UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

Design of strategies to prevent synthesis of *S. pneumoniae* capsular polysaccharide at the bacteria division septum

Joana Silva Figueiredo

DISSERTAÇÃO

MESTRADO EM BIOLOGIA MOLECULAR E GENÉTICA

2012

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

Design of strategies to prevent synthesis of *S. pneumoniae* capsular polysaccharide at the bacteria division septum

Dissertação orientada por Dr. Sérgio Filipe (ITQB/UNL) e

Prof. Mário Santos (FCUL)

Joana Silva Figueiredo

MESTRADO EM BIOLOGIA MOLECULAR E GENÉTICA

2012

**INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
/UNL**

Design of strategies to prevent synthesis of *S. pneumoniae* capsular polysaccharide at the bacteria division septum

Joana Silva Figueiredo

MASTER THESIS

2012

This thesis was performed at the Bacterial Cell Surfaces and Pathogenesis laboratory, located in the ITQB/UNL institute, under the direct supervision of Dr. Sérgio Filipe.

Prof. Mário Santos was the internal designated supervisor in the scope of the *Master in Genetics and Molecular Biology* of the Faculty of Sciences of the University of Lisbon.

Acknowledgments

The completion of this dissertation was only possible with the support of many to whom I wish to thank.

Working in the Bacterial Cell Surfaces and Pathogenesis laboratory was both an honor and a pleasure. I would like to thank Dr. Sérgio Filipe, my external supervisor, for giving me the opportunity to work in his lab and especially for helping me whenever I needed, encouraging me to move forward, and offering me his teachings, ideas and suggestions. I would also like to thank to Professor Mário Santos, my internal supervisor, for supporting my thesis and for the attention and suggestions given. I also thank to Dr. Mariana Pinho for her suggestions and to the researchers of Bacterial Cell Biology Laboratory, with whom I worked side by side. Finally, I want to thank ITQB for receiving me and Faculdade de Ciências from Universidade de Lisboa for being my second home for the last five years.

To Mafalda for being both my "overseer", my friend and for teaching me how to work with these particular bacteria, *pneumococcus.* For her patience, even when her time was short and for the full readiness showed for helping and motivating me. To Maria João for sharing her wisdom, helping me in defeat and treasuring my victories.

To the remaining members of the BCSP lab (Filipa, Gonçalo, Magda, Rita, Tatiana and Vânia) for always providing a good working environment and sharing a good mood. For the jokes and strength to overcome the hard times, while providing continuous learning.

I also wish to express my gratitude towards Cristina Mendes, who as a teacher and a friend who taught me how to truly work in a laboratory, and how to think as a scientist.

I cannot end without thanking to my parents, grandmother and family (Golias included) for always supporting and guiding me so I could follow my heart, even when I doubted myself. To the friends I have made during my academic years for always being by my side and whose concern and strength has helped me to evolve in every aspects of my being/myself. Therefore, Carlota, David, Francisco Rafael, Sofia and Vânia, and many others, a sincere thank you. To my "sister" Catarina and "second parents" Palmira and Álvaro for everything and to Joana for being there for me and for being a "bad influence" whenever I was needing to chill out.

To you all, for the encouragement, trust and unconditional support, I dedicate this thesis.

Abstract

Keywords: *Streptococcus pneumoniae;* capsular polysaccharide synthesis; bacterial two-hybrid system; fluorescent proteins.

Streptococcus pneumoniae is a common respiratory bacterial pathogen and a frequent cause of community-acquired pneumonia in developed countries. The genes encoding for the capsule polysaccharide (CPS), one of the most important virulence factors of these bacteria, are organized in an operon and in almost all the serotypes the two conserved Wzd and Wze proteins are expressed. Previous results suggest that if these two proteins cannot interact, forming a Wzd/Wze protein complex, pneumococcal bacteria will be prevented from producing capsule at the septum, which was shown to abolish the ability of these bacteria to cause bacteremia in mice after intranasal challenge.

In this work, we aimed to find a method capable of screening and identifying small inhibitory (SI) peptides that prevent the interaction between Wzd and Wze. This could, consequently, represent a breakthrough in the development of strategies to replace vaccines against this important clinical pathogen.

Initially, a derivative of the *Escherichia coli* bacterial two-hybrid assay was used. Here, T25- and T18-tagged proteins Wzd and Wze were expressed in the presence of a protein that should compete and interfere with their interaction. However, expression of this control protein, untagged and fully functional Wze, did not prevent the interaction between T25-Wzd and T18-Wze when expressed in a different plasmid.

Afterwards we decided to screen for SI peptides directly in *S. pneumoniae*. For that purpose, we constructed a mutant strain that encodes in the chromosome both proteins, Wzd and Wze, functional and fused to different fluorescent proteins. Accordingly, we observed that Wzd and Wze were localized at the division septum of bacteria and that this localization was lost when a competitor was expressed from a replicative plasmid. We will now screen for SI peptides that can cause delocalization of Wzd and/or Wze and determine their effect on the synthesis of pneumococcal CPS.

Palavras-chave: *Streptococcus pneumoniae;* polissacarídeo capsular; bacterial two-hybrid; proteínas fluorescentes;

Streptococcus pneumoniae é um agente bacteriano patogénico que causa frequentemente pneumonia em países desenvolvidos. Um dos factores de virulência mais importantes é a cápsula (CPS), um polissacárido que reveste as bactérias. Os genes que codificam para a síntese da cápsula encontram-se organizados num operão que contém os genes *wzd* e *wze*, conservados em quase todos os serotipos conhecidos. As proteínas Wzd e Wze interagem formando um complexo proteico que é recrutado para o septo, local de divisão da bactéria, induzindo e regulando a síntese da cápsula. Resultados anteriores sugerem que a inibição da ligação entre estas duas proteínas pode impedir a produção de CPS no septo. Por esta razão, a descoberta de pequenos péptidos, denominados de péptidos SI (pequenos péptidos inibitórios – small inibitory peptides), que inibam a interacção entre as proteínas Wzd e Wze, pode significar uma revolução na criação de estratégias alternativas para substituir as vacinas desenvolvidas contra este patogéneo.

O objectivo deste trabalho consistiu no desenvolvimento de um método de identificação de péptidos SI. Começou-se por usar um derivado do sistema "bacterial two-hybrid", em *Escherichia coli*, em que ambas as proteínas Wzd e Wze, contendo os tags T25 e T18 respectivamente, são expressas na presença de uma proteína competidora capaz de inibir a interacção Wzd/Wze. Contudo, este método não se revelou o mais adequado.

De seguida decidiu-se desenvolver um sistema alternativo que pudesse identificar péptidos SI directamente em pneumococos. Para isso, construíram-se mutantes que expressam no cromossoma os genes *wzd* e *wze* em fusão com sequências que codificam para diferentes proteínas fluorescentes CFP e Citrine, respectivamente. Esta ferramenta revelou-se capaz de identificar péptidos SI que inibam a interacção entre as proteínas Wzd e Wze. Posteriormente procuraremos outros péptidos SI capazes de deslocalizar as proteínas Wzd e/ou Wze e determinar o seu efeito na síntese de CPS de pneumococos.

Resumo

Streptococcus pneumoniae é uma bactéria Gram-positiva presente na flora do tracto respiratório superior do Homem. No entanto, quando esta bactéria se alastra para o tracto respiratório inferior, sangue ou cérebro, pode causar doenças graves, tais como meningite, septicemia, pneumonia, sinusite ou otite. Trata-se do agente etiológico mais prevalente em infecções respiratórias adquiridas, sendo a pneumonia a principal causa de morte em crianças com idade inferior a cinco anos, superando o conjunto das percentagens de mortalidade causadas pela sida, malária e tuberculose.

A superfície desta bactéria está coberta por um polissacárido que constitui a cápsula formando uma camada exterior com 200-400nm de espessura e é o principal factor de virulência de *S. pneumoniae*. O papel fundamental da cápsula para a virulência e patogenicidade deste microrganismo deve-se à sua capacidade em evitar a activação do sistema do complemento e em impedir o reconhecimento de antigénios da bactéria. O acesso de componentes do sistema imunitário ao peptidoglicano, ácidos teicóicos e proteínas de superfície ao ser dificultado pela presença da cápsula, permite a evasão ao sistema imunitário do hospedeiro e a sua colonização.

O polissacárido capsular é composto por monossacáridos interligados por ligações glicosídicas e a variação dos seus componentes estabelece a heterogeneidade entre as diferentes cápsulas conhecidas em pneumococos. Hoje em dia são conhecidos noventa e três serotipos diferentes, e no caso particular do serotipo 14, que foi utilizado na realização deste trabalho, o polissacárido é composto por unidades repetidas de um tetrassacárido contendo D-glucose, N-acetil-D-glucosamina, e D-galactose.

Os genes que codificam para a síntese da cápsula estão organizados no operão *cps* entre os genes *dexB* e *aliA* e podem ser divididos em duas classes – os genes conservados entre os diferentes serótipos e os genes que são específicos de cada serotipo.

Na região 3' do operão *cps* encontra-se a região variável que codifica para proteínas específicos dos serotipos, tais como glicosiltransferases, e para duas proteínas membranares: Wzx, a flipase que transporta as unidades repetidas pela membrana plasmática e Wzy que é a polimerase que liga as unidades repetidas individuais.

Na região 5` do operão *cps* encontram-se quatro genes conservados nos vários serotipos (com excepção dos serotipos tipo 3 e 37) – *wzg, wzh, wzd* e *wze* (também denominados de *cpsA*, *cpsB*, *cpsC* e *cpsD*). Devido à conservação e especificidade dos genes conservados, estes constituem excelentes alvos para o desenvolvimento de novas terapêuticas de prevenção e de tratamento de infecção por *S.pneumoniae*.

O primeiro gene, *wzg*, codifica para a ligase que faz a ligação covalente entre a cápsula e o peptidoglicano. O gene *wzh* codifica para uma fosfatase de tirosinas fosforiladas

que, vai actuar na proteína Wze, desfosforilando-a. Os genes *wzd* e *wze*, nos quais se foca este trabalho, codificam para proteínas que funcionam como co-polimerases na biossíntese da cápsula:

- *wze* codifica para uma cinase de tirosinas, sendo uma proteína citoplasmática que tem a capacidade de se autofosforilar na presença da proteína codificada por *wzd*. Esta proteína contém domínios conservados Walker A e B que permitem a ligação de ATP. A região C-terminal tem 4 resíduos de tirosinas que podem ser fosforiladas. Inicialmente a fosforilação ocorre à custa de ATP, mas, de seguida, dá-se por transfosforilação.

- *wzd* codifica para uma proteína membranar pertencente à família das PCP (do inglês polysaccharide co-polymerase) e é necessária para que a fosforilação da proteína Wze possa ocorrer.

Actualmente, o modelo para a regulação da síntese da cápsula em *S. pneumoniae* propõe que a proteína Wze na sua forma activa, desfosforilada, interage com a proteína Wzd permitindo que o ATP se ligue aos seus domínios específicos e levando assim à interacção do complexo Wzd/Wze com outras proteínas. Neste estado, a síntese de cápsula ocorre em níveis elevados. Contudo, a presença do Wzd vai induzir a autofosforilação do Wze que desta forma vai reduzir os níveis de síntese da cápsula, por consequência regulando negativamente a produção da mesma. A posterior desfosforilação da proteína Wze pela fosfatase Wzh permite que todo este ciclo se repita.

Através de experiências realizadas anteriormente foi demonstrado que quando derivados fluorescentes das proteínas Wzd e Wze são expressos numa estirpe capsulada, conseguem localizar-se no septo, local onde as bactérias sintetizam a nova parede celular e por onde se dividem. Por outro lado, quando estas proteínas são expressas individualmente e numa estirpe não capsulada, a proteína Wzd localiza-se na membrana citoplasmática enquanto que a proteína Wze está espalhada pelo citoplasma da bactéria. No entanto, quando as proteínas Wzd e Wze são expressas conjuntamente numa estirpe capsulada mas com o operão *cps* deletado, não possuindo por isso cápsula, ambas as proteínas se localizam no septo. Isto indica que a localização do complexo Wzd/Wze não está dependente de outras proteínas codificadas no operão.

Quando os genes *wzd* ou *wze* são deletados numa estirpe capsulada, a cápsula continua a ser produzida, contudo está ausente no septo, demonstrando que a activação da síntese de cápsula neste local está dependente da interacção e subsequente localização destas duas proteínas no septo. A síntese de cápsula no septo encontra-se em coordenação com a síntese da parede celular, incluindo o peptidoglicano. Assim sendo, especula-se que as proteínas Wzd e Wze funcionam como reguladores da síntese da cápsula e que se a sua interacção for impedida não ocorrerá formação desta no septo das bactérias. Consequentemente, as bactérias poderão ficar mais susceptíveis/vulneráveis ao

reconhecimento pelo sistema imune do hospedeiro, diminuindo drasticamente a sua virulência.

Assim, o objectivo deste trabalho consistiu em desenvolver um método para identificar péptidos capazes de inibir a interacção entre as proteínas Wzd e Wze. Estes péptidos, denominados péptidos SI (pequenos péptidos inibitórios – small inibitory peptides) e expressos a partir de um plasmídeo ou importados do meio, evitariam a interacção entre Wzd e Wze e inibiriam a consequente produção de cápsula no septo, através de três modos: sequestrando a proteína Wze ou alterando a sua estrutura; bloqueando a acessibilidade da região C-terminal da proteína Wzd ou interagindo com a região externa da proteína Wzd de modo a alterar a sua conformação e a sua capacidade de interagir com a proteína Wze.

Inicialmente começou-se por testar um método derivado do sistema "Bacterial Two-Hybrid" em *Escherichia coli*. Neste ensaio construíram-se fusões proteicas entre as proteínas Wzd e Wze e os diferentes domínios, T25 e T18 respectivamente, que pertencem à proteína CyaA (uma ciclase de adenilato de *Bordetella pertussis*). Estas fusões foram expressas em conjunto com uma proteína que poderia inibir a interacção de T25-Wzd com T18-Wze. No entanto, a expressão da proteína competidora, Wze não marcado, não conseguiu inibir a interacção T25Wzd/T18Wze quando expressa num plasmídeo diferente. Apesar de vários esforços terem sido feitos de forma a alterar e melhorar o método, este não resultou para o efeito pretendido. Tal poderá dever-se ao facto da interacção entre estas duas proteínas ser muito estável ou pela ausência de outras proteínas de *S. pneumoniae* que podem ser necessárias para uma interacção dinâmica.

Por esta razão, decidiu-se desenvolver um novo método directamente em pneumococos. Para isso, foi necessário construir uma estirpe mutante que codificasse no cromossoma os genes *wzd* e *wze* ligados a sequências que codificam para proteínas fluorescentes diferentes (*wzd* ligado ao *CFP* e o *wze* ligado ao *citrine*). Desta forma seria possível visualizar a localização das proteínas Wzd-CFP e Wze-Citrine e confirmar a sua co-localização em células vivas a dividirem-se. Quando observada ao microscópio de fluorescência esta estirpe apresenta ambas as fluorescências no septo, tal como esperado. Posteriormente, ao transformar a estirpe mutante com um plasmídeo que codifica para a proteína Wze foi possível confirmar que esta actua como competidora com a proteína Wze-Citrine codificada pelo cromossoma. O sinal de fluorescência proveniente da proteína Wze-Citrine passou a estar espalhado por todo o citoplasma ao invés de se localizar apenas na zona septal. Assim, comprovou-se que o método proposto permitirá identificar novos péptidos capazes de inibir a interacção Wzd/Wze e provavelmente a síntese da cápsula no septo bacteriano.

Table of contents:

Introduction

Streptococcus pneumoniae

Streptococcus pneumoniae is an aerotolerant anaerobic Gram-positive bacterium that is a normal inhabitant of the human upper respiratory tract. However, when this cocci spreads to the inner ear, lungs, bloodstream or brain a inflammatory response arises causing several diseases that range in severity from meningitis, septicemia and pneumonia to sinusitis and acute otitis [1]. This pathogen is a major cause of invasive disease in the human population throughout the world, which is associated with high morbidity and mortality. Pneumonia is the leading cause of death in children worldwide and kills approximately 1.4 million children under the age of five years, every year – more than AIDS, malaria and tuberculosis combined (October 2011 World Health Organization). In addition, according to the Instituto Nacional de Estatística (National Statistics Institute, Portugal), pneumonia (bacterial and viral) has a 5% rate of mortality of the total deaths in Portugal, being the main cause of mortality by respiratory diseases, including lung cancer, with a rate of mortality of 32% (Relatório do Observatório Nacional das doenças respiratórias 2011).

Like other bacteria, pneumococcus is surrounded by a capsular polysaccharide (capsule or CPS), which forms a thick layer of 200-400 nm in depth [2]. This is a major and essential virulence factor that enables the bacteria to colonize the host by blocking antibody deposition, which binds to deeper cell structures such as teichoic acids and cell surface proteins, and consequently attenuates opsonophagocytosis. It also reduces the activation of the complement system [3] and the trapping of bacteria in neutrophil extracellular traps [4].

Pneumoccoci undergo a bidirectional phase variation between two colonial morphologies, described as "transparent" and "opaque", which correlate with the thickness of the CPS. The transparent form is associated with reduced levels of CPS expression and occurs during the initial colonization phase. At this stage some important pneumococcal surface structures, such as adhesins and teichoic acids, may be better exposed to external recognition allowing the establishment of these bacteria in the nasopharynx, and enhancing its capacity to cross the blood-brain barrier. The second phase, the opaque form, exhibits a massively increased virulence that seems to be correlated with an apparent increase of CPS production because of its antiphagocytic properties [5,6,7,8]. This suggests that the capacity to regulate CPS production is crucial for the survival of this pathogen in different host environments.

S. pneumoniae is also capable of exchanging genetic material by natural transformation. This process is induced by a competence-stimulating peptide (CSP) that is released by growing bacteria and its concentration is dependent on cell density, consistent

with a quorum sensing model in which a specific molecule, when present above a certain concentration threshold, modifies the metabolic status of part of the bacterial population. These characteristics can lead to the appearance of new bacteria resistant to different antibiotics and, consequently, increases the complexity of pneumococcal infection control [9]. Furthermore, the *cps* locus organization, the abundance of transposable elements at this locus and the presence of important genes to the capsule formation outside this site enhance the genetic variability of the capsule. However, there are some barriers that control the appearance of new serotypes, for example the lack of homology between serotype-specific genes in different serogroups limits recombination and the change of multiple CPS genes is required [10].

Up to date, ninety three serotypes that express structurally and antigenically different CPS have been reported to be associated with *S. pneumoniae*, but not all are associated with pneumococcal disease [11]. This association of specific serotypes with pneumococcal disease has recently influenced the development of anti-pneumococcal vaccines based on formulations of various capsular antigens [12]. A possible alternative to these vaccines is the use of drugs that can inhibit the synthesis of the capsule.

Chemical studies have shown that capsular structure and immunological specificity are correlated. The CPS is composed of basic building blocks, monosaccharides, that are linked by glycosidic bonds and its variation in type or constituent establishes the heterogeneity among capsules [13]. Serotype 14 CPS (one of the predominant serotypes in South American countries before implementation of vaccination programs [14]) has been extensively studied. In this case, the polysaccharide is composed of a tetrasaccharide repeating unit containing D-glucose, N-acetyl-D-glucosamine, and D-galactose [15]. All the different CPS, except those associated with serotypes 3 and 37 that are synthesized by the synthase pathway, are synthesized by the Wzy dependent pathway where a lipid-linked repeat unit on the intracellular face of the membrane is exported by Wzx. This protein is associated to the outer face of the cytoplasmic membrane. Here, the polysaccharide repeat unit is polymerized into its mature form and, as it was demonstrated in some serotypes (including serotype 14), covalently attached to the peptidoglycan present in the bacterial cell wall by Wzg [2,10,16]. The peptidoglycan is a heterogeneous polymer that preserves cell integrity as its main function, but also maintains the cell shape and acts as a scaffold to other elements such as proteins and teichoic acids. It is made of glycan chains of β-(1-4)-linked N-acetylglucosamine and N-acetylmuramic acid that are cross-linked by short peptides [17]. In type III *Streptococcus agalacticae* CPS is linked via an additional oligosaccharide and a phosphodiester bond to N-acetylglucosamine residues on the peptidoglycan and it is thought that in *S. pneumoniae* this mechanism could be analogous [18].

The cps *operon*

The genes coding for the synthesis of the capsular polysaccharide are located, in almost all the serotypes, at a specific region of the pneumococcal chromosome, between the *dexB* and *aliA* genes (Figure 1) [19]. This operon can be divided into two different sets of genes: a variable region at 3' end and a conserved region at 5' end. The 3' region of this locus encodes for serotype-specific genes such as enzymes that produce the different sugar nucleotide precursors for the synthesis of the capsule. These include glycosyltransferases and two membrane proteins: Wzx, a flippase that transports the repeat units across the cytoplasmic membrane, and Wzy which is a polymerase that links individual repeat units to form high-molecular weight CPS [10,20,21].
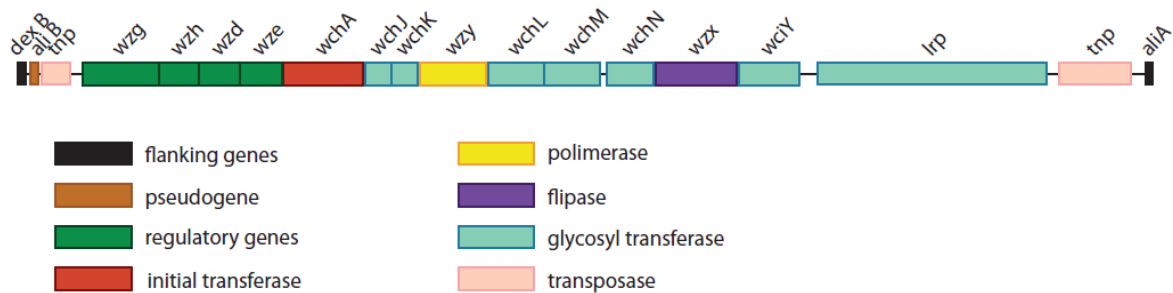
Figure 1: Representation of the *cps* gene clusters for serotype 14. Genes are represented on the forward and reverse strands by boxes coloured according to the gene key, with gene designations indicated above each box. Adapted from Bentley *et al* [10].

At the 5' region there are four genes that are conserved in all pneumococcal serotypes, except types 3 and 37: *wzg*, *wzh*, *wzd* and *wze* (also known as *cpsA*, *cpsB*, *cpsC* and *cpsD*). Because of their conserved nature it is expected that these genes encode for proteins that play general but necessary functions in the production, processing or regulation of the polysaccharide capsule [22].

The first gene of the *cps* operon, *wzg*, encodes for a member of the LytR-Cps2A-Psr (LCP) protein family (comprising Wzg, LytR and Psr) which is widespread in Gram-positive bacteria and does not occur in Gram-negative bacteria. Initially, it was suggested that these proteins were transcriptional regulators of cell wall processes because of pleiotropic phenotypes of LCP mutant strains and its homology with other transcriptional regulators [23]. However, and because these are integral membrane proteins that are enriched at mid-cell, it is not immediately apparent how they can play any role in DNA binding. Kawai *et al*. recently reported strong genetic and biochemical evidence that LCP proteins are, in fact, enzymes that catalyze the covalent attachment of anionic cell wall polymers, like teichoic acids and capsular polysaccharides, to peptidoglycan [24]. Wzg is required for the full expression of the

3

capsule, but as Eberhardt *et al.* showed all three LCP proteins seem to have a semiredundant role in the transfer of teichoic acid and/or capsular polysaccharides onto peptidoglycan [25].

The other three conserved genes encode for proteins, Wzh, Wzd and Wze that are indispensable for encapsulation of bacteria and regulation of the synthesis of the capsule. Wzh is a cytoplasmic manganese-dependent phosphotyrosine-protein phosphatase required to dephosphorylate Wze and prevent its phosphorylation [26,27,28].

The proteins that can function as co-polymerases in the biosynthesis of the capsule and extracellular polysaccharides often belong to the family of bacterial tyrosine kinases. These proteins usually have two domains: a transmembrane domain and an intracellular catalytic domain. Contrasting with Gram-negative bacteria, where the kinase is composed by only one protein, in the Gram-positive bacteria the two domains are divided into two different proteins [29,30]. The expression in Gram-positive bacteria of these two domains in two different proteins may have introduced a different level of regulation to the synthesis of the capsular polysaccharide in regulatory mechanisms for kinase function and possibly surface polysaccharide transport [31]. In *S. pneumoniae* these proteins are encoded by the third and fourth genes of the *cps* operon: *wzd* and *wze*, in which we will focus our attention in this work. Wzd, a membrane protein, constitutes the transmembrane domain and belongs to the polysaccharide co-polymerase family [32], being necessary for tyrosine phosphorylation of Wze but not for the transphosphorylation between Wze proteins [27]. Wzd has two membrane spanning hydrophobic domains and both N and C termini are located in the cytoplasm whereas the central portion is exposed on the external side of the cell membrane [33]. On the other hand, Wze, a cytoplasmic protein that constitutes the catalytic domain, is an autophosphorylating protein-tyrosine kinase, similar to Wzc from *E. coli*. It has conserved Walker A and B ATP binding motifs and a C-terminal tyrosine rich region. Multiple tyrosine residues at the C-terminal region become phosphorylated, initially via the ATP-binding domain and then via transphosphorylation. Although the binding of ATP to the protein Wze is essential for the CPS synthesis, the tyrosine phosphorylation of this protein does not appear to be necessary for capsular polysaccharide production [32]. Recently, Morona *et al.* proposed that phosphorylated Wze has a role in the attachment of CPS to the cell wall [33].

In many serotypes the fifth gene of the *cps* operon, *wchA*, codes for an enzyme that catalyse the transfer of the initial monosaccharide, in most serotypes glucose, from an UDP-glucosyl phosphate to a membrane associated lipid carrier [34]. Then a sequential transfer of monossacharides occurs to form the capsular polysaccharide repeat units until the flippase Wzx transfers these repeating units to the outer side of the cytoplasmic membrane. After that, the lipid linked CPS is formed in the external membrane by polymerization of individual

repeat units by Wzy [35,36]. Finally, the attachment of the mature CPS to the peptidoglycan is accomplished by the ligase Wzg [24].

The model for the capsule synthesis regulation that is currently accepted states that non-phosphorylated Wze interacts with Wzd allowing ATP to bind and interactions between capsule proteins to occur. This enables the capsule synthesis to proceed at its maximal level. However, with Wzd present, the autophosphorylation of Wze is then induced by a still unknown signal and when its C-terminal Tyrosine-rich region is phosphorylated the levels of capsule synthesis are slowed down, probably because the interactions with Wzy (capsule polymerase) change. This means that the phosphorylation of Wze reduces the level of encapsulation, therefore negatively regulates CPS production [32,37]. The reduction in the CPS synthesis allows the CPS polymer to be linked to the peptidoglycan by the ligase Wzg. Finally, dephosphorylation of Wze by Wzh occurs and the cycle can be repeated [1]. Assembly of high-molecular-weight capsule requires a switching between Wze active state (non-phosphorylated) and its inactive form (phosphorylated) [33].

In our laboratory we wanted to study the localization of the regulators of the capsule synthesis: Wzd and Wze. To accomplish this, genes *wzd* and *wze* were substituted by genes encoding the fluorescent derivatives of Wzd and Wze. Expression of fluorescent proteins may take place at levels that are probably similar to those observed with unlabeled proteins as these genes are cloned in the nature *cps* operon. Moreover, this method allows the expression of other proteins, encoded within the *cps* operon, suffers minimal alterations and can be used to observe localization of proteins in live and dividing encapsulated cells. Preliminary work carried out in our laboratory showed that, when expressed in unencapsulated strains, a fluorescent derivative of the protein Wzd localizes all over the cell membrane while the protein Wze fluorescent derivative is spread throughout the cytoplasm of the bacteria. However, when these two proteins were expressed separately, in an encapsulated strain ATCC6314, both proteins were recruited to the bacteria division septum early in the cell cycle, before the onset of invagination in newborn bacteria. Furthermore, co-expression of both derivative proteins in unencapsulated strains also led to its localization at the septum of the cells. This surprising result indicates that these two proteins localize at the bacterial division septum without requiring the presence of additional proteins encoded by the capsule operon[38]. Our laboratory has also determined that the interaction between the two proteins and their location in the septum is dependent on a functional ATP binding domain of Wze (Walker A domain) and the presence of ATP. Additionally, in the absence of Wzd or Wze the capsule is still produced, but it is absent from the division septum, which proves that only when Wzd and Wze interact, they can localize to the septum allowing the formation of CPS there. As a result, Henriques *et al.* proposed that Wzd and Wze are spatial-temporal regulators of capsular polysaccharide synthesis and, in the presence of ATP,

localize at the division site [38]. The formation of capsule at the bacteria division septum and its coordination with the cell wall synthesis seems to be of major importance and one possibility is that the newly synthesized cell wall, at the division site, has to be covered by a capsular polysaccharide so it is not recognized by the host immune system. Therefore, the capsule is produced in coordination with cell wall synthesis, which results in full encapsulation of the bacteria. These results suggest that if the interaction between Wzd/Wze is blocked, pneumococcal bacteria will be prevented from producing capsule at the division septum making it much more susceptible to be recognized by the host complement system or innate immune system. This is in accordance to previous reports that have shown that *S. pneumoniae* mutant strains unable to produce Wzd or Wze are not capable of causing bacteremia in mice after intranasal challenge [5,33]. Consequently, finding strategies that could result in the lack of capsule synthesis at the bacterial septum would not kill the bacteria but greatly decrease its virulence. The results obtained in our attempt to find strategies to prevent the interaction between the Wzd and Wze proteins are potentially of great importance not only for *S. pneumoniae* but also to other encapsulated pathogens.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains used in this study are listed in Table 1 of Appendix. *Escherichia coli* was routinely grown in LB or LA medium at 37ºC, unless otherwise indicated. When needed, antibiotics were used at the following concentrations: 100 µg/ml ampicillin and 50 µg/ml kanamycin. Isopropyl-b-D-thiogalactopyranoside (IPTG, Apollo Scientific) was used at 0.5 mM and 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-gal, Apollo Scientific) at 40 µg/ml.

Streptococcus pneumoniae was grown in C + Y liquid medium [39] at 37ºC, without aeration, or in trypic soy agar (TSA, Difco) supplemented with 5% sheep blood (Probiológica). Tetracycline (Sigma-Aldrich), streptomycin and erythromycin (Apollo Scientific) were used at 1µg/ml, 100 µg/ml and 0.25 µg/ml, respectively. For white/blue selection of *S. pneumoniae* colonies, X-gal was used at 120 µg/ml.

Lactococcus lactis was grown in M17 broth (Difco), supplemented with sucrose (0.5M) and glucose (0.5% w/v), at 30ºC. Erythromycin (Apollo) was used at 100 µg/ml.

DNA purification and manipulation

The preparation and subsequent transformation of *E. coli* and *S. pneumoniae* competent cells was performed as previously described [40, 41]. PCR products and plasmid DNA were purified using the WizardR SV Gel and PCR clean-up System (Promega) and WizardR Plus SV Minipreps (Promega) kits, respectively. PCR fragments were amplified using Phusion high-fidelity DNA polymerase (Finnzymes). DNA was digested with restriction enzymes purchased from New England Biolabs. DNA ligations were performed following standard molecular biology techniques using T4 DNA ligase (Fermentas). Plasmids and primers used in this study are listed in Tables 2 and 3, respectively, of Appendix.

Construction of plasmids for Bacterial Two-Hybrid Assays

To test the interaction between proteins Wze and Wzd *in vivo*, a Bacterial Two-Hybrid (BTH) assay was employed. Plasmid pBCSJF006 was constructed by amplification of the gene *wze* using the primers 1 and 2, which was then cloned into plasmid pBCSMH037. The plasmid pBCSJF007 was constructed in a similar way except that amplification of *wze* was carried out with primers 1 and 3, which added 6 histidines to the N-terminus end of Wze. The fragment of DNA encoding the genes *wzd* and *wze* was amplified from genomic DNA extracted from *S. pneumoniae* strain ATCC6314, with the primers 4 and 5 and then cloned in the plasmid pKT25. The resulting plasmid was named pBCSJF008. The *wze* gene was

amplified from the genomic DNA extracted from *S. pneumoniae* strain ATCC6314 with the primers 2 and 6 and ligated to the product of amplification of pBCSMC006 with the primers 7 and 5, producing the plasmid pBCSJF009. In this process, *Eco*RI and *Kpn*I restriction sites were added upstream and downstream of the gene *wze*, respectively. The plasmid pBCSJF010 was constructed by amplification of the gene *wzh,* using genomic DNA from *S. pneumoniae* strain ATCC6314 as template, with the primers 8 and 9. This amplified fragment was ligated to the product of amplification of pBCSJF008 with the primers 10 and 11, originating the restriction sites *Xho*I and *Kpn*I upstream and downstream of the fragment *wzh*, respectively. The nucleotide sequences of the modified regions of the constructed plasmids were confirmed by sequencing.

Bacterial Two-Hybrid Assays

Interactions between Wzd and Wze were tested by transformation of BTH101 cells with previously constructed bacterial two-hybrid plasmids. These plasmids express derivatives of Wze, Wzd and/or Wzh with T18 and T25 complementary fragments of the catalytic domain of adenylate cyclase of *Bordetella pertussis*. Plates were incubated at 30ºC and screened for blue/white colonies, in which blue indicated a positive interaction.

Construction of plasmids for protein expression in Streptococcus pneumoniae

The coding sequence of Wze was amplified using ATCC6314 genomic DNA, as template, and primers 12 and 13. The amplified fragment was then cloned downstream the improved Citrine sequence (with the first 10 amino acids of the protein Wze at the N-terminus) present in plasmid pBCSJC001[42], thus resulting in the expression of the fluorescent derivative iCitrine-Wze. The resulting plasmid was named pBCSJF001[42]. For expression of Citrine-Wze, *wze* was amplified with primers 12 and 13 and cloned into pBCSMH002, to produce plasmid pBCSJF002[42]. Plasmids pBCSJF003 and pBCSJF004, which allowed the expression of iCFP-Wzd and CFP-Wzd, respectively, were constructed through amplification of *wzd* with primers 16 and 17 and cloning in plasmids pBCSMH031 and pBCSMH018, respectively. In order to generate plasmid pBCSJF005, the gene *wzd* was amplified by PCR with primers 14 and 15, and then cloned in the plasmid pBCSMH004. This plasmid encodes for the fluorescent derivative Wzd-CFP. Screening of positive transformants was carried out by PCR using primer pairs 18 and 19. The coding sequences for all fusion proteins were confirmed by sequencing.

Substitution of capsule wild-type genes for fluorescent derivatives

In order to construct the mutant strain encoding the CFP fluorescent derivatives of the protein Wzd encoded in the *S. pneumoniae* (strain ATCC6314*wze::wzecitrine*) genome the plasmid pORI280 was used. For the construction of the strain BCSJF012 it was necessary to construct the plasmid pBCSJF011 by the following way: primers 20 and 21 were used to amplify the 3' part of *wzh* (~500bp), using the ATCC6314 chromosomal DNA as template. Primers 22 and 23 were used to amplify the entire coding sequence of CFP and the 5' part of the gene *wzd* (~500bp) from plasmid pBCSJF003. Primers 21 and 22 possess an overlapping region of 33 base pairs, allowing the above fragments to be joined by overlap PCR using the primers 20 and 23. The product of this reaction was restricted with *Bam*HI and *Eco*RI, cloned into plasmid pORI280. For the construction of strain BCSJF013 a fragment comprising the 3' terminal of the gene *wzd* (~500bp) and the entire CFP coding sequence was amplified with primers 24 and 25 from plasmid pBCSJF005. The downstream region of *wzd* (~750bp) was amplified with primers 26 and 27, from the chromosomal DNA of strain ATCC6314. Primers 25 and 26 contain an overlapping region of 39 base pairs so that the two fragments could be joined by overlap PCR. The resulting fragment was restricted with *Bam*HI and *Eco*RI and cloned in pORI280, originating the plasmid pBCSJF012. Plasmids were routinely propagated in *Lactococcus lactis* LL108 and purified before being transformed into the encapsulated ATCC6314 *S. pneumoniae* strain (serotype 14). The strains BCSJF012 and BCSJF013 were obtained by excision of the plasmid as previously described [43].

Fluorescence Microscopy visualization

For fluorescence microscopy visualization, *S. pneumoniae* strains were grown until early exponential phase ($OD_{600nm} \sim 0.3$). Aliquots of 1 ml were first washed three times with fresh C + Y media at 37ºC and then the remaining pellet resuspended in 50 µl of fresh C+Y media. About 2-3 µl of the 50µl cell suspension were loaded on to a Pre-C + 1% agarose microscope slide [39]. Images were obtained using a Zeiss Axio Observer microscope equipped with a Photometrics CoolSNAP HQ2 camera (Roper Scientific) with the appropriate filters. Exposure times were: Phase contrast 100 msec, YFP 5000 msec and CFP 5000 msec. After acquisition, these images were analyzed using Metamorph (Meta Imaging series 7.5) and Image J softwares [44]. Quellung reactions were performed using 1 ml aliquots of liquid cultures (OD600 ~ 0.5). Cells were washed 3 times with fresh C + Y medium at 37°C and resuspended in a final volume of 50 ml of C + Y medium. A volume of 2 µl of this suspension was mixed with 2 µl of CPS14 pneumococcal antisera (SSI Diagnostica) and the resulting reaction was observed under the microscope.

Dot-Blot

Cell samples were prepared by harvesting cells at early exponential growth-phase (OD600 ~ 0.3), and then resuspended in water. After adjusting the samples to the same cell density, cells were lysed with deoxycholate and boiled for 3 minutes before use. Samples were loaded into Hybond PVDF (Amersham) membranes, pre-equilibrated in PBS and placed on top of PBS-soaked Hybond Blotting Paper. The membranes were allowed to air-dry for 30 min and then blocked during 1 hour in Blocking Buffer (5% non-fat dried milk in PBS). Membranes were washed in PBS-T (PBS + 0.05% Tween 20) and incubated overnight at 4°C with primary antibody Anti-CPS14 purified as previously described [45] diluted 1/1000 in PBS-T. After washing with PBS-T, membranes were incubated during 1 hour at room temperature with the secondary antibody Anti-Rabbit IgG peroxidase linked diluted 1/100000 in PBS-T. Membranes were again washed with PBS-T and detected using the ECL Plus Western Blotting Detection Reagents (Amersham).

Results and Discussion

Screening methods to identify small inhibitory/interacting (SI) peptides capable of preventing the interaction between Wzd and Wze capsular proteins

In *Streptococcus pneumoniae*, *wzd* and *wze* genes encode for the Wzd and Wze capsular proteins, which are required to ensure that synthesis of new capsule takes place at the division septum, where the new cell wall is being assembled [16]. For this to happen, Wzd and Wze have to interact with each other and form a protein complex, Wzd/Wze, which can find the division septa [38]. At this particular sub-cellular region, the Wzd/Wze complex may activate and regulate the synthesis of the capsular polysaccharide to ensure full encapsulation of bacteria. When a *cps* null mutant is transformed with a plasmid expressing Wzd and Wze, the protein complex Wzd/Wze still localizes at the septum, meaning that its septal localization occurs without requiring the presence of additional proteins encoded in the capsule operon [38].

In addition, it has been shown that in *wzd* and *wze* null mutant strains the capsule is still produced, linked to the cell surface in regions of mature cell wall, but it is not present at the division septum of pneumococcal bacteria [38]. Prompted by these results we hypothesized that if interaction between Wzd and Wze is prevented or disrupted, the complex Wzd/Wze will not be recruited to the division site. In this scenario, CPS would not be synthesized at the division septum of encapsulated pnemococcus, which could result in bacteria not fully encapsulated and more prone to be recognized by the host immune system. Studies in mice infected with *S. pneumoniae wzd* or *wze* null mutants have shown that expression of Wzd and Wze is essential for bacteria to cause bacteremia after intranasal challenge [5,33]. Therefore, finding a screening method that could easily identify peptides capable of preventing the interaction between Wzd and Wze could represent an important milestone in the discovery of strategies aimed to resolve pneumococcal infections. *Streptococcus pneumoniae* is found associated with the production of more than 90 different capsular polysaccharides, which are synthesized by different proteins. However, the synthesis of CPS in all serotypes, except serotype 3 and 37, seems to be dependent on Wzd and Wze proteins making these proteins attractive targets for the development of therapies that impair *S. pneumoniae* virulence.

The observations described above led to the main question of this thesis: Is it possible to prevent, in encapsulated bacteria, the interaction between Wzd and Wze capsular proteins? If this was possible, small inhibitory (SI) peptides, capable of preventing this interaction, should reduce virulence of encapsulated *S. pneumoniae* bacteria. The

presence of these SI peptides in the cytoplasm of encapsulated pneumococcal bacteria (expressed from a plasmid or imported from the surrounding growth medium) could prevent the formation of the complex Wzd/Wze in three different ways (Figure 2):

A- SI peptides could sequester Wze, spreading it throughout the bacterial cytoplasm, or alter its structure so that it would stop interacting with Wzd located at the surrounding membrane;

B- SI peptides could block the accessibility of the C-terminus end of Wzd, which is assumed to be required for its interaction with Wze, therefore preventing Wzd from interacting with Wze;

C- SI peptides could interact with the extracellular part of the Wzd protein, if an SI peptide is secreted or added to the surrounding medium, and change the structure or accessibility of Wzd C-terminus end in such a way that would prevent it from interacting with Wze.

In order to take this endeavor we designed two different methods that could screen SI peptides capable of interfering with the interaction of Wzd and Wze proteins (Figure 2). The first tested method was based on the *Escherichia coli* bacterial two-hybrid (BTH) system that allow screening for successful interactions through the colony colour while the other was based in co-localization of fluorescent derivatives of interacting proteins when expressed in *S. pneumoniae*.

Screening method based on a bacterial two-hybrid assay in Escherichia coli

We first designed a screening method based on the bacterial two-hybrid (BTH) system developed by Dr. D. Ladant (Institut Pasteur) [46]. This system is based on the cAMP signaling cascade of *E. coli*, which was constructed to identify protein-protein interactions. The catalytic domain of adenylate cyclase from the Gram-negative bacteria *Bordetella pertussis* consists of two complementary fragments, T18 and T25, which are not active when physically separated. However, when these two fragments are fused to interacting proteins, T18 and T25 are brought close to each other and their interaction activates the synthesis of cAMP. As cAMP is a pleiotropic regulator of gene transcription in *E. coli*, including the gene of the *lac* operon that encodes for β-galactosidase, bacteria expressing interacting proteins can be screened as their colonies become blue in plates containing X-Gal.
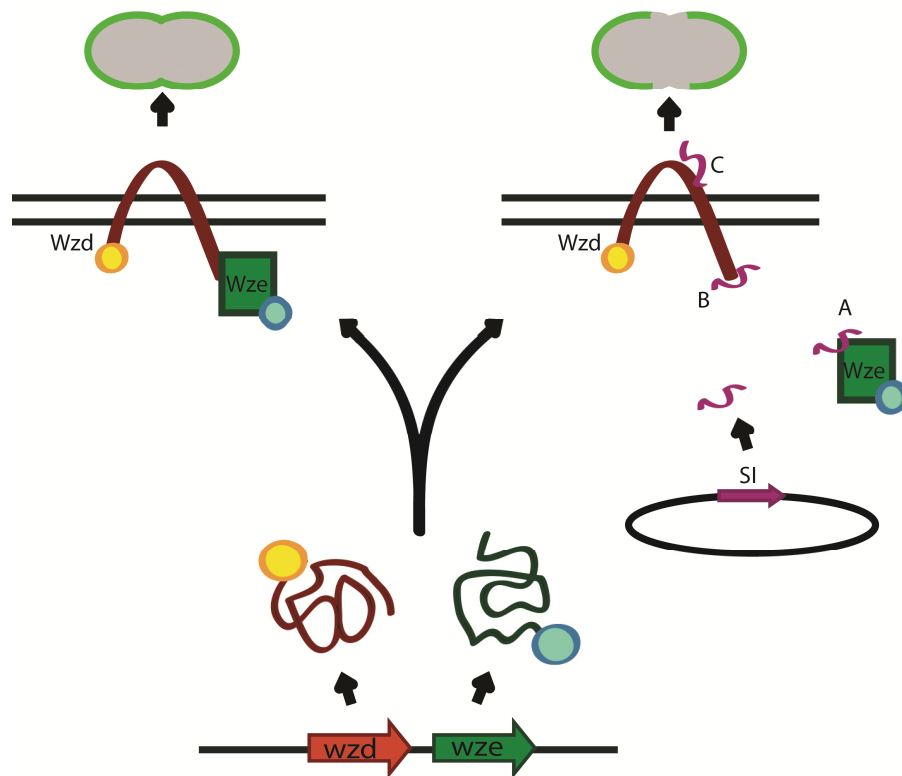
Figure 2: Schematic view of a screening method to find small inhibitory (SI). Both proteins Wzd and Wze are linked to two different tags (yellow and blue circles) that, when close to each other, allow the observation that Wzd and Wze are interacting (ex. production of a specific compound or co-localization of fluorescent signals). When membrane tagged-Wzd and cytoplasmic tagged-Wze are expressed in bacteria, their interaction brings both tags close to each other which allow the visualization of a specific signal. However, when an SI peptide is present, expressed from a plasmid or added to the growth medium, the interaction of the proteins Wzd and Wze is prevented and no signal is observed. The disruption of the complex Wzd/Wze should result in the lack of capsule at the bacteria division septum and in the reduction of virulence in encapsulated bacteria. This disruption can be caused by SI peptides if these molecules interact with Wze (A), with the C-terminus end of Wzd (B) or change the structure of Wzd by interacting with the external part of this protein (C).

In order to validate this approach to screen SI peptides, we asked whether untagged Wze could prevent interaction of T25-Wzd with T18-Wze when expressed in the same *E. coli* strain. T25-Wzd and T18-Wze proteins were encoded in a single low-copy number plasmid (derivative of the plasmid pSU40, which carries a p15A origin of replication), while the Wze was expressed in a second, high copy number, plasmid (pUC19 derivative, which carries a Col E1 origin of replication). If Wze bound to T25-Wzd, T18-Wze could not interact with T25-Wzd and colonies should be white since they are unable to produce β-galactosidase. On the other hand, if such disruption was not successful, bacteria would be able, through the interacting partners T25-Wzd and T18-Wze, to activate the production of cAMP and

consequently the production of β-galactosidase, which would originate the formation of blue colonies when plated in appropriate media.

In Figure 3, plasmids used in the BTH assay are represented schematically and the results obtained when bacteria were transformed with these plasmids are shown. We did not observe any result that indicated the ability of Wze to prevent the interaction of T25-Wzd with T18-Wze (Figure 3, combination 7). The colonies of *E. coli* transformed with the plasmid that allowed the expression of Wze were blue and not white as it was expected if the T25-Wzd/T18-Wze complex was disrupted. We then hypothesized that Wze was unable to compete with T18-Wze for the interaction with T25-Wzd. The inability of Wze to prevent the formation of the T25-Wzd/T18-Wze protein complex could be due to different hypotheses:

i) As T18-Wze is being translated from the same mRNA molecule that also encodes T25-Wzd, the two proteins may form a protein complex, immediately after their translation, too stable to be disrupted by the presence of an untagged Wze;

ii) Wze alone is competing and interacting with the T25-Wzd, but the transient formation of the complex T25-Wzd/T18-Wze results in sufficient cAMP molecules capable of activating the expression of β-galactosidase and turn *E. coli* colonies blue. Therefore the colonies would remain blue even if this interaction was later disrupted.

In order to elucidate which of the two previous hypotheses was most likely to be true, we expressed T25-Wzd and the competing Wze from the same low-copy number plasmid and T18-Wze from the high-copy number plasmid (Figure 3, combinations 9 and 10). This would allow us to address if genetic proximity of the encoding sequences for the proteins tested was important. We observed that when Wze was encoded in the proximity of the encoding sequence of T25-Wzd, it prevented the later from interacting with T18-Wze, as colonies transformed with these plasmids were white (Figure 3, combinations 9 and 10). These results showed that the untagged Wze was being expressed in *E. coli* and that the genetic proximity of the encoding sequences for the proteins involved in the T25-Wzd/T18-Wze complex and for the competing protein, in this case Wze, is an important factor to consider in our screening method. In other words, we showed that if the sequence encoding for the competitor peptide is cloned in a plasmid different from that encoding the T25-Wzd and T18-Wze interacting partners, it cannot disrupt the interaction of these two proteins which probably occurs immediately upon synthesis. It should be emphasized that in *S. pneumoniae* it is not likely that the interaction between Wzd and Wze is irreversible and probably other pneumococcal proteins are necessary for a dynamic association and dissociation of Wzd and Wze.

We also asked if we could delay the timing of the interaction of T25-Wzd with T18-Wze, and consequently the expression of β-galactosidase and the appearance of blue colonies, by expressing Wze and T18-Wze from the same high-copy number plasmid. This plasmid was transformed together with a low-copy number plasmid, which allowed the expression of T25-Wzd (Figure 3, combination 11). In this case, the proteins Wze and T18-Wze were not translated from the same mRNA molecule encoding the T25-Wzd, and they should have similar probabilities of finding interacting partners. We were expecting to see variability in the intensity or the pattern of the color of the colonies (a less intense blue color or the appearance of a mixture of white and blue colonies) because both Wze and T18-Wze had the same probability to interact with T25-Wzd. However, this was not the case as all the colonies were blue and no differences were observed between them (Figure 3, combination 11), suggesting that this method is not adequate to screen for SI peptides capable of preventing the interaction of T25-Wzd and T18-Wze. This is in accordance with the expectation that a bacterial two-hybrid assay should only be used to enquire if two proteins may interact and cannot be used to quantify protein interactions.

Finally, we enquired whether we could reduce the stability of the interaction between T25-Wzd and Wze, to facilitate the interaction with SI peptides and the consequent disruption of the T25-Wzd/Wze protein complex, by co-expressing them with the Wzh phosphatase (Figure 3, combination 12). In encapsulated *S. pneumoniae* bacteria, this phosphatase has been proposed to dephosphorylate the protein Wze, which is interacting with Wzd, to allow dynamic association/dissociation reactions between Wzd and Wze [47]. If that was the case, expression of Wzh from the same plasmid encoding T25-Wzd and Wze, would decrease the stability of the T25-Wzd/Wze complex allowing T25-Wzd to interact with T18-Wze, encoded in a second plasmid. In this case, *E. coli* colonies should become blue. We observed that transformation of *E.coli* with these plasmids also resulted in white colonies, indicating that the presence of the protein Wzh does not influence the stability of the T25-Wzd/Wze protein complex (Figure 3, combination 12).
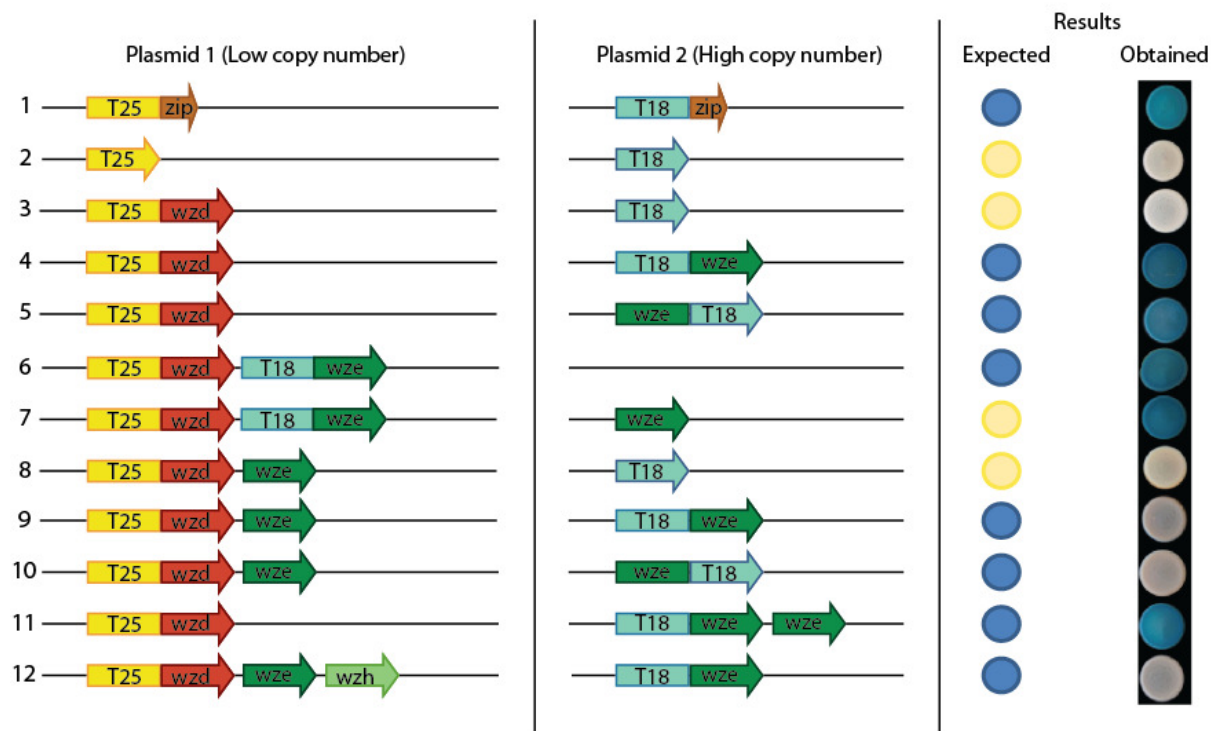
Figure 3: A screening method based in an *E. coli* bacterial two-hybrid assay is not adequate to find candidates for SI peptides. Schematic representation of the plasmids used to transform the strain BTH101 in the bacterial two-hybrid assay is shown in the left and central panels. The expected results, in case of a successful screening method, and those observed are shown on the right panel. The controls used for this experiment were the simultaneous expression of T25-zip and T18-zip (positive control, combination 1), which originates blue colonies as the zip tag can form dimers and bring together the T25 and T18 CyaA fragments; the expression of the CyaA fragments T25 and T18, or the simultaneous expression of T25-Wze and T18 (negative controls, combination 2 and 3, respectively), which originate white colonies as they cannot interact; the simultaneous expression of T25-Wzd with T18-Wze or Wze-T18 (combination 4 and 5), which originates blues colonies as it has been shown that Wzd and Wze can interact and bring close together the CyaA T25 and T18 fragments [38]. Expression of T25-Wzd and T18-Wze from the same plasmid showed that these proteins could still interact, as they originated blue colonies (combination 6), but the expression of untagged Wze could not prevent this interaction (combination 7). Expression of Wze from the same plasmid that encoded T25-Wzd did not originate blue colonies (combination 8), as they cannot interact. However, this prevented T25-Wzd from interacting with T18-Wze, or Wze-T18, that are encoded in the high-copy number plasmid (combination 9 and 10, respectively). These results indicate that the genetic proximity of the encoding sequences for T25-Wzd, T18-Wze and Wze is relevant for a successful screening of SI peptides. Expression of Wze from the same plasmid encoding T18-Wze did not impair its interaction with T25-Wzd (combination 11), as colonies transformed with the encoding plasmids were still blue. Expression of Wzh from the same plasmid encoding T25-Wzd and Wze did not interfere with their ability to interact (combination 12), as colonies transformed with the encoding plasmids were still white.

In summary, we concluded that we should not use a method based in the *E. coli* bacterial two-hybrid assay to screen for SI peptides that were capable of preventing the interaction of Wzd and Wze. Furthermore, we learned that the genetic proximity of the sequences encoding for protein interacting partners seem to be important in this *E. coli* system. One explanation for the results obtained could be that we were expressing these proteins in *E. coli*, which does not have the capsule operon encoding other pneumococcal proteins which could be necessary for the dynamic interaction between Wzd and Wze, allowing other peptides to interfere and compete. Hence, by using a simple method in *E. coli* some crucial protein interactions might not occur. Therefore, we set to design a new screening method in encapsulated *S. pneumoniae* bacteria, where all the complex reactions and interactions between proteins may take place.

Screening method based on the expression and localization of fluorescent proteins in Streptococcus pneumoniae

As the initial *E. coli* based method could not be used to screen for SI peptides, we then designed an alternative method based in the expression and localization of fluorescent proteins in *S. pneumoniae* (Figure 2).

In this new method functional fluorescent derivatives of Wzd and Wze were expressed from their respective genes inserted into the chromosome, at the *cps* operon, of encapsulated *S. pneumoniae* bacteria. It has been observed that Wzd is attached to the membrane while Wze is cytoplasmic, and only when both interact they are able to localize at the division septum [38]. We hypothesized that by attaching a different fluorescent protein to each Wzd and Wze interacting partner, it would be possible to determine when they were interacting with each other (as both should co-localize at the division septum) or when this interaction was prevented (and therefore Wze should diffuse into the cytoplasm while Wzd should spread through the entire membrane). In this way, we would be able to find SI peptides through the transformation of plasmids expressing different peptides capable of interfering with the co-localization of Wzd and Wze, assessed by the visualization of both fluorescent signals.

As the attachment of a fluorescent domain to the N- or C-terminus of Wzd and Wze proteins may result in a non-functional protein or prevent its ability to interact with other proteins involved in the synthesis of CPS, we decided to construct pneumococcal plasmids that allow the expression of different fluorescent proteins in various strains to choose the best combination, where both proteins are functional (Figure 4). The two fluorescent proteins chosen were the Citrine fluorescent protein (excitation and emission peaks of 516 and 529 nm, respectively) and the Cyan fluorescent protein (CFP, excitation and emission peaks of

458 and 480 nm, respectively). The plasmid encoding a functional Wze-Citrine had been already reported [38], therefore we had only to construct plasmids for the expression of Citrine-Wze, Wzd-CFP and CFP-Wzd.

Surprisingly, transformation of the unencapsulated strain R36A with the plasmids encoding the proteins CFP-Wzd and Citrine-Wze did not result in fluorescent bacteria. This may be due to the proposed role of the folding of the 5' coding region of mRNA, immediately after the start codon, which may shape expression levels. It seems that tightly folded messages, for example with long hairpin loops, obstruct translation initiation and thereby reduce protein synthesis [47].

Meanwhile, we have observed that the presence of the first ten aminoacids of the protein Wze allows the expression of fluorescence, when fused to the N-terminal of the protein in study in *S. pneumoniae* [42]. The presence of the nucleotide sequence, encoding the first ten aminoacids of the protein Wze and named "i-tag", is essential for the fluorescence probably by destabilizing the mRNA structure of this region, which facilitates the ribosome binding to the mRNA molecule. This may increase the rate of initiation of the translation of the protein. Therefore, we constructed new plasmids that permitted the expression of iCitrine-Wze and iCFP-Wzd (plasmids pBCSJF001 and pBCSJF003, respectively).

To determine whether the different fluorescent proteins were functional, R36A, a laboratory unencapsulated strain, and ATCC6314, a serotype 14 encapsulated strain, were transformed with the plasmids expressing the different Wzd and Wze fluorescent derivatives (Figure 4 – i, ii, vii and viii).

In the unencapsulated R36A strain, iCFP-Wzd and Wzd-CFP proteins were distributed all over the membrane, while iCitrine-Wze and Wze-Citrine proteins were spread throughout the entire cytoplasm, as expected.

In the encapsulated ATCC6314 strain, both iCFP-Wzd and Wzd-CFP proteins were able to localize at the bacterial division septum. The same result was observed when the strain ATCC6314 was transformed with plasmids encoding for iCitrine-Wze and Wze-Citrine (Figure 4 – iii, iv, ix and x).

We further enquired the functionality of the fluorescent derivatives of Wzd and Wze by expressing them in ATCC6314Δ*wzd* (null mutant for the *wzd* gene) and ATCC6314Δ*wze* (null mutant for the *wze* gene), respectively (Figure 4 – v, vi, xi and xii). As these strains do not produce the native Wzd and Wze proteins, which cannot interfere or compete with the fluorescent Wzd and Wze derivatives, we expected a better septal localization of the fluorescent signals if the proteins were functional. The results obtained for the expression of Wzd fluorescent derivatives were identical to the ones observed with the strain ATCC6314. No significant difference in the fluorescence between the two transformed strains was

observed (Figure 4). However, a different result was observed for the expression of Wze fluorescent derivatives. While Wze-Citrine fluorescent signal was correctly located at the bacterial division septum, the protein iCitrine-Wze had an unexpected localization since it was distributed all over the membrane. As Wze is a cytoplasmic protein, this suggests that when a fluorescent protein is fused at the N-terminus of the protein Wze, the resulting iCitrine-Wze may be able to interact with Wzd, which can recruit it to the membrane. However, iCitrine-Wze protein is not fully functional, as the protein complex made of Wzd/iCitrine-Wze cannot migrate to the division septum of dividing bacteria. The attachment of the fluorescent protein at this end may change the protein's conformation and conceal the surface required for interactions with other proteins. We concluded that iCitrine-Wze cannot be used to construct the *S. pneumoniae* mutant required for this new screening strategy.

In order to enquiry about the ability of the fluorescent derivatives of Wzd and Wze to regulate the synthesis of the capsular polysaccharide, we determined whether ATCC6314 *wzd* and *wze* null mutant strains expressing these proteins could agglutinate when in the presence of antibodies capable of binding the bacterial capsule (Quellung reaction). In this assay cells were incubated with rabbit-serum raised against pneumococcal type 14 capsular polysaccharide and then observed by microscopy (Figure 5).

Expression of both Wzd-CFP and iCFP-Wzd in ATCC6314Δ*wzd* seemed to complement the production of CPS, as these strains showed a similar ability to agglutinate as the encapsulated ATCC6314 strain, the positive control. As expected, the strain ATCC6314Δ*wzd* presented a decreased number of agglutinated cells and no agglutination was observed with the unencapsulated strain R36A, which does not produce any capsule (Figure 5). In addition, expression of Wze-Citrine in the ATCC6314Δ*wze* strain resulted in an improved ability to agglutinate, in the presence of the serum. The agglutination observed was very similar to the wild-type, again indicating that the capsule is being synthesized. On the other hand, the expression of iCitrine-Wze resulted in an apparent reduction of agglutination, as expected for a protein not fully functional. In order to confirm these results we decided to perform a Dot-blot assay (Figure 5) to measure the amount of capsule produced when *wzd* and *wze* null mutants were transformed with plasmids that allowed the expression of the different fluorescent derivatives of Wzd and Wze. We observed a clear recovery of the ability to produce capsule when the *wze* null mutant was transformed with the plasmid that allowed the expression of Wze-Citrine (Figure 5, vii).
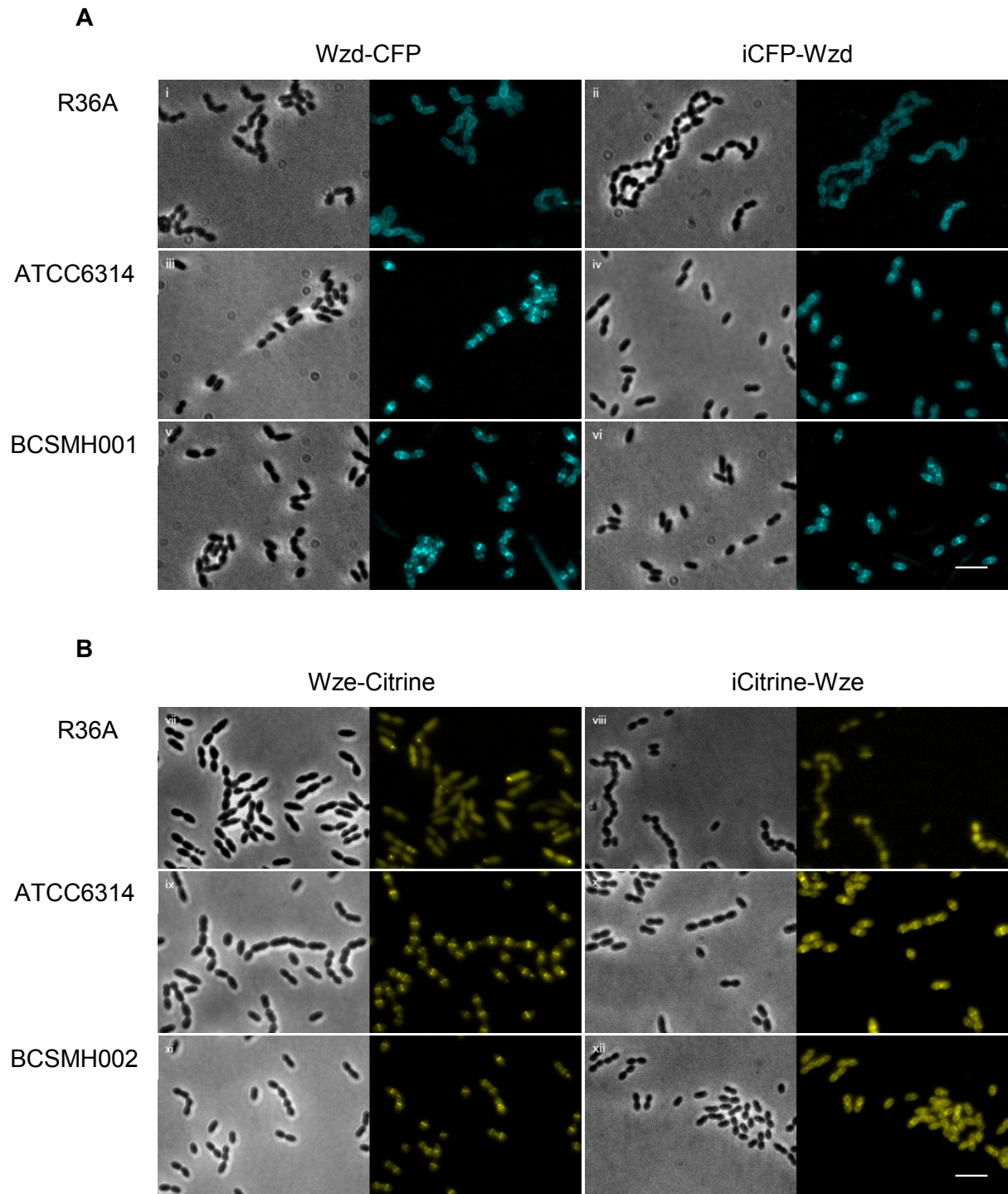
Figure 4: Localization of Wzd and Wze fluorescent derivatives in encapsulated and unencapsulated pneumococcal strains. A. Microscopy images of *S. pneumoniae* strains expressing constitutively Wzd-CFP and iCFP-Wzd in the unencapsulated R36A strain (i. strain BCSJF006 and ii. strain BCSJF003, respectively), in the encapsulated ATCC6314 strain (iii. strain BCSJF008 and iv. strain BCSJF007, respectively) and in the BCSMH001 (ATCC6314 *wzd* null mutant) (v. strain BCSJF011 and vi. strain BCSJF010, respectively). B. Microscopy images of *S. pneumoniae* strains expressing constitutively Wze-Citrine and iCitrine-Wze in the unencapsulated R36A strain (i, strain BCSMH007 and ii. strain BSCJF005, respectively), in the encapsulated ATCC6314 strain (iii. strain

BCSMH016 and iv. strain BCSJF001, respectively) and in the BCSMH002 (ATCC6314 *wze* null mutant) (v. strain BCSMH023 and vi. strain BCSJF009, respectively). Scale bar, 2μm.

When *wzd* null mutant was transformed with plasmids that permitted the expression of iCFP-Wzd and Wzd-CFP (Figure 5, iv and v) we observed a small increase in the amount of produced capsule, which was not as high as in the parental encapsulated ATCC6314 strain. These results show that the proteins Wzd-CFP, iCFP-Wzd and Wze-Citrine are, at least, partially functional and therefore their genes are suitable for the genomic substitution of the *wzd* and *wze* genes in the *cps* operon.
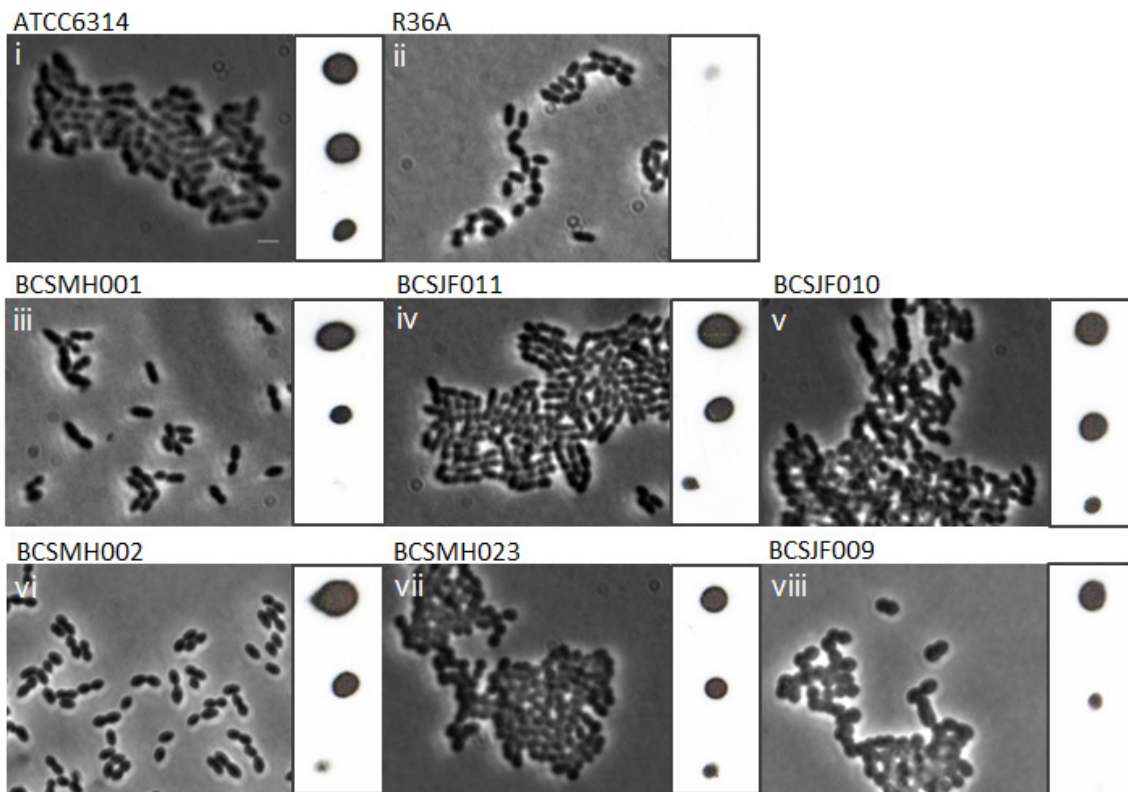
Figure 5: Confirmation of production of capsular polysaccharide due to expression of fluorescent derivatives of Wzd and Wze by agglutination and Dot-blot. Brightfield images of the agglutination resulting from Quellung reaction. i) ATCC6314, encapsulated parental strain, ii) R36A, unencapsulated strain, iii) BCSMH001 strain, ATCC6314Δ*wzd*, iv) BCSJF011 strain, ATCC6314Δ*wzd* expressing Wzd-CFP, v) BCSJF010 strain, ATCC6314Δ*wzd* expressing iCFP-Wzd, vi) BCSMH002 strain, ATCC6314Δ*wze*, vii) BCSMH023 strain, ATCC6314Δ*wze* expressing Wze-Citrine and viii) BCSJF09 strain, ATCC6314Δ*wze* expressing iCitrine-Wze. Scale bar, 2μm. Next to each panel are shown the results obtained by Dot-Blot analysis using undiluted (top), 10 fold diluted (middle) and 100 fold diluted culture (bottom).

The strain ATCC6314*wze::wze_citrine* from Henriques *et al.* [39] has the chromosomal gene *wze* mutated so that Wze has the fluorescent protein Citrine fused to it. With the intention of testing both fluorescences, this strain was transformed with the plasmids pBCSJF003 and pBCSJF005 separately, and both proteins Wzd and Wze could be visualized and its co-localization analyzed. Figure 6 shows the fluorescence from the Wze-Citrine encoded in the chromosome and Wzd-CFP and iCFP-Wzd encoded in the plasmids transformed. The fluorescence from CFP was not as strong as that observed with expression of Citrine. Moreover, the percentage of cells that had the fluorescence signal localized at the septa was inferior to that observed in the parental strain. This could be explained by the presence of the wild-type protein Wzd, encoded in the chromosome, that can interfere with the interaction of Wzd-CFP with Wze-Citrine. The results demonstrate that Wzd-CFP and Wze-Citrine proteins co-localize. Similar results were obtained with transformation of plasmids pBCSJF003 (iCFP-Wzd) and pBCSJF005 (Wzd-CFP).
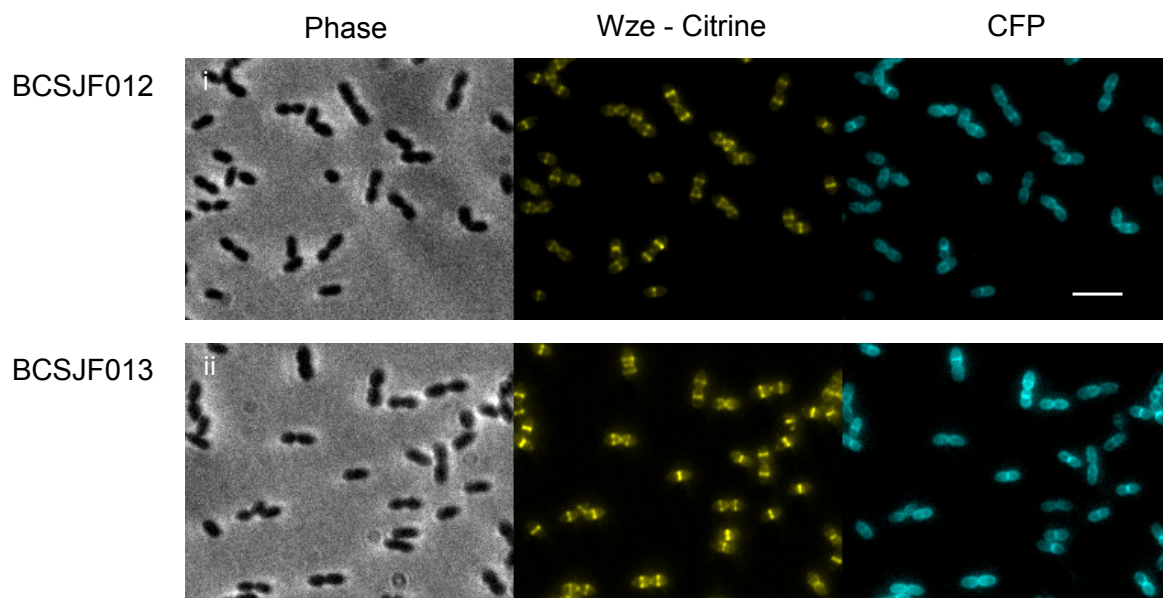
Figure 6: Fluorescence microscopy images of the encapsulated strain ATCC6314*wze*::*wze_citrine* expressing the plasmids i) iCFP-Wzd (pBCSJF003) and ii) Wzd-CFP (pBCSJF005). Scale bar, 2 µm.

All these results are evidence that the Wzd-CFP or iCFP-Wzd should be functional when expressed in the strain BCSMH004 (ATCC6314*wze*::*wze_citrine*). Cloning both genes should result in a suitable strain for the screening method proposed (both proteins Wzd and Wze are fluorescent and are able to localize at the septum of the cells).

In order to construct a strain expressing both fluorescent derivatives of Wzd and Wze from their native locus we exchanged the *wzd* gene for the sequence encoding Wzd-CFP or iCFP-Wzd in BCSMH004 strain [38]. Figure 7 shows a schematic representation of the integration of a plasmid encoding a truncated Wzd-CFP in the bacterial chromosome.

Excision of this plasmid allows the construction of a strain expressing Wzd-CFP at the *cps* locus.

As shown previously, both Wzd-CFP and iCFP-Wzd proteins seemed functional and therefore we expected that the resulting strains ATCC6314 *wzd*::*iCFP_wzd wze*::*wze_citrine* (strain BCSJF014) and ATCC6314 *wzd*::*wzd_*CFP *wze*::*wze_citrine* (strain BCSJF015) would be suitable strains for our screening method (they expressed functional and fluorescent Wzd and Wze derivatives that are able to localize at the division septum of bacteria and that could delocalize in the presence of an SI peptide).
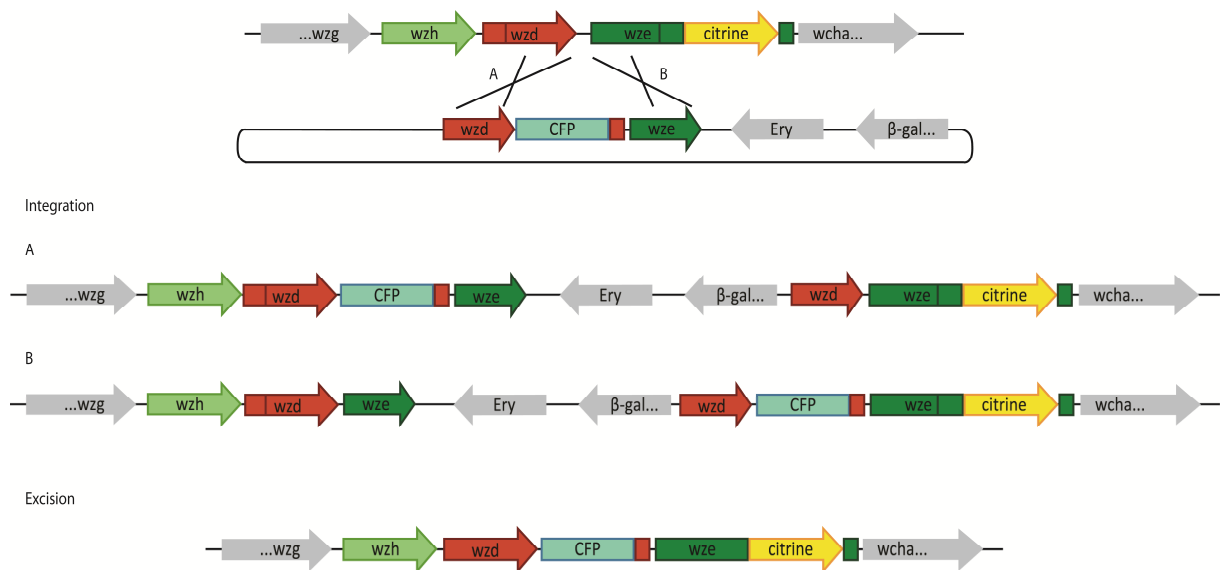
Figure 7: Schematic representation of the construction of the mutant showing the two forms of integration of the plasmid into the chromosome (in this case the *wzd* gene is being replaced in the chromosome by *wzd-CFP*).

In spite of having both Wzd and Wze proteins fused to fluorescent protein tags, the strain BCSJF014, constructed in different attempts, was non-fluorescent. This was a surprising result that is still not well understood. However, strain BCSJF015 was fluorescent and the fluorescence signal from both Wzd and Wze proteins co-localized at the septum of dividing bacteria (Figure 8, i). This allowed us to conclude that the BCSJF015 mutant strain may be used as a tool to determine which SI peptides can prevent the Wzd and Wze interaction.

Validation of the use of BCSJF015 mutant strain, which produces Wzd-CFP and Wze-Citrine localized at the division septum, to screen SI peptides

We confirmed that BCSJF015 (ATCC6314*wzd*::*wzd_CFPwze*::*wze_citrine*) grows normally. Growth curves were made and compared with the parental encapsulated ATCC6314 and the strain ATCC6314*wze::wze_citrine* (strain BCSMH004). No significative difference in the growth rates was observed (data not shown). We further enquired whether this mutant strain was capable of producing capsule. This strain was incubated with rabbit-serum raised against pneumococcal capsular polysaccharide serotype 14, which resulted in a positive reaction of agglutination (quellung reaction). Strain BCSJF015 seemed to produce more capsule than the null mutant ATCC6314Δ*wzd*, but not as much as the encapsulated wild type strain ATCC6314 (data not shown). This suggests that although both Wzd-CFP and Wze-Citrine can co-localize at the division septum of BCSJF015 strain. However, the production of the capsule is not taking place at parental levels. This may be due to the possibility that the Wzd-CFP is not fully functional or that expression of Wzd-CFP is not taking place at the correct levels.
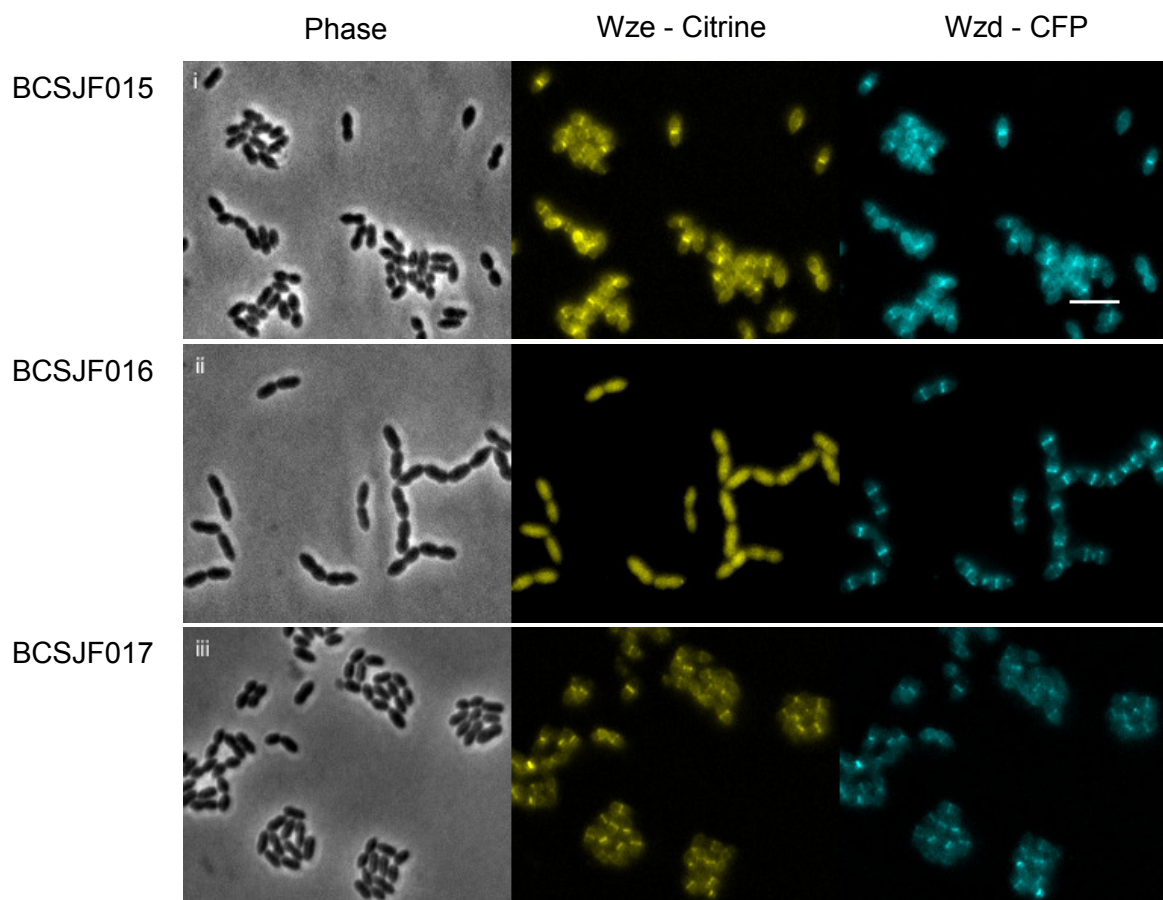
Figure 8: Constitutive expression of untagged Wze can delocalize the fluorescent Wze-Citrine produced from it native locus, the *cps* operon. Brightfield and fluorescence microscopy images of BCSJF015 (i, a derivative of the encapsulated ATCC6314 modified to express *wzd-CFP* and *wze-Citrine* from the *cps operon*). Taking into account that the CFP protein requires excitation of a higher energy to fluoresce, and to avoid the photobleaching of the Citrine, all the photographs were taken in the specific order Citrine → CFP. Also shown microscopy images of BCSJF016 (ii, BCSF015 transformed with plasmid pBCSMH017 that permit the constitutive expression of untagged Wze) and BCSJF017 (iii, BCSF015 transformed with plasmid pBCSLF001, empty vector used to express proteins constitutively). Scale bar, 2 μm.

Considering whether we could use the BCSJ015 strain to screen SI peptides that can prevent the interaction between the proteins Wzd and Wze, we decided to transform the mutant strain BCSJF015 with a plasmid encoding for the untagged Wze protein. Figure 8 shows that as expected, Wzd-CFP kept its septal localization since it is interacting with an untagged Wze. In addition, the untagged Wze protein when expressed from the plasmid could compete with the chromosomal protein Wze-Citrine. Accordingly, the fluorescence signal from the Wze-Citrine appeared all over the cytoplasm (strain BCSJF016). This delocalization was not observed when BCSJF015 was transformed with the plasmid that did not have encoded the protein Wze (strain BCSJF017).

Conclusion

Pneumococcal bacteria are surrounded by a capsular polysaccharide, which forms a shield that prevents recognition of bacterial cell structures, such as peptidoglycan, teichoic acids and cell surface proteins, by the host. As unencapsulated pneumococci are non-virulent strains, this polysaccharide structure is considered to be a major virulence factor that allows the bacteria to evade the host immune system.

Recently, it was shown that Wzd and Wze proteins are essential for the synthesis of CPS at the division site of bacteria [38]. When one of these proteins is not present, the capsule is not synthesized at the division septum but it is still present in regions of mature cell wall. Hence, it has been proposed that the synthesis of the CPS occurs in two modes, septal and non-septal [38], in a similar way of what had been observed for peptidoglycan synthesis [49].

Our first attempt to design a screening method to identify peptides capable of disrupting the interaction of Wzd with Wze was carried out using the Gram-negative *Escherichia coli* model. Namely, we used a derivative of a bacterial two-hybrid assay and tried to validate it by testing whether untagged Wze protein would be identified as a competitor peptide capable of disrupting the interaction between T25-Wzd and T18-Wze. The method involved the transformation of two plasmids into *E. coli*: the first expressed the two interacting proteins T25-Wzd and T18-Wze, encoded by adjacent sequences, while the second plasmid expressed the candidate peptide capable of preventing or disrupting their interaction.

The results showed that the strategy using a screening method based on the bacterial two-hybrid assay in *E.coli* is not appropriate. The interaction between T25-Wzd and T18-Wze could only be prevented when the expression of a competitor protein, in this case the untagged Wze protein, took place from the same plasmid that expressed T25-Wzd. This means that in this case the competitor Wze can interact with T25-Wzd but it may be unable to disrupt a previously assembled T25-Wzd/T18-Wze protein complex.

The second screening method was based in the expression of fluorescent derivatives of Wzd and Wze in encapsulated *S. pneumoniae*, allowing us to determine their temporal and spatial localization in live cells. We decided to construct a *S. pneumoniae* mutant strain, which produces the fluorescent derivatives of Wzd and Wze, from their native genetic locus in the *cps* operon. Then protein should be capable of interacting with each other, and probably also with other pneumococcal proteins, so that they had a septal localization.

We used the strain BCSMH004, an encapsulated serotype 14 strain that express Wze-Citrine from its native locus at the *cps* operon, to construct the mutants BCSJF014 and BCSJF015, which express iCFP-Wzd and Wzd-CFP, respectively, together with Wze-Citrine

from the same locus. Only BCSJF015, the strain that express Wzd-CFP and Wze-Citrine, was fluorescent and therefore was used in the development of the screening method.

Finally, we confirmed that the mutant BCSJF015 strain could be used as a tool to screen inhibitory peptides that would prevent the assembly of the Wzd/Wze protein complex. This was tested by transforming the mutant strain BCSJF015 with a plasmid encoding the protein Wze. We observed that the fluorescence signal of Wze-Citrine was found all over the cytoplasm and not at the division septum, which indicates that expression of untagged Wze can compete with Wze-Citrine.

This screening method can be performed using small peptides that can be integrated into the bacteria or using chemical compounds that can be added to the surrounding medium. Morona *et al.* showed a screening method to identify inhibitors of the protein Wzh [51]. This class of anti-virulence therapeutics should not inhibit the growth of bacteria but will suppress virulence, turning them susceptible to the immune system of the host.

References

1. Kadioglu, A., Weiser, J. N., Paton, J. C. & Andrew, P. W. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. Nature reviews Microbiology 6, 288–301 (2008).

2. Sorensen, U. B. S., Blom, J., Birch-andersen, A. & Henrichsenl, J. Ultrastructural localization of capsules, cell wall polysaccharide, cell wall proteins, and F antigen in Pneumococci. Infection and immunity 56, 1890–1896 (1988).

3. Abeyta, M., Hardy, G. G. & Yother, J. Genetic alteration of capsule type but not pspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. Infection and immunity 71(1):218 (2003).

4. Wartha, F. et al. Capsule and D-alanylated lipoteichoic acids protect *Streptococcus pneumoniae* against neutrophil extracellular traps. Cellular microbiology 9, 1162–71 (2007).

5. Morona, J. K., Miller, D. C., Morona, R. & Paton, J. C. The effect that mutations in the conserved capsular polysaccharide biosynthesis genes *cpsA, cpsB,* and *cpsD* have on virulence of *Streptococcus pneumoniae*. The Journal of infectious diseases 189, 1905–13 (2004).

6. Cundell, D. R., Weiser, J. N., Shen, J., Young, A. & Tuomanen, E. I. Relationship between colonial morphology and adherence of Streptococcus pneumoniae. Infection and immunity 63, 757–61 (1995).

7. Talbot, U. M., Paton, a W. & Paton, J. C. Uptake of *Streptococcus pneumoniae* by respiratory epithelial cells. Infection and immunity 64, 3772–7 (1996).

8. Kim, J. O. & Weiser, J. N. Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of *Streptococcus pneumoniae*. The Journal of infectious diseases 177, 368–77 (1998).

9. García, E., Llull, D. & López, R. Functional organization of the gene cluster involved in the synthesis of the pneumococcal capsule. International microbiology : the official journal of the Spanish Society for Microbiology 2, 169–76 (1999).

10. Bentley, S. D. et al. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS genetics 2, e31 (2006).

11. Babl, F. E., Pelton, S. I., Theodore, S. & Klein, J. O. Constancy of distribution of serogroups of invasive pneumococcal isolates among children: experience during 4 decades. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America 32, 1155–61 (2001).

12. Pletz, M. W., Maus, U., Krug, N., Welte, T. & Lode, H. Pneumococcal vaccines: mechanism of action, impact on epidemiology and adaption of the species. International journal of antimicrobial agents 32, 199–206 (2008).

13. Kamerling, J. P. et al. "Pneumococcal polysaccharides: a chemical view." In Streptococcus pneumoniae: Molecular Biology and Mechanisms of Disease. Tomasz, A. (ed.). New York: Larchmont, pp. 81–114. (1999).

14.   Hortal, M. et al. Capsular type distribution and susceptibility to antibiotics of *Streptococcus pneumoniae* clinical strains isolated from Uruguayan children with systemic infections. Pneumococcus Study Group. Microbial drug resistance (Larchmont, N.Y.) 3, 159–63 (1997).

15.   Powell, A., Lindberg, B. and Lonngrer J. Structural studies on the specific type-14 pneumococcal polysaccharide. Carbohydrate research 58, 177-186 (1977).

16.   Sørensen, U. B., Henrichsen, J., Chen, H. C. & Szu, S. C. Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of *Streptococcus pneumoniae* revealed by immunochemical methods. Microbial pathogenesis 8, 325–34 (1990).

17.   Vollmer, W., Blanot, D. & de Pedro, M. A. Peptidoglycan structure and architecture. FEMS microbiology reviews 32, 149–67 (2008).

18.   Deng, L., Kasper, D. L., Krick, T. P. & Wessels, M. R. Characterization of the linkage between the type III capsular polysaccharide and the bacterial cell wall of group B *Streptococcus*. The Journal of biological chemistry 275, 7497–504 (2000).

19.   Mavroidi, A. et al. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. Journal of bacteriology 189, 7841–55 (2007).

20.   Aanensen, D. M., Mavroidi, A., Bentley, S. D., Reeves, P. R. & Spratt, B. G. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. Journal of bacteriology 189, 7856–76 (2007).

21.   Yother, J. Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation. Annual review of microbiology 65, 563–81 (2011).

22.   García, E. and López R., Molecular biology of the capsular genes of *Streptococcus pneumoniae*. FEMS Microbiology letters 149, 1-10 (1997).

23.   Cieslewicz, M. J., Kasper, D. L., Wang, Y. & Wessels, M. R. Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci. The Journal of biological chemistry 276, 139–46 (2001).

24.   Kawai, Y. et al. A widespread family of bacterial cell wall assembly proteins. The EMBO journal 30, 4931–41 (2011).

25.   Eberhardt, A. et al. Attachment of capsular polysaccharide to the cell wall in *Streptococcus pneumoniae*. Microbial drug resistance (Larchmont, N.Y.) 18, 240–55 (2012).

26.   Morona, J. K., Morona, R., Miller, D. C. & Paton, J. C. *Streptococcus pneumoniae* Capsule biosynthesis protein CpsB is a novel manganese-dependent. Journal of bacteriology 184, 577–583 (2002).

27.   Bender, M. H. & Yother, J. CpsB is a modulator of capsule-associated tyrosine kinase activity in *Streptococcus pneumoniae*. The Journal of biological chemistry 276, 47966–74 (2001).

28.   Standish, A. J., Salim, A. a, Zhang, H., Capon, R. J. & Morona, R. Chemical inhibition of bacterial protein tyrosine phosphatase suppresses capsule production. PloS one 7, e36312 (2012).

29.   Grangeasse, C., Cozzone, A. J., Deutscher, J. & Mijakovic, I. Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. Trends in biochemical sciences 32, 86–94 (2007).

30. Grangeasse, C., Terreux, R. & Nessler, S. Bacterial tyrosine-kinases: structure-function analysis and therapeutic potential. Biochimica et biophysica acta 1804, 628–34 (2010).

31. Lee, D. C. & Jia, Z. Emerging structural insights into bacterial tyrosine kinases. Trends in biochemical sciences 34, 351–7 (2009).

32. Morona, R., Van Den Bosch, L. & Daniels, C. Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. Microbiology (Reading, England) 146 (Pt 1, 1–4 (2000).

33. Morona, J. K., Morona, R. & Paton, J. C. Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. Proceedings of the National Academy of Sciences of the United States of America 103, 8505–10 (2006).

34. Cartee, R. T., Forsee, W. T., Bender, M. H., Ambrose, K. D. & Yother, J. CpsE from Type 2 Streptococcus pneumoniae Catalyzes the Reversible Addition of Glucose-1-Phosphate to a Polyprenyl Phosphate Acceptor , Initiating Type 2 Capsule Repeat Unit Formation. Journal of bacteriology 187, 7425–7433 (2005).

35. López, R. & García, E. Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. FEMS microbiology reviews 28, 553–80 (2004).

36. Type, S., Jiang, S., Wang, L., Reeves, P. R. & Wang, L. E. I. Molecular Characterization of *Streptococcus pneumoniae* Type 4, 6B, 8, and 18C capsular polysaccharide gene clusters. Infection and immunity 69, 1244-1255 (2001).

37. Morona, J. K., Morona, R., Miller, D. C. & Paton, J. C. Mutational analysis of the carboxy-terminal $(YGX)_4$ repeat domain of CpsD, an autophosphorylating tyrosine kinase required for capsule biosynthesis in *Streptococcus pneumoniae*. Journal of bacteriology 185, 3009–3019 (2003).

38. Henriques, M. X., Rodrigues, T., Carido, M., Ferreira, L. & Filipe, S. R. Synthesis of capsular polysaccharide at the division septum of Streptococcus pneumoniae is dependent on a bacterial tyrosine kinase. Molecular microbiology 82, 515–34 (2011).

39. Lacks, S. & Hotchkiss, R. D. A study of genetic material determining an enzyme activity in *pneumococcus.* Biochimica et biophysica acta 39, 508–518 (1959).

40. Green M. R. and Sambrook J., Molecular Cloning : A Laboratory Manual (Fourth Edition).

41. Martin, B., García, P., Castanié, M. P. & Claverys, J. P. The recA gene of *Streptococcus pneumoniae* is part of a competence-induced operon and controls lysogenic induction. Molecular microbiology 15, 367–79 (1995).

42. Henriques, M. Catalão, M. J. Figueiredo, J. Gomes, J. P. & Filipe, S. Construction of improved tools for protein localization in *Streptococcus pneumoniae*. Submitted.

43. Kloosterman, T. G., Bijlsma, J. J. E., Kok, J. & Kuipers, O. P. To have neighbour's fare: extending the molecular toolbox for *Streptococcus pneumoniae*. Microbiology (Reading, England) 152, 351–9 (2006).

44. Abràmoff, M. D., Magalhães, P. J. & Abràmoff, Ram, S. J., Image processing with ImageJ. Biophotonics Int 11, 36-42 (2004).

45. Xayarath, B. & Yother, J. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of

suppressor mutations and can affect polymer transfer to the cell wall. *Journal of bacteriology* 189, 3369–81 (2007).

46.   Adant, D. A. L. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc. Natl. Acad. Sci. USA 95, 5752–5756 (1998).

47.   Kudla, G. et al. Coding sequence determinants of gene expression in *Escherichia coli.* Science 324, 255–258 (2009).

48.   Zapun, A., Vernet, T. & Pinho, M. G. The different shapes of cocci. FEMS microbiology reviews 32, 345–60 (2008).

49.   Mccarty, M. et al. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. The journal of experimental medicine 79, 137-158 (1943).

50.   Leenhouts, K. et al. A general system for generating unlabelled gene replacements in bacterial chromosomes. Molecular & general genetics: MGG 253, 217–24 (1996).

51.   Standish, A.J., Salim, A.A., Zhang, H., Capon, R.J., & Morona, R. et al. Chemical inhibition of bacterial protein tyrosine phosphatase suppresses capsule production. PloS ONE 7(5):e36312 (2012).

Appendix 1

Table 1: Bacterial strains used in this work

Strains	Relevant characteristics	Source/Reference
E. coli		
DH5α	Cloning strain	Lab strain
BTH101	Reporter strain for BTH system, F- cya-99, araD139, galE15, galK16, rpsL1 (Str[r]), hsdR2, mcrA1, mcrB1	Karimova *et al.* (1998) [46]
S. pneumoniae		
R36A	Unencapsulated laboratory strain	Avery *et al.* (1944) [50]
ATCC6314	Encapsulated strain, serotype 14	American Type Culture Collection
BCSMH001	Strain ATCC6314 null mutant for the *wzd* gene	Henriques *et al.* (2011) [38]
BCSMH002	Strain ATCC6314 null mutant for the *wze* gene	Henriques *et al.* (2011) [38]
BCSMH004	Strain ATCC6314 with *wze* substituted by *wze_citrine* derivative	Henriques *et al.* (2011) [38]
BCSMH007	R36ApBCSMH004, Tet[r]	Henriques *et al.* (2011) [38]
BCSMH016	ATCC6314pBCSMH004, Tet[r]	Henriques *et al.* (2011) [38]
BCSMH023	ATCC6314Δ*wze*pBCSMH004, Tet[r]	Henriques *et al.* (2011) [38]
BCSJF001	ATCC6314pBCSJF001, Tet[r]	Henriques *et al.* (2012) submitted [42]
BCSJF002	R36ApBCSJF002, Tet[r]	Henriques *et al.* (2012) submitted [42]
BCSJF003	R36ApBCSJF003, Tet[r]	Henriques *et al.* (2012) submitted [42]
BCSJF004	R36ApBCSJF004, Tet[r]	Henriques *et al.* (2012) submitted [42]
BCSJF005	R36ApBCSJF001, Tet[r]	This work
BCSJF006	R36ApBCSJF005, Tet[r]	This work
BCSJF007	ATCC6314pBCSJF003, Tet[r]	This work
BCSJF008	ATCC6314pBCSJF005, Tet[r]	This work
BCSJF009	ATCC6314Δ*wze*pBCSJF001, Tet[r]	This work
BCSJF010	ATCC6314Δ*wzd*pBCSJF003, Tet[r]	This work
BCSJF011	ATCC6314Δ*wzd*pBCSJF005, Tet[r]	This work
BCSJF012	ATCC6314*wze::wze_citrine*pBCSJF003, Tet[r]	This work
BCSJF013	ATCC6314*wze::wze_citrine*pBCSJF004, Tet[r]	This work
BCSJF014	ATCC6314*wzd*::iCFP_*wzdwze::wze_citrine*	This work
BCSJF015	ATCC6314*wzd*::*wzd*_CFP*wze::wze_citrine*	This work
BCSJF016	ATCC6314*wzd*::*wzd*_CFP*wze::wze_citrine* pBCSMH017, Tet[r]	This work
BCSJF017	ATCC6314*wzd*::*wzd*_CFP*wze::wze_citrine* pBCSLF001, Tet[r]	This work
Lactococcus lactis		
LL108	RepA[+] MG1363, Cm[r]	Leenhouts *et al.* (1996) [51]

Table 2: Bacterial plasmids used in this work

Plasmids	Relevant characteristics	Source/Reference
Plasmids for BTH assay in E. coli		
pKT25	Expression of T25 fragment as C-terminal fusions, Kan[r]	Karimova *et al.* (1998) [46]
pUT18C	Expression of T18 fragment C-terminal fusions, Amp[r]	Karimova *et al.* (1998) [46]
pKT25zip	BTH control plasmid, Kan[r]	Karimova *et al.* (1998) [46]
pUT18Czip	BTH control plasmid, Amp[r]	Karimova *et al.* (1998) [46]
pBCSMC001	Expression of T25-Wzd protein as C-terminal fusions, Kan[r]	Henriques *et al.* (2011) [38]
pBCSMC002	Expression of T18-Wze protein as C-terminal fusions, Amp[r]	Henriques *et al.* (2011) [38]
pBCSMC006	Expression of Wze-T18 protein as N-terminal fusions, Amp[r]	Carido and Filipe, unpublished data
pBCSMH037	Derivative of pUT18C plasmid in which T18 was removed, Amp[r]	Henriques, *et al.* unpublished data
pBCSJF006	Expression of Wze protein, Amp[r]	This work
pBCSJF007	Expression of Wze6histidines protein, Amp[r]	This work
pBCSMH038	Expression of T25-Wzd and T18-Wze protein as C-terminal fusions, Kan[r]	Henriques, *et al.* unpublished data
pBCSJF008	Expression of T25-WzdWze proteins as C-terminal fusions, Kan[r]	This work
pBCSJF009	Expression of WzeWze proteins as C-terminal fusions, Amp[r]	This work
pBCSJF010	Expression of T25-WzdWzeWzh proteins as C-terminal fusions, Kan[r]	This work
Replicative plasmids *in S. pneumoniae*		
pBCSMH002	Expression of pneumococcal proteins fusions to Citrine, Tet[r]	Henriques *et al.* (2011) [38]
pBCSMH004	pBCSMH002 containing *wze-citrine*, Tet[r]	Henriques *et al.* (2011) [38]
pBCSMH018	Expression of pneumococcal protein fusions to CFP, Tet[r]	Henriques *et al.* (2012) submitted [42]
pBCSMH031	Derivative of pBCSMH018 allowing expression of CFP containing the first 10 a.a. of Wze at N-terminus, Tet[r]	Henriques *et al.* (2012) submitted [42]
pBCSMH017	Derivative of pBCSMH004 in which citrine was removed, Tet[r]	Henriques *et al.* (2011) [38]
pBCSJC001	Expression of pneumococcal proteins as N-terminal fusions to Citrine improved with 10aa. of Wze, Tet[r]	Henriques *et al.* (2011) [38]
pBCSLF001	High-copy-number vector, contains the -10 constitutive promoter of SigA from S. pneumoniae, Tet[r]	Henriques *et al.* (2011) [38]
pBCSJF001	Derivative of pBCSJC001 containing *icitrine-wze*, Tet[r]	Henriques *et al.* (2012) submitted [42]
pBCSJF002	Derivative of pBCSMH002 containing *citrine-wze*, Tet[r]	Henriques *et al.* (2012) submitted [42]
pBCSJF003	Derivative of pBCSMH031 containing *iCFP-wzd*, Tet[r]	Henriques *et al.* (2012) submitted [42]
pBCSJF004	Derivative of pBCSMH018 containing *CFP-wzd*, Tet[r]	Henriques *et al.* (2012) submitted [42]

pBCSJF005	Derivative of pBCSMH018 containing *wzd-CFP*, Tet[r]	This work
Non replicative plasmids *in S. pneumoniae*		
pORI280	Em[r], LacZ[+], ori[+] of pWV01	Leenhouts *et al.* (1996) [51]
pBCSJF011	pORI280 containing the 3′ of *wzh*, the entire CFP sequence, and the 5' region of *wzd*, Em[r], LacZ[+]	This work
pBCSJF012	pORI280 containing the 3′ of *wzd*, the entire CFP sequence, the downstream region of *wzd*, and the 5' region of *wze* Em[r], LacZ[+]	This work

Table 3: Primers used in this work

Primer	Sequence 5'→3'	Features/ restriction sites
1	AAAA<u>CTGCAG</u>TAAGGAGGTGCTGCGGCCATGCCGACATTAGAAATAGC	PstI
2	CGG<u>GGTACC</u>TTATTTTTTTACCATAATTTCC	KpnI
3	GG<u>GGTACC</u>TCA**GTGGTGGTGGTGGTGGTG**CTCGAGAGAACCTCTAGGTACTAATTTTTTACCATAATTTCC	6histines / KpnI
4	CTAG<u>TCTAGA</u>GATGAAGGAACAAAACACTTTGG	XbaI
5	G<u>GAATTC</u>CCCTAAGTTATTTTTTACC	EcoRI
6	GGG<u>GAATTC</u>GGATGCCGACATTA	EcoRI
7	GGG<u>GGTACC</u>CCGGGCTAAGTAAT	KpnI
8	GGG<u>CTCGAG</u>GATGATAGACATCCATTCGCA	XhoI
9	GGGG<u>GGTACC</u>GCCGGCCTAAATTAATTGATCCATTATAA	KpnI
10	CCG<u>GGTACC</u>AATTCGGCCGTCGTTTTACAACGTCGTG	KpnI
11	CG<u>CTCGAG</u>AGAACCTCTAGGTACTAATTTTTTACCATAATTTCCATAGGAAGC	XhoI
12	ATAAGAAT<u>GCGGCCGC</u>AATGCCGACATTAGAAATAGC	NotI
13	CATG<u>AGATCT</u>TTATTTTTTTACCATAATTTCC	BglII
14	CTA<u>GCTAGC</u>ATGAAGGAACAAAACACTTTGG	NheI
15	GGG<u>GGTACC</u>TTTCAACTTACTCAAG	KpnI
16	ATTT<u>GCGGCCGC</u>TCATATGATGAAGGAACAAAACAC	NotI
17	CATG<u>CCATGG</u>CTATTTCAACTTACTCAAG	NcoI
18	CC<u>GAATTC</u>GGATCTAAAGAG	EcoRI
19	CGGGGCAGGTTAGTGACATTAG	
20	CGC<u>GGATCC</u>GGCGCAGAGATATACTATAC	BamHI
21	CTAATGTCGGCATATTTTCTCCTAAATTAATTGATCCATTATAATTTTTCTGGG	Region overlaping primer 22
22	CAATTAATTTAGGAGAAAATATGCCGACATTAGAAATAGCAC	Region overlaping primer 21
23	CCG<u>GAATTC</u>GTGATAATTACCGCATCAATAAC	EcoRI
24	CGC<u>GGATCC</u>GAAGTCTGGTTTAACCAATC	BamHI
25	CTCCCAAAAGTGTCATCTGCTTACTTGTACAGCTCGTCCATGCCGAGAGTG	Region overlaping primer 26
26	GGACGAGCTGTACAAGTAAGCAGATGACACTTTTGGGAGTTGTACC	Region overlaping primer 25
27	CCG<u>GAATTC</u>CCGCATCAATAACAATTCCAACAGG	EcoRI