **193**(24): p. 6834-42.

Siegel, S.D., C. Wu, and H. Ton-That, *A Type I Signal Peptidase Is Required for Pilus Assembly in the Gram-Positive, Biofilm-Forming Bacterium Actinomyces oris.* J Bacteriol, 2016. **198**(15): p. 2064-73.

Sillanpää, J., et al., *A collagen-binding adhesin, Acb, and ten other putative MSCRAMM and pilus family proteins of Streptococcus gallolyticus subsp. gallolyticus (Streptococcus bovis Group, biotype I).* J Bacteriol, 2009. **191**(21): p. 6643-53.

Six, A., et al., *Srr2, a multifaceted adhesin expressed by ST-17 hypervirulent Group B Streptococcus involved in binding to both fibrinogen and plasminogen.* Mol Microbiol, 2015. **97**(6): p. 1209-22.

Six, A., et al., *Molecular Characterization of Nonhemolytic and Nonpigmented Group B Streptococci Responsible for Human Invasive Infections.* J Clin Microbiol, 2016. **54**(1): p. 75-82.

Spellerberg, B., et al., *Identification of genetic determinants for the hemolytic activity of Streptococcus agalactiae by ISS1 transposition.* J Bacteriol, 1999. **181**(10): p. 3212-9.

Spellerberg, B., et al., *Lmb, a protein with similarities to the LraI adhesin family, mediates attachment of Streptococcus agalactiae to human laminin.* Infect Immun, 1999. **67**(2): p.

871-8.

Springman, A.C., et al., *Pilus distribution among lineages of group b streptococcus: an evolutionary and clinical perspective.* BMC Microbiol, 2014. **14**: p. 159.

Springman, A.C., et al., *Selection, recombination, and virulence gene diversity among group B streptococcal genotypes.* J Bacteriol, 2009. **191**(17): p. 5419-27.

Stålhammar-Carlemalm, M., L. Stenberg, and G. Lindahl, *Protein rib: a novel group B streptococcal cell surface protein that confers protective immunity and is expressed by most strains causing invasive infections.* J Exp Med, 1993. **177**(6): p. 1593-603.

Stoner, T.D., et al., *Group B streptococcal infection and activation of human astrocytes.* PLoS One, 2015. **10**(6): p. e0128431.

Swierczynski, A. and H. Ton-That, *Type III pilus of corynebacteria: Pilus length is determined by the level of its major pilin subunit.* J Bacteriol, 2006. **188**(17): p. 6318-25.

Sørensen, U.B., et al., *Emergence and global dissemination of host-specific Streptococcus agalactiae clones.* MBio, 2010. **1**(3).

Takamatsu, D., et al., *Genetic organization and preferential distribution of putative pilus gene clusters in Streptococcus suis.* Vet Microbiol, 2009. **138**(1-2): p. 132-9.

Tazi, A., et al., *Group B Streptococcus surface proteins as major determinants for meningeal tropism.* Curr Opin Microbiol, 2012. **15**(1): p. 44-9.

Tazi, A., et al., *The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates.* J Exp Med, 2010. **207**(11): p. 2313-22.

Tazi, A., et al., *[A breakthrough in the understanding of neonatal group B streptococcus meningitis].* Med Sci (Paris), 2011. **27**(4): p. 362-4.

Teatero, S., et al., *Clonal Complex 17 Group B Streptococcus strains causing invasive disease in neonates and adults originate from the same genetic pool.* Sci Rep, 2016. **6**: p. 20047.

Telford, J.L., et al., *Pili in gram-positive pathogens.* Nat Rev Microbiol, 2006. **4**(7): p. 509-19.

Tenenbaum, T., et al., *Adherence to and invasion of human brain microvascular endothelial cells are promoted by fibrinogen-binding protein FbsA of Streptococcus agalactiae.* Infect Immun, 2005. **73**(7): p. 4404-9.

Tenenbaum, T., et al., *Streptococcus agalactiae invasion of human brain microvascular endothelial cells is promoted by the laminin-binding protein Lmb.* Microbes Infect, 2007. **9**(6): p. 714-20.

Tettelin, H., et al., *Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome".* Proc Natl Acad Sci U S A, 2005.

102(39): p. 13950-5.

Tettelin, H., et al., *Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae.* Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12391-6.

Ton-That, H., L.A. Marraffini, and O. Schneewind, *Protein sorting to the cell wall envelope of Gram-positive bacteria.* Biochim Biophys Acta, 2004. **1694**(1-3): p. 269-78.

Ton-That, H. and O. Schneewind, *Assembly of pili on the surface of Corynebacterium diphtheriae.* Mol Microbiol, 2003. **50**(4): p. 1429-38.

van Sorge, N.M., et al., *The group B streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier.* J Infect Dis, 2009. **199**(10): p. 1479-87.

Vengadesan, K., et al., *Purification, crystallization and halide phasing of a Streptococcus agalactiae backbone pilin GBS80 fragment.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. **66**(Pt 12): p. 1666-9.

Vitiello, L., et al., *Immunoregulation through extracellular nucleotides.* Blood, 2012. **120**(3): p. 511-8.

Vornhagen, J., et al., *Bacterial Hyaluronidase Promotes Ascending GBS Infection and Preterm Birth.* MBio, 2016. **7**(3).

Wang, N.Y., et al., *Group B streptococcal serine-rich repeat proteins promote interaction with fibrinogen and vaginal colonization.* J Infect Dis, 2014. **210**(6): p. 982-91.

Weerkamp, A.H., H.C. van der Mei, and R.S. Liem, *Structural properties of fibrillar proteins isolated from the cell surface and cytoplasm of Streptococcus salivarius (K+) cells and nonadhesive mutants.* J Bacteriol, 1986. **165**(3): p. 756-62.

Weidenmaier, C., S.A. Kristian, and A. Peschel, *Bacterial resistance to antimicrobial host defenses-an emerging target for novel antiinfective strategies?* Curr Drug Targets, 2003. **4**(8): p. 643-9.

Weisner, A.M., et al., *Characterization of group B streptococci recovered from infants with invasive disease in England and Wales.* Clin Infect Dis, 2004. **38**(9): p. 1203-8.

Whidbey, C., et al., *A hemolytic pigment of Group B Streptococcus allows bacterial penetration of human placenta.* J Exp Med, 2013. **210**(6): p. 1265-81.

Xia, F.D., et al., *Capsular polysaccharide of Group B Streptococcus mediates biofilm formation in the presence of human plasma.* Microbes Infect, 2015. **17**(1): p. 71-6.

Yamaguchi, T. and N. Matsunoshita, *Isolation and some properties of fimbriae of oral Streptococcus intermedius.* Curr Microbiol, 2004. **49**(1): p. 59-65.

Yancey, M.K., et al., *The accuracy of late antenatal screening cultures in predicting genital group B streptococcal colonization at delivery.* Obstet Gynecol, 1996. **88**(5): p. 811-5.

Yeung, M.K., et al., *Identification of a gene involved in assembly of Actinomyces naeslundii T14V type 2 fimbriae.* Infect Immun, 1998. **66**(4): p. 1482-91.

Yeung, M.K. and P.A. Ragsdale, *Synthesis and function of Actinomyces naeslundii T14V type 1 fimbriae require the expression of additional fimbria-associated genes.* Infect Immun, 1997. **65**(7): p. 2629-39.

Young, P.G., et al., *Structural conservation, variability, and immunogenicity of the T6 backbone pilin of serotype M6 Streptococcus pyogenes.* Infect Immun, 2014. **82**(7): p. 2949-57.

Young, P.G., et al., *Structure and activity of Streptococcus pyogenes SipA: a signal peptidase-like protein essential for pilus polymerisation.* PLoS One, 2014. **9**(6): p. e99135.

Zähner, D. and J.R. Scott, *SipA is required for pilus formation in Streptococcus pyogenes serotype M3.* J Bacteriol, 2008. **190**(2): p. 527-35.

Zangwill, K.M., A. Schuchat, and J.D. Wenger, *Group B streptococcal disease in the United States, 1990: report from a multistate active surveillance system.* MMWR CDC Surveill Summ, 1992. **41**(6): p. 25-32.

Zheng, C., et al., *Two Spx regulators modulate stress tolerance and virulence in Streptococcus suis serotype 2.* PLoS One, 2014. **9**(9): p. e108197.

Zong, Y., et al., *A 'Collagen Hug' model for Staphylococcus aureus CNA binding to collagen.* EMBO J, 2005. **24**(24): p. 4224-36.

Annexes