UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

RECOGNITION OF *BACILLUS SUBTILIS* RECEPTORS BY BACTERIOPHAGE SPP1 – A GENETIC AND BIOCHEMICAL APPROACH

CATARINA DUARTE GALHARDO BAPTISTA

TESE ORIENTADA POR:

PROF. DOUTOR MÁRIO DE ALMEIDA SANTOS

DOUTOR PAULO TAVARES

DOUTORAMENTO EM BIOLOGIA
ESPECIALIDADE BIOLOGIA MOLECULAR

2009

O trabalho da presente dissertação foi realizado no Instituto para a Ciência Aplicada e a Tecnologia (ICAT), na Faculdade de Ciências da Universidade de Lisboa (FCUL) e na Unité de Virologie Moléculaire et Structurale (VMS) do Centre National de la Recherche Scientifique (CNRS) em Gif-sur-Yvette, França. Este trabalho teve o apoio da Fundação para a Ciência e a Tecnologia através da concessão de uma bolsa de doutoramento à autora (SFRH/BD/19675/2004), co-financiada pelo POCI 2010 e Fundo Social Europeu (FSE), e através do financiamento dos projectos onde esteve envolvida (POCTI/BIA-MIC/57088/2004 e PTDC/BIA-MIC/66412/2006). Deslocações ao estrangeiro e participação em reuniões científicas receberam financiamento do Programa Pessoa (GRICES) e da Fundação Luso-Americana para o Desenvolvimento (FLAD), respectivamente.

Na presente dissertação apresentam-se os trabalhos das seguintes publicações:

Baptista, C., Santos, M. A. and São-José, C. (2008) Phage SPP1 reversible adsorption to *Bacillus subtilis* cell wall teichoic acids accelerates virus recognition of membrane receptor YueB. J. Bacteriol. 190: 4989-4996.

São-José, C., Baptista, C. and Santos, M. A. (2004) *Bacillus subtilis* operon encoding a membrane receptor for bacteriophage SPP1. J. Bacteriol. 186: 8337-8346.

Para efeitos do disposto no n.º 2 do Artigo 8º do Decreto-Lei n.º 388/70, a autora da dissertação declara que participou na concepção e execução do trabalho experimental descrito, bem como na análise e discussão dos resultados e na redacção dos textos.

"The capacity to blunder slightly is the real marvel of DNA. Without this special attribute, we would still be anaerobic bacteria and there would be no music…". Lewis Thomas

Para os meus pais e irmão.

Adoro-vos!

ACKNOWLEDGMENTS

I would like to thank deeply to my supervisors Prof. Dr. Mário de Almeida Santos and Dr. Paulo Tavares for their scientific support and for this amazing opportunity to join their groups and study bacteriophages and their biology. It has been an incredible journey!

I am most indebt to Dr. Carlos São-José for his support, commitment, advises, supervising and friendship during this period. He was a key element and without him this would have been much harder… Thank you for believing in me since day one (or two…).

I would like also to express my gratitude to Dr. Isabelle Auzat for her indispensable experimental and scientific support in the gp21 work that was done at CNRS.

I am grateful to V. Lazarevic, R. Losick, J. D. Helmann and E. Brown for the gift of *B. subtilis* cell wall mutants. I thank K. Nishihara for the kind gift of chaperone vector pG-KJE3.

I thank J. G. Nascimento for valuable discussions of the work. He is a very good scientist with an amazing capacity to store apparent unusable information that has helped me in several occasions!

I have to write a special word to Dr. Ricardo Parreira. Besides your permanent friendship and helpful advices, if it wasn't for you I would probably not be writing this thesis now. A million thanks for that day, back in the summer of 2004, where you helped me compose the most beautiful letter for FCT… I hope I have not disappointed you in any circumstances.

To all my lab colleagues/friends, António, Bernardo, Carolina, Filipe, Ivo, Líbia, Liliana, Margarida, Petra, Ricardo, Sandra and Susana (and Mário, of course). Sharing the lab with you was always a pleasure. Without your company this would not have been half as fun! Also, to all the people in ICAT who helped throughout these years, thank you.

Thanks to the people from Dr. Paulo Tavares laboratory at CNRS. Sandrine, it was a pleasure knowing you and its amazing how you remember my birthday every year! Isabelle Petitpas, thank you for your collaboration and for anti-gp21 N-ter antibodies. Leonor Oliveira and Inês Vinga, you have always welcomed me. Thank you for your friendship and help during those periods. You helped to make my stays so much pleasant.

I would also like to thank my students, who had to put up with me in some difficult phases of this work. To them, Claúdia Vicente and Filipa Marta, sorry! Wish you all the luck in your science carriers.

To all my friends who always believed in me and which have been a fantastic support extra-lab. Francisca, Rui and Eurico, we have been a part of each others lives for the past twenty five years. It is a long time… hope we face a lot more! Your friendship means everything to me. To all my dear friends André, Catarina, Fred, Gonçalo, Helena, Inês, João, João Francisco, Lena, Leonor, Marta, Marta Joana and Rita... Beijinhos grandes! I cannot thank you enough for your friendship… Hope you enjoy reading this manuscript. And you still have to explain me the nickname "amoeba"; after all these time you still haven't got it… It's phages, people!

Finally some words to my family. I can not express how much I love you all and what you mean to me. Thank you so much for your words, advices and just for caring. A special thanks to my mother Georgete, father Acácio and brother João, who are wonderful people and I love so much. To Carlos, my special someone… sorry for my stress behaviour in the last months (or years…). Thank you for your dedication and for being there. I love you. And to my very special girl Matilde. I also have to apologize to you for the numerous weekends we stayed home because I had to write this thesis. Still you would sit next to me and start your own thesis "Eu gostava de ser cozinheira num restaurante de França muito famoso com 21 cozinheiros…" I wish you all the best and hope all your dreams come true and that I am there to witness and love you.

Thank you all!

RESUMO

Reconhecimento dos Receptores de *Bacillus subtilis* pelo Bacteriófago SPP1 – uma Abordagem Genética e Bioquímica

Bacteriófagos ou fagos são vírus que infectam especificamente bactérias. São parasitas obrigatórios com diferentes graus de dependência de funções celulares codificadas pelo seu hospedeiro. A grande maioria dos fagos descritos possui uma cauda ligada a uma nucleocápside que, por sua vez, contém o genoma de ADN em dupla cadeia. O início da infecção fágica compreende uma série de eventos que culminam com a entrada do genoma viral no citoplasma bacteriano. Estes eventos podem ser divididos, por uma questão meramente prática, em três passos principais: adsorção fágica à superfície bacteriana, ejecção do genoma da partícula viral, e finalmente a sua entrada no citoplasma da célula hospedeira. O passo da adsorção envolve a interacção específica entre proteínas estruturais da partícula viral (proteínas de ligação aos receptores, RBP) e componentes do hospedeiro expostos à superfície celular (receptores). Na grande maioria dos sistemas estudados pode-se diferenciar dois tipos de adsorção: primeiro os fagos adsorvem reversivelmente a um componente do invólucro celular, normalmente um componente grandemente representado, e depois adsorvem irreversivelmente a um outro (ou ao mesmo) receptor, ligação esta que compromete o fago à infecção.

Com este trabalho propusemo-nos aprofundar o actual conhecimento sobre o primeiro passo da infecção fágica de hospedeiros Gram-positivos, usando para tal a bactéria modelo *Bacillus subtilis* e o fago virulento SPP1. O bacteriófago SPP1 é um fago com cauda que pertence à família *Siphoviridae*. Possui uma cauda longa e não contráctil, sem apêndices, e que termina numa longa fibra distal. SPP1 reconhece a proteína membranar YueB do hospedeiro para adsorver irreversivelmente e infectar *B. subtilis*.

A análise do contexto genómico do gene *yueB* sugeriu que este podia fazer parte de um operão, conjuntamente com os genes *yukE*, *yukD*, *yukC* e *yukBA* situados a montante. Esta estrutura em operão é suportada por vários estudos apresentados nesta tese. De maneira a avaliar uma possível função na infecção por SPP1, cada gene do operão yukE foi estrategicamente inactivado e, cada mutante assim produzido, estudado tendo em conta a adsorção fágica, eficiência de plaqueamento e fenótipo das placas fágicas. Em relação a estes parâmetros apenas observámos um efeito negativo no tamanho das placas fágicas. Este fenótipo poderia ser explicado se genes

do operão desempenhassem um papel na infecção de SPP1 num passo pós-adsorção. Ainda que esta questão não tenha sido completamente esclarecida, os resultados obtidos parecem não apoiar esta hipótese. O que parece claro é que *yueB* é o único gene do operão essencial para a infecção de SPP1.

Embora a investigação do papel celular do operão yukE não estivesse contemplada nos objectivos deste trabalho, a análise bioinformática dos seus produtos revelou homologias com componentes do sistema de secreção do tipo VII (T7SS) de *S. aureus*, nomeadamente com o *locus* Ess. Os T7SS são maquinarias celulares especializadas no transporte de proteínas extra-celulares e são usados por bactérias patogénicas, como *M. tuberculosis* e *S. aureus*, para activamente secretarem os seus factores de virulência. T7SSs putativos estão presentes em várias bactérias não patogénicas e em bactérias patogénicas sem serem implicados na virulência. Apesar de não estar presentemente nenhuma informação disponível sobre a localização celular dos T7SSs, o facto destes sistemas incluírem um membro da família FtsK/SpoIIIE pode sugerir uma localização preferencial deste complexo ao nível da superfície celular. A exposição de YueB à superfície celular foi prevista por análise bioinformática e confirmada experimentalmente pelo nosso grupo. Neste trabalho utilizámos fusões à proteína GFP para estudar a sua localização subcelular. Os resultados apresentados mostram que YueB se distribui por todo o invólucro celular. Contudo, a análise mais precisa das imagens de microscopia de fluorescência evidenciaram zonas de maior fluorescência nos pólos e na zona central da célula (futura região polar).

Embora resistentes à infecção por SPP1, os mutantes de *B. subtilis* no gene *yueB* ainda apresentam uma elevada taxa de adsorção. Esta adsorção é contudo completamente reversível. Neste trabalho desenvolvemos estudos genéticos e de cinética de adsorção, quer para identificar os receptores envolvidos na adsorção reversível de SPP1, quer para avaliar a importância deste passo no processo geral de adsorção. Os estudos de adsorção efectuados com mutantes de *B. subtilis* deficientes na biosíntese dos principais polímeros da parede indicaram que SPP1 utiliza como receptor reversível o polímero glucosilado de poliglicerol fosfato, o qual constitui o ácido teicóico maioritário da parede (WTA). A necessidade da glucosilação dos WTA na adsorção de SPP1 foi confirmada pela utilização da lectina Concanavalina A, que se liga especificamente aos resíduos de α-D-glucose na parede, no tratamento de células de *B. subtilis*.

Os estudos de cinética de adsorção mostraram que a taxa inicial de ligação reversível dos fagos não varia grandemente com a temperatura. Pelo contrário, a taxa de dissociação aumenta

exponencialmente com a temperatura, evidenciando uma dependência que segue a lei de Arrhenius, e que permitiu calcular um valor de energia de activação de 22.6 kcal/mol para a libertação fágica. O fago SPP1 é capaz de plaquear nos mutantes de *B. subtilis* afectados no passo de adsorção reversível, embora com um impacto na eficiência de plaqueamento e na morfologia das placas fágicas que depende do tipo de mutação testada. Em meio líquido a taxa de adsorção irreversível a estes mutantes foi drasticamente afectada, demonstrando uma importância crítica da adsorção reversível no eficiente reconhecimento do receptor YueB. A necessidade da adsorção reversível foi suplantada com uma superprodução de YueB, indicando que este primeiro passo de ligação não constitui um pré requisito essencial que "activa" o fago para a ligação subsequente a YueB. O resultado mais interessante deste estudo foi a descoberta de que as taxas de adsorção irreversível são principalmente determinadas pela taxa de libertação dos fagos dos receptores reversíveis, de modo que uma libertação mais rápida permite uma maior adsorção irreversível. No seu conjunto estes resultados permitiram-nos propor um modelo para a adsorção de SPP1 em que os ciclos rápidos de adsorção/libertação dos WTA glucosilados permitem o fago "examinar" rapidamente a superfície celular para a ligação ao YueB. Este modelo poderá aplicar-se a outros sistemas em que dois receptores diferentes estejam envolvidos em cada tipo de adsorção.

Com este trabalho também se pretendeu investigar as proteínas ou estruturas da cauda de SPP1 que especificamente reconhecem os receptores em *B. subtilis* (YueB e WTA glucosilados). Nos fagos com cauda a actividade de ligação aos receptores está codificada nas subestruturas da cauda, tais como as placas basais, fibras e espículas. O produto do gene *21* foi determinado como sendo ou fazendo parte da estrutura que interage com YueB, após uma captura selectiva de proteínas fágicas baseada na ligação específica a uma matriz de afinidade de YueB. Na tentativa de confirmar *in vitro* a interacção entre gp21 e YueB, produziram-se e purificaram-se parcialmente fusões da proteína gp21. Estudos preliminares parecem suportar esta ligação. Prevê-se que a proteína gp21 seja o principal constituinte da fibra terminal de SPP1. Neste trabalho estávamos também interessados em perceber se o fago SPP1 utiliza a mesma RBP para os dois tipos de adsorção. Tendo por base os resultados que obtivemos sobre a cinética de adsorção de SPP1 elaborámos um protocolo para isolar selectivamente mutantes de SPP1 que não possuem a capacidade para adsorver reversivelmente. A sequenciação do gene *21* em vários destes mutantes revelou duas mutações distintas na região que codifica para extremidade C-terminal da proteína. Estas mutações foram corrigidas após recombinação com uma cópia selvagem do gene *21*, conferindo a estes fagos recombinantes a capacidade de adsorção reversível. Os resultados obtidos parecem indicar que a mesma RBP de SPP1, gp21, está envolvida na adsorção reversível

e irreversível. Este resultado é suportado pela análise bioinformática que evidencia semelhanças estruturais entre a gp21 e RBPs conhecidas de outros fagos.

Palavras chave: adsorção reversível; adsorção irreversível; proteínas fágicas de ligação aos receptores; receptores fágicos; cinética de adsorção; poliglicerol fosfato glucosilado; fibra da cauda.

ABSTRACT

Recognition of *Bacillus subtilis* Receptors by Bacteriophage SPP1 – a Genetic and Biochemical Approach

Bacteriophages or phages are viruses that infect bacteria. They are obligatory parasites with different extents of dependence from cellular functions of the host. The vast majority of known phages possess a tail attached to a nucleocapsid that contains the dsDNA genome. The onset of phage infection comprehends a series of events that lead to viral genome delivery to the bacterial cell cytoplasm. For study purposes these can be presented as three sequential major steps: phage adsorption to bacterial surface, genome ejection from the virus particle, and finally its entry into the host cell. The adsorption step involves the specific interaction between phage structural proteins (receptor binding proteins, RBP) and surface-exposed components (receptors) of the host cell envelope. In the great majority of the studied systems two types of adsorption can be considered: first, phages adsorb reversibly to a component of the cell surface, usually one that is largely represented, and secondly, irreversible adsorption occurs to another (or the same) receptor that commits phage to infection.

With this work we aimed to deepen the understanding of the initial step of phage infection of Gram-positive hosts using the model bacterium *Bacillus subtilis* and its virulent phage SPP1. Bacteriophage SPP1 is a tailed phage that belongs to the *Siphoviridae* family. It possesses a long, non-contractile tail with a long distal tail fiber and no appendages. SPP1 uses the host cell membrane protein YueB to irreversibly adsorb and infect *B. subtilis*.

The analysis of the *yueB* genomic context suggested that it could be part of an operon structure together with the upstream genes *yukE*, *yukD*, *yukC* and *yukBA*. Several studies presented in this thesis support this operon organization. To assess its possible role in SPP1 infection we have strategically disrupted each gene of the yukE operon and studied each mutant regarding phage adsorption, efficiency of platting and phage plaque phenotype. Concerning these parameters we only observed a negative effect in phage plaque size, a phenotype that could be explained if some genes of the operon played a role in a post-adsorption step of the phage infection cycle. Although still being an open question the results obtained seem not to support this hypothesis. What seems clear is that *yueB* is the only gene of the operon that is essential for SPP1 infection.

Although investigation of the cellular role of the yukE operon was out of the scope of this work, the bioinformatics analysis of its products revealed similarities with components of the type VII secretion system (T7SS) of *S. aureus* locus Ess. T7SS are specialized apparatus for the transport of extracellular proteins and are used by pathogenic bacteria, such as *M. tuberculosis* and *S. aureus* to actively secrete virulence factors. Putative T7SSs are present in several non-pathogenic bacteria and in pathogenic bacteria without being used for virulence. Although no information on the cellular localization of T7SSs is available, the fact that such systems comprise an FtsK/SpoIIIE-like member might suggest a preferable subcellular localization of this complex at the bacterial cell surface. The surface localization of YueB was predicted by bioinformatics analysis and confirmed experimentally by our group. In this work we studied its subcellular localization by GFP fusions. The results showed that YueB distributes all over the cell envelope. However, acute analysis of the fluorescence microscopy images suggested small patches localized at the cell poles or at mid-cell region (future polar regions).

Although resistant to SPP1 infection, *B. subtilis yueB* mutants still exhibit a high rate of phage adsorption. This binding is however completely reversible. We have performed genetic and adsorption kinetics studies to identify the receptor(s) involved in SPP1 reversible binding and to evaluate the relevance of this interaction in the whole adsorption process. Adsorption studies performed with *B. subtilis* mutants affected in the biosynthesis of the major cell wall polymers indicated that SPP1 uses as receptor for reversible adsorption the glucosylated polyglycerol phosphate, the principal cell wall teichoic acid (WTA). The requirement of WTA glucose substitution for SPP1 reversible adsorption was further evidenced by treating *B. subtilis* cells with Concanavalin A, a lectin that specifically targets α-D-glucosyl residues in the cell wall.

The adsorption kinetics studies showed that the initial rate of reversible adsorption did not vary significantly with temperature whereas phage dissociation showed a temperature-dependent increase that followed an Arrhenius law. This allowed us to determine an activation energy of 22.6 kcal/mol for phage release. Phage SPP1 could plate in the *B. subtilis* mutants affected in reversible adsorption although with impact in the efficiency of platting and phage plaque morphology that depended on the type of mutation. In liquid medium the rate of irreversible adsorption to these mutants was severely impaired, demonstrating that the step of reversible adsorption is critical for efficient recognition of receptor YueB. Interestingly, overproduction of YueB bypassed the requirement for reversible adsorption, indicating that this first interaction is not an essential pre-requisite that "activates" the phage for the subsequent irreversible step. The

most striking result of this study was the finding that the rate of SPP1 irreversible adsorption is mainly determined by the rate of phage release from reversible receptors with faster dissociation promoting higher irreversible adsorption. Overall, the results allowed us to propose a model where fast adsorption/desorption cycles to glucosylated WTA allows SPP1 to rapidly scan the cellular surface for YueB binding. This model likely applies to other systems where two different receptors are involved in the two-step adsorption process.

In this work we also searched for the SPP1 proteins or tail structures that specifically target the *B. subtilis* receptors (YueB and glucosylated WTA). In tailed phages, receptor binding activity is carried in tail substructures such as baseplates, fibers and spikes. We have determined the product of gene *21* as being the, or part of the, structure that interacts with YueB after selective capture of phage proteins based on specific binding to a YueB-affinity matrix. As an attempt to confirm *in vitro* the gp21/YueB interaction we have produced and partially purified gp21 fusions. The preliminary studies seem to support gp21 binding to YueB. gp21 is predicted to be the major constituent of the SPP1 tail fiber. We were also interested to disclose if SPP1 uses the same or distinct RBPs in each type of adsorption. Based on the knowledge obtained on the kinetics of SPP1 adsorption, we developed a protocol to selectively isolate SPP1 mutants impaired in reversible adsorption. Sequencing of gene *21* in several mutants revealed two distinct mutations in the region coding for the protein C-terminal end. These mutations were corrected after recombination with a wild type copy of gene *21*, which restored normal adsorption to the recombinant phages. Thus, the results obtained seem to indicate that the same SPP1 RBP, gp21, is involved both in reversible and irreversible adsorption. This result is supported by gp21 bioinformatics analysis that suggests structural similarities with known RBPs from other phages.

Key words: reversible adsorption; irreversible adsorption; phage receptor binding proteins; phage receptors; kinetics of adsorption; glucosylated polyglycerol phosphate; tail fiber.

TABLE OF CONTENTS

PART III – PHAGE SPP1 REVERSIBLE ADSORPTION

PART IV – PHAGE SPP1 RECEPTOR BINDING PROTEINS

PART I

GENERAL INTRODUCTION

I.1. AN HISTORICAL VIEW ON BACTERIOPHAGES

Since the discovery of bacteriophages (phages) in the early 20th century these "virus parasitic on bacteria" as Felix d'Herelle wrote, have been in the origin and part of the major advances in the world of molecular biology and molecular genetics. Soon after the independent discovery by Frederick Twort (1915) and d'Herelle (1917), the later started using these yet uncharacterized filtrates to initiate what today is known as phage therapy. He begun in 1919 treating chicken with typhus and successfully moved to heal dysenteric human patients. In the subsequent years, d'Herelle spread phage therapy throughout the world from India to Egypt and, five years before Helmut Ruska first saw phages under the electron microscope, he joined George Eliava and founded the today's Eliava Institute of Bacteriophage, Microbiology, and Virology in Tiblisi, Georgia (1934). The report of penicillin in 1928 by Alexander Fleming started the era of antibiotic discovery and by the late 30s phage therapy was but an old fashion idea. Phage biologists around the world started using bacteriophages as model systems *per se* rather than therapeutic agents. Early contributions in phage biology came from names like Martin Schlesinger, who showed the chemical nature of these viruses as formed of proteins and nucleic acids, Sir Macfarlane Burnet that identified several species of phages and reported their antigenic properties, and Max Delbrück, a physicist fascinated by the simplicity of doing "experiments in something like atoms in biology" (Hayes, 1984). Delbrück was introduced to phage biology by Emory Ellis and in 1940 met Salvador Luria, an Italian microbiologist refugee, and they started working together in bacteria resistance to phage infection. Their paper in 1943 stating that phage resistance arises in bacteria by mutation is considered nowadays and the beginning point of molecular biology. These two phage biologists met Alfred Hershey and soon after the three formed a nucleus known as the Phage Group. The Phage Group expanded and based their work on three fundamental areas of phage biology, i.e., the mechanisms of attack, multiplication and lysis (Pennazio, 2006). Lysogeny was left under the study of André Lwoff in Paris. Here started the phage era and during the 40s and 50s major advances were obtained in molecular and phage biology including the contribution to the understanding of the physic nature of genes (Calendar, 2007). After the disclosure of the molecular structure of DNA by Francis Crick and James Watson in 1953, both members of the Phage Group and the later a Ph.D. student of Luria, our view on biological problems turned to a different direction. In the next years the finding that RNA, rather than DNA, is directly involved in protein synthesis, the discovery of plasmids and restriction enzymes opened a new period in molecular biology, the Recombinant DNA era. The legacy of phages ranged through disclosure of genome replication, mRNA and transcription,

regulation of protein synthesis and modification, and macromolecular assembly mechanisms. Phages have also been proved to be involved in pathogenesis, bacterial evolution and ecology (Calendar, 2007). More recently, the deep understanding of the molecular details of phage assembly pathways have placed them at the vanguard of structural biology (Johnson and Chiu, 2007). Finally, and somewhat unexpectedly, we presently assist to a renewed interest in the use of phages, or derived products, as potential agents to combat bacterial pathogens, namely those presenting resistance to antibiotherapy. Understanding bacteriophages and using their fantastic properties to explore molecular biology has gain phage biologists four Nobel awards and many more nominations (Norrby, 2008).

I.2. GENERAL FEATURES OF BACTERIOPHAGES

Phages are by far the most abundant biological entities on earth and probably the most diverse (Hendrix, 2002; Pedulla *et al.,* 2003). They are obligatory bacteria parasites that exist extracellularly as supramolecular structures, the virions. These structures are optimized to ensure viral propagation by protecting their genome and efficiently delivering it to the host bacteria. Once their genomic information is inside the bacterial cytoplasm, viral genome replication and protein synthesis takes place to assemble new infective virions. Except for the case of filamentous phages, the virion progeny is then released to the media through lysis of the host cell. In addition to this lytic pathway, some phages (temperate phages as opposed to virulent) have also the ability to integrate their DNA in the genome of the host, allowing the virus to persist within the bacteria progeny (Fig. I.1).

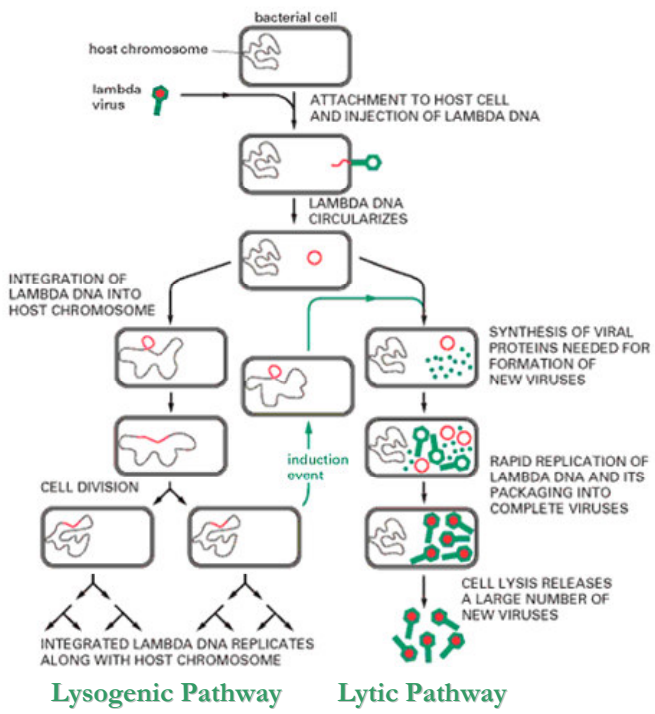


Figure I.1 | Lifestyles of bacteriophages. The temperate phage lambda is used as an example to illustrate the two possible lifestyle pathways. After recognition and efficient attachment to the host cell wall, phages deliver their genetic information to the cytoplasm. In the presented case, a circularization of the genome is needed both for integration in the host chromosome (lysogenic pathway) or for viral DNA transcription and replication (lytic pathway). The decision to enter the lysogenic state is taken immediately after the viral genome has reached the host cytoplasm. In the lysogenic lifestyle the integrated viral DNA remains in the host chromosome for several generations until a certain stimulus induces a genetic switch that causes the entry in the lytic cycle. After production of the phage proteins, phage particles are assembled and the newly replicated DNA is encapsidated. Complete viruses are released, usually by induced host cell lysis, causing cell death.

Phages were first classified, according to Bradley´s proposal, into six morphological types based on gross morphology and nucleic acid type (Bradley, 1967). Viral genomes can be circular or linear, double stranded (ds) DNA (vast majority of morphotypes), single stranded (ss) DNA (*Microviridae*, *Inoviridae*), dsRNA (*Cystoviridae*) or ssRNA (*Leviviridae*). Until today, at least 5,568 bacteriophages have been examined by electron microscopy for morphological analysis, 96% of

which possess a nucleocapsid attached to a tail (Ackermann, 2009). Other than tailed, phages can be polyhedral, filamentous, or pleomorphic (PFP). The three families of tailed phages constitute the order *Caudovirales* and the various families of PFP phages are not grouped into orders.

I.3. THE ORDER *CAUDOVIRALES*

The virions of this order consist of an icosahedral capsid, or head, that contains one molecule of linear dsDNA, attached to a tail that possesses structures involved in recognition and fixation to the host. For practical reasons rather than taxonomical, tailed phages were subdivided according to head shape, length and flexibility of the tail (Ackermann & DuBow, 1987). According to tail morphology tailed phages fall into three families, the *Myoviridae* (rigid and contractile tail; 25%), *Siphoviridae* (flexible, long and non-contractile tail; 61%), and *Podoviridae* (short tail; 14%), and each family has several genus (Table I.1). A few morphological types illustrating tailed phages variety are presented in Figure I.2. Virion assembly is complex and includes separate pathways for each major structure (Ackermann, 1998):

- The basic building units of phage capsids are the capsomers that appear as hexamers and pentamers. The icosahedral heads can be isometric (identical length and width) or prolate (length longer than width) and have a 5-fold symmetry at their vertices. One of these is asymmetrical and designated to serve as the portal vertex. Besides being more or less elongated, heads can have fibers distributed homogeneously or just on the vertices. Capsid assembly always involves a major head protein and a portal protein (the later one being part of the connector; see below). Although the use of a scaffold protein in capsid assembly is generally observed, exceptions occur as in the case of phage HK97, which uses a domain of the capsid protein as scaffold (Duda *et al.*, 1995). Accessory proteins from viral or host origin can also be involved. The assembly of the head is completed with the packaging of DNA that imperatively enters through the portal vertex powered by the ATPase activity of the terminase complex.

- The tails have generally a 6-fold symmetry with a variable number of stacked rows of subunits in a rod or helical-based shape, with a facultative sheath. Tails have several other facultative structures as collars (at the head-tail junction), baseplates, spikes, adornments and fibers, which may or may not have a fundamental function. Addition of tail is always the last step of the phage assembly. Tails of *Podoviridae* are assembled by sequential binding of proteins to the head portal vertex whereas tails of long-tailed phages are built separately and joined later to this head vertex. The tail is built from the distal end upwards (to the direction of the head) and the process is completed by addition of tail fibers or spikes. Length of the tail is determined by the tape measure protein around which tail tube monomers polymerise. The recognition and adsorption apparatus is localized in the tail.

- Located between the capsid and the tail, the connector brings them together. The connector is located at the portal vertex. It is usually composed of 12-fold symmetry small disk(s) with a hole; it localizes inside the head serving to correct the mismatch between the head and tail symmetries. The connector is as well of extreme importance to the head assembly and DNA encapsidation and serves also as the site for tail attachment and later for DNA ejection. It can also have fibers. Usually assembly of the connector sets the beginning of head assembly although it was suggested for phage T7 that the portal protein is added to an already assembled procapsid (Cirreteli and Studier, 1996).

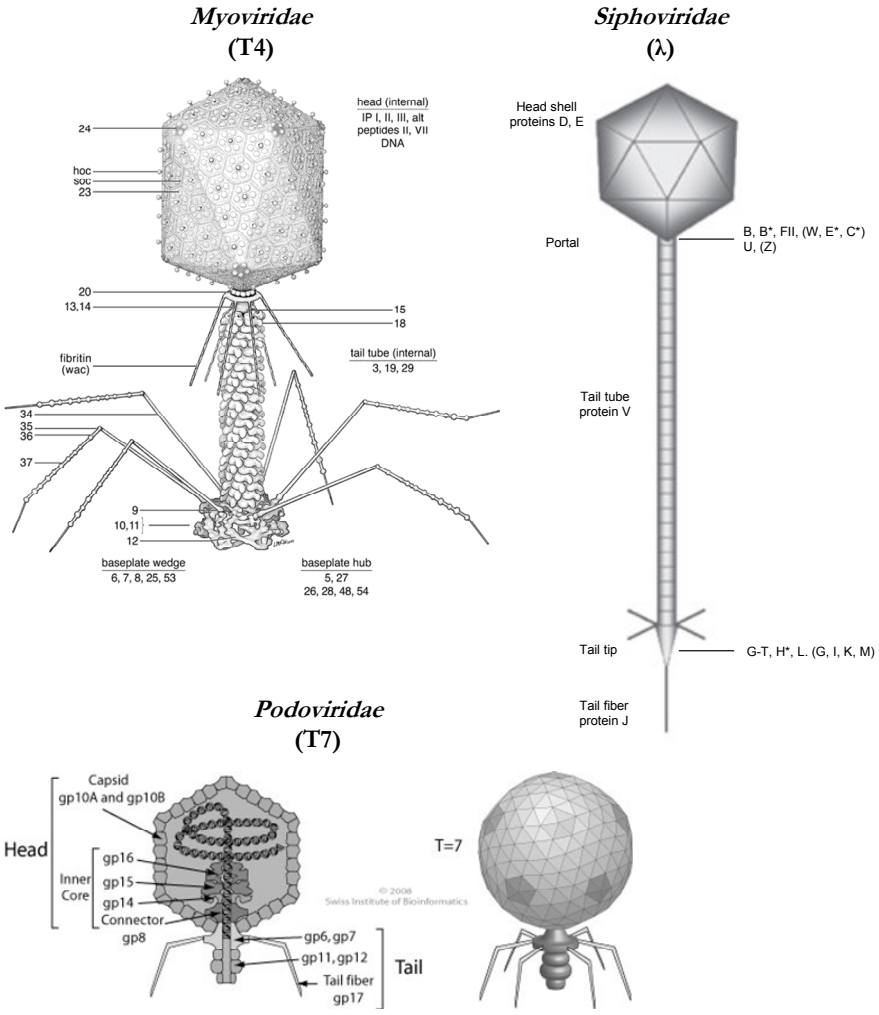


Figure I.2 | Schematic representation of tailed bacteriophage virions. One of the major studied examples of each family is presented: *Myoviridae* (T4), *Siphoviridae* (λ) and *Podoviridae* (T7). Phage T4 has a prolate and λ and T7 have isometric heads. Head and tail components are indicated as well as the proteins that constitute them. Phage protein numbers or letters are specific for the examples presented here and not for all members of each family. Pictures of T4 and T7 are from Leiman *et al.,* 2003 and http://www.expasy.ch/viralzone/all_by_protein/230.html, respectively. Image of phage λ is adapted from http://www.bigdna.com/index.html.

Table I.1 schematizes the organization of the order *Caudovirales*, the three families and the genus within each family, which include a large number of species. This includes not only the phages of eubacteria, but also the virus from archea. Whereas bacteriophages are a polyphyletic group, tailed phages appear as monophyletic and as the oldest known virus group (Ackermann, 2003). The three families of tailed bacterial viruses have been assigned based on similarities in virus morphology. The genera classification has been established for some species in each family, based on properties involving viral DNA replication and packaging, and on some features specific to particular genera (e.g., DNA-termini linked proteins, virus-encoded polymerases, and ability to establish temperate infections) (Maniloff and Ackerman, 1998). This classification scheme tries to reflect their evolutionary history (Fauquet *et al.,* 2004), which has been traced thanks to the impressive number of sequenced bacteriophage genomes accomplished in the last years. Comparative analysis of these genomes have highlighted a very high degree of genetic diversity, suggesting early evolutionary origins, and the particular mosaic nature of phage genomes, which reflects an unusually high degree of horizontal genetic exchange in their evolution (Hatful, 2008; Casjens, 2005; Hendrix *et al.,* 2003). In spite of this, enough common features survive to indicate their fundamental relatedness

Table I.1 | Taxonomic structure of the order *Caudovirales* (source Fauquet *et al.,* 2004 Virus Taxonomy: VIIIth Report ICTV, International Committee on Taxonomy of Viruses). In gray are the extra genus classified as such in the NCBI taxonomy database. Each family has a large number of unclassified phages.

Order	*Caudovirales*		
Family	*Myoviridae*	*Siphoviridae*	*Podoviridae*
Genus	T4-like Viruses	λ-like Viruses	T7-like Viruses
	P1-like Viruses	T1-like Viruses	P22-like Viruses
	P2-like Viruses	T5-like Viruses	ϕ29-like Viruses
	Mu-like Viruses	L5-like Viruses	N4-like Viruses
	SP01-like Viruses	c2-like Viruses	BPP-1--like Viruses
	ϕH-like Viruses	ψM1-like Viruses	ε-15-like Viruses
		ϕC31-like Viruses	119X-like Viruses
		N15-like Viruses	VP2-like Viruses
			Autographivirinae (includes T7-like Viruses)
			Picovirinae (includes ϕ29-like Viruses)

I.4. THE BACTERIAL CELL ENVELOPE: PHAGE LANDING FIELD AND BARRIER TO INFECTION

The cell envelope may be defined as the cell membrane and cell wall plus an outer membrane if one is present. The bacterial cell envelope is the first and major line of defence against threats from the environment. It has the major function of maintaining the cell homeostasis. It is an essential and yet vulnerable structure that provides the cell with structural integrity and counteracts the high internal osmotic pressure. It also provides an important sensory interface and molecular sieve, mediating both information flow and the controlled transport of solutes. The cell envelope is also the target for numerous antibiotics (Jordan *et al.,* 2008; Delcour, 2008). In a contradictory function, the cell envelope is at the same time a barrier to phage infection and a field of structures that are exploited by phages as receptors. The basic structure and composition of most bacterial cell envelopes fall into two major categories that are the base of Gram staining: Gram-negative and Gram-positive (Fig. I.3). Other types are found in a few bacterial species such as the Mycobacteria.

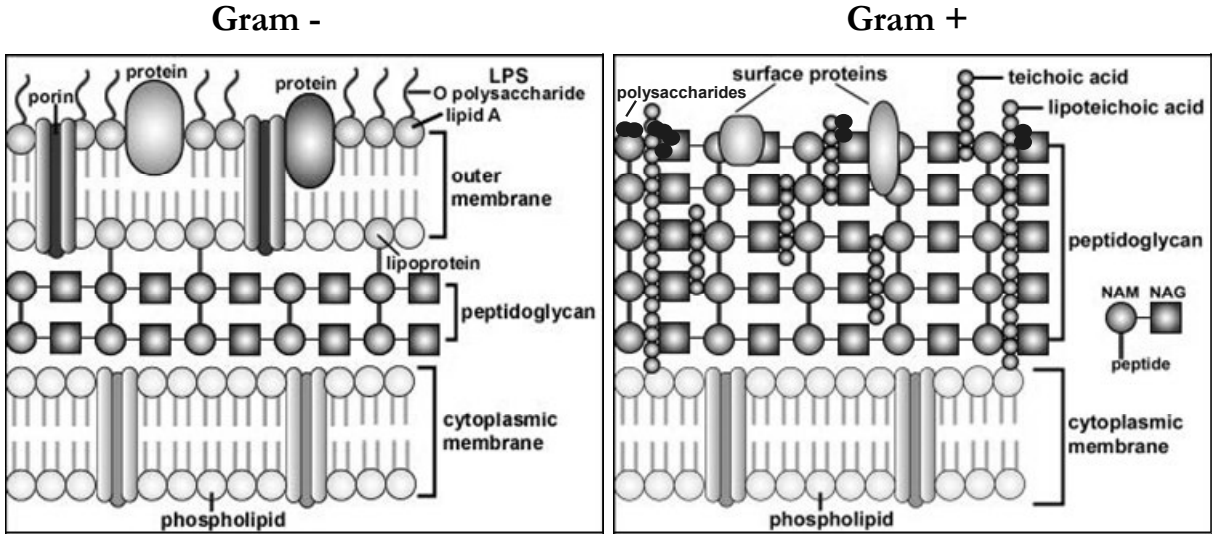


Figure I.3 | Schematic representation of the Gram-positive and -negative cell envelope architecture. The major constituents are indicated (see text for details). Scheme not drawn to scale. Pictures adapted from Prof. Gary E. Kaiser lectures available online at http://student.ccbcmd.edu/~gkaiser/welcome.html. Structure of Gram negative and Gram positive cell walls can be found at http://student.ccbcmd.edu/courses/bio141/labmanua/lab16/diseases/nmeningitidis/u1fig10b.html and http://student.ccbcmd.edu/courses/bio141/lecguide/unit4/innate/u1fig9b.html, respectively.

The cell envelope of Gram-negative bacteria consists of two membranes, the cytoplasmic or inner membrane (IM) and the outer membrane (OM), which are separated by the periplasm

containing the peptidoglycan layer (see below). The Gram-positive cell envelope consists only of two functional layers (compared with three in Gram-negative bacteria) that enclose the cellular contents: the inner membrane, surrounded by a thick cell wall composed of peptidoglycan. It lacks an outer membrane. In Mycobateria, its particular type of cell envelope is thicker than in many other bacteria, waxy, and rich in mycolic acids/mycolates. The cell wall consists of the hydrophobic mycolate layer and a peptidoglycan layer held together by a polysaccharide, arabinogalactan (Goude and Parish, 2008; Daffé and Draper, 1998).

The major constituents of the bacteria cell envelope will be presented below. Although an interesting and prolific field, biogenesis and regulation will not be discussed here as it is over the scope of this thesis.

I.4.1 – CYTOPLASMIC MEMBRANE

The cytoplasmic or inner membrane (IM) is a fluid phospholipid bilayer that surrounds the cytoplasm and protects it from the environment. It has associated or imbedded (integral) proteins, typically spanning the membrane in the form of hydrophobic α-helices. It also contains lipoproteins that remain attached to the phospholipids while facing the periplasm. The IM joins the functions of selective nutrient transport, protein translocation, lipid biosynthesis, and oxidative phosphorylation (Bos *et al.,* 2007; Doung *et al.,* 1997). Molecules can move across the bacterial IM by passive diffusion (osmose or facilitated diffusion) or by active transport (dispensing metabolic energy). Traffic of bigger molecules like DNA is carried out by specific transport proteins.

I.4.2 – CELL WALL

The term cell wall refers to the structures that surround the cytoplasmic membrane. The major component is the peptidoglycan (or murein), a highly cross-linked macromolecule that provides rigidity and counteracts the internal pressure of the bacterial cell, which is approximately 3 to 5 atmospheres in Gram-negative bacteria and can reach up to 25 atmospheres in Gram-positive bacteria (Jordan *et al.,* 2008). The peptidoglycan polymer is a glycan (polysaccharide) backbone consisting of β-(1,4) linked N-acetyl muramic acid (NAM) and N-acetyl glucosamine (NAG),

with peptide side chains linked to the former. These are composed of L-alanine, D-glutamic acid, D-alanine and either L- lysine or diaminopimelic acid (DAP), depending on species. The side chains are frequently cross-linked by peptide bridges between DAP and the D-Ala. In *Staphylococcus aureus* (Gram-positive) the cross-link is achieved by a peptide interbridge composed of five glycines that connect the D-Ala of one chain to the L-Lys of the other. The peptidoglycan meshwork has an average cut-off value for the passage of globular proteins of about 50 kDa and permits the passage of particles with approximately 2 nm (Demchick and Koch, 1996).

In Gram-negative bacteria the peptidoglycan consists of a thin layer with about 2.5 to 7.5 nm in thickness. The Gram-positive peptidoglycan sacculus is a multi-layered net-like structure of about 50-nm thickness that can withstand the high turgor pressures. The structural rigid sugar chains perpendicularly cross-linked with flexible peptide bridges, give this mesh a strong, but also elastic stress-bearing structure (Jordan *et al.,* 2008). It is a highly dynamic macromolecule that allows maintenance of cell shape, cellular growth and division at the same time (Archibald *et al.,* 1993).

Covalently bound to the sugar backbone of peptidoglycan in Gram-positive bacteria are teichoic acids or, functionally similar polymers such as teichuronic acids and lipomannan (Neuhaus and Baddiley, 2003). Teichoic acids are most frequently found as polymers of glycerol-phosphate but can also be constituted of ribitol-phosphate (e.g. *Listeria monocytogenes*, *Staphylococcus aureus*, *Bacillus subtilis* W23), galactosylglycerol-phosphate (*Bacillus coagulans*) or N-acetylglucosamine-phosphate (*Bacillus pumilus*) (Yokoyama *et al.,* 1989). Lipoteichoic acids (LTA) are structurally related to cell wall teichoic acids (WTA) but are tethered to the membrane by a glicolipid anchor with its poly(glycerol-phosphate) (Gro-P) chain extending into the wall. They are all poly-anionic, phosphate-rich linear polymers, mainly responsible for the overall negative net charge and physicochemical properties of the cell envelope (Bhavsar *et al.,* 2004). Teichoic acids are also known to be involved in the activity control of autolytic enzymes, one of the steps in cell wall biosynthesis (Neuhaus and Baddiley, 2003). They frequently incorporate α-glucosyl and D-alanyl substituents, the latter contributing to balance the negative charge of the cell surface.

Also embedded in the peptidoglycan meshwork are sugar polymers, proteins and lipoproteins. Peptidoglycan and anionic polymers constitute the bulk of the cell wall and are present in approximately equal proportions (Bahvsar and Brown, 2006). Most of the Gram-positive cell wall components referred above are employed by phages to adsorb to their host cells.

I.4.3 – OUTER MEMBRANE (GRAM-NEGATIVE BACTERIA)

Gram-negative bacteria are exclusively endowed with an outer membrane (OM). The OM is an asymmetrical bilayer that covers the peptidoglycan, consisting of phospholipids and lipopolysaccharides (LPS) in the inner and outer leaflet, respectively. It comprises both integral OM proteins (OMPs), which generally consist of antiparallel amphipathic β-strands, and lipoproteins, which are anchored to the membranes with the protein moiety usually facing the periplasm (Bos *et al.,* 2007). The OM is usually the first barrier in a Gram-negative bacterium and for that must function as a selective protection from external threats, such as antibiotics.

Unlike the IM, the OM is not energized by a proton gradient and ATP is not available in the periplasm. Thus, nutrients usually pass the OM by solute-specific permeases (e.g. LamB) or by passive diffusion via an abundant class of OMPs called porins. Porins are usually trimers forming water-filled channels that allow the passage of small hydrophilic solutes with molecular weights up to ~600 Da. Nevertheless, energy-requiring transport processes in the OM have also been described. Such processes are dependent on complex energy coupling systems, such as the TonB system, which couples the proton-motive force of the IM to receptor-mediated uptake processes in the OM (Wandersman and Delepelaire, 2004). The OMPs are one of the structures frequently exploited by phages to use as their receptors. In fact, the highly specific binding of phages to receptors has been exploited by geneticists to identify and characterize components of the bacterial surface. Several OMP have been identified this way, e.g. TonA (phage T1), Bfe (or BtuB; phage BF23), LamB (phage λ) and Tsx (phage T6) (Heller, 1992).

Lipopolysaccharides are complex glycolipids that contribute greatly to increase the negative charge of the cell membrane and help to stabilize the overall membrane structure also protecting from certain kinds of chemical attack. They are anchored to the outer membrane through their lipidic portion, known as lipid A, while decorating the cell surface with their hydrophilic constituents known as the core and O-antigen. These structures have significant variations in their composition among different species and also within several strains of the same species (especially considering O-antigen variations) (Bos *et al.,* 2007). They produce strong immunogenic responses in animals and are thus considered endotoxins. In lipid A, the principal responsible for the toxic properties, the fatty acids are linked to N- acetylglucosamine phosphate that in its turn binds to the ketodoxyoctonate (KDO) portion of the core. The rest of the core is composed of sugars and derivatives and, when there is one, links the O-antigen (or O-polysaccharide). This is

an often branched network of sugar sequences (Madigan *et al.,* 1997). LPS, especially the O-antigen, serve as receptors for many Gram-negative infecting phages, because of their privileged position on the cell surface.

Also important in the OM are the lipoproteins. They are anchored to the inner leaflet side via an amino-terminal N-acyl-diacylglycerylcysteine, with the protein moiety usually facing the periplasm. A very abundant group of these proteins, Braun's lipoproteins, link tightly the OM to the peptidoglycan layer, through covalent peptidoglycan bounds (Hayashi and Wu, 1990).

In Gram-negative bacteria, the space between the two membranes is named the periplasm. This gel-like region comprises the thin peptidoglycan layer and several other molecules, such as components of transport and substrate binding systems, proteins responsible for hydrolysis and reception of extracellular signals, chaperones, amongst others (Doung *et al.,* 1997). Although devoid of OM, a periplasmic space has also been referred to exist between the IM and the cell wall in Gram-positive bacteria (Bahvsar and Brown, 2006; Matias and Beveridge, 2005). Additionally, Matias and Beveridge (2008) observed by cryo-EM immuno labelled frozen-hydrated sections from *B. subtilis* that the major component of this "space" is in fact LTA.

I.4.4 – SURFACE STRUCTURES AND CELL INCLUSIONS

In addition to these basic features, a number of additional cell envelope structures are present in many bacteria. These structures often play an important role in virulence, antibiotic resistance, surface adhesion, colonization, protection of desiccation (and engulfment for some pathogens) and multicellular differentiation. They include extracellular polysaccharide capsules and slime layers, biofilm matrices, as well as proteinaceous S-layers. Although these are all structures designed to protect the bacterial cells from environmental stresses, bacteriophages have been reported to use the capsule as their docking field.

Optional structures also include flagella and fimbriae. Flagella are long rotating appendages attached to the IM and used for locomotion. Fimbriae or pili are structurally similar to flagella. Fimbriae are numerous thin protein tubes mostly involved in bacterial adhesion (Li and Mobley, 2001). Some bacteria produce a special pilus called a conjugation or sex pilus that permits conjugation, i.e. the transfer of DNA from a donor to a recipient bacterium to enable genetic

recombination. Pili are often visualized under the electron microscope because they serve as bacteriophage receptors and are coated with them (Madigan *et al.,* 1997). Phages use the proprieties of polymerization/depolymerisation of pili to approach the bacterial surface.

I.5. PHAGE GENOME DELIVERY TO BACTERIAL CELLS

Like any other virus, tailed phages must deliver their genome to the site of replication inside the host, in this case the bacterial cell cytoplasm. However, in contrast to the majority of eukaryotic viruses, phage virions do not enter the host cell. The only component that reaches the cytoplasm is the viral nucleic acid and eventually a few structural proteins that, among other functions, may assist DNA entry. As mentioned above, the phage tail is the specialized structure that mediates the attachment of virions to the bacterial surface and that serves as conduit for the viral DNA transfer into the bacteria.

In spite of several decades devoted to the study of the various stages of the phage infection cycle, the phenomenon of phage DNA entry to bacterial cells remains poorly characterized, still being a mystery how a dsDNA molecule of several megadaltons crosses the bacterial cell wall and membranes in a few seconds or minutes. Although important work has been carried out with a few *Escherichia coli* phages, the study of this process in phages infecting Gram-positive bacteria has remained elusive until recently.

The following sections summarize the current fundamental aspects related to tailed phage binding to host cells and ejection of the viral DNA into bacteria.

I.5.1 – PHAGE ADSORPTION

The first vital step for bacteriophage propagation is the recognition and binding to a susceptible host. This process is classically denominated as adsorption and in principle guarantees that phages deliver their genetic information to an active and fit host cytoplasm. Phage adsorption has been a study subject since the discovery of bacteriophages in the early 20th century and is still a delicate issue nowadays. There are multiple solutions for phages to adsorb to their hosts but at least one structural phage protein (anti-receptor or phage receptor-binding proteins (RBP)) has to be in contact with, at least, one bacterial surface element (receptors). Early studies have readily highlighted that phage adsorption could involve both a reversible interaction with bacterial receptors and an irreversible binding leading to phage inactivation. Reversible adsorption corresponds to binding to a non-saturable cell wall receptor in a way that phages can be released from cells as infectious particles by dilution. Irreversible adsorption is generally characterized by binding to a specific saturable receptor and phages are not recoverable as they become

committed to infection. Reversible adsorption has been interpreted as a dynamic process that facilitates irreversible binding either as a transitory step that precedes phage locking to the bacterial surface or as a pre-requisite required by phages to be structurally able to adsorb irreversibly. A more elaborate discussion of the phage adsorption process is presented in Part II.

I.5.1.1 – Bacterial receptors

Potentially all structures presented at the bacterial cell surface can work as phage receptors. The two types of adsorption can take place with the same receptor but, more frequently distinct surface components are involved, obeying the rule non-saturable *vs* saturable receptor. The recent review of Vinga *et al.* (2006a) compiles a list of known receptors.

Gram-negative hosts

Surface structures such as the conjugative pili and the flagella are often used by bacteriophages to adsorb to their hosts (this is especially the case of filamentous, ssRNA and dsRNA phages; see Vinga *et al.,* 2006a for a review). Tailed phages attaching to Gram-negative hosts seem though to preferentially use the LPS as the first contact structure. LPS serve mainly as reversible receptors, with the studied exception being phage T7 that seems to target only the LPS for both steps of adsorption. Usually LPS are used in conjunction with one outer membrane protein (OMP) for the complete adsorption process. OMPs can be porins (e.g., OmpC, OmpF, PhoE), specific transport proteins (e.g., LamB, BtuB, Bfe, FhuA, SidK, Tsx; Heller, 1992) or have other roles such as cell shape stabilization (e.g., OmpA). Examples of phages that target these two types of receptors are phages T4, K20, T5, and TLS, which, in addition to LPS, also interact with the non-specific porins OmpC, and OmpF, FhuA, and the multipurpose OMP TolC, respectively (Yu and Mizushima, 1982; Traurig and Misra, 1999; Heller and Broun, 1982; German and Misra, 2001).

Interestingly, a report from Williams *et al.,* 1986 stated that, apart from the well described adsorption of the *E. coli* Lambda (λ) to the maltoporin LamB, the inner membrane protein ManY, was also involved in the initial steps of phage infection. ManY is involved in mannose uptake through a specific II-Man phosphotransferase system (PTS). In fact, these authors demonstrated that *E. coli pel* mutants are affected in ManY. These mutants had been previously described as allowing λ adsorption but not DNA entry. Recently, Moldovan *et al.,* 2007 showed that phage λ targets LamB for reversible and irreversible interactions and described the adsorption kinetics for each event. So, infection dependence on IM protein ManY seems to be

strictly at the DNA passage level, possibly showing the way to the DNA crossing channel constitution. Also, irreversible adsorption and infection of coliphage N4 is dependent on two proteins, an OMP (NfrA) and an IM protein (NfrB), respectively, being the reversible adsorption possibly to the LPS (MacPartland and Rothman-Denes, 2009). Uniquely, a report by Kiino *et al.* (1993) refers the requirement of a cytoplasmic protein, NfrC, for N4 irreversible adsorption. It was suggested that NfrC should be important for the correct function of the cell envelope receptors, either at the adsorption step or during the translocation of both viral genome and virion RNA polymerase into the cell interior. FhuA is an *E. coli* OMP that acts as the receptor for ferrichrome-iron and, together with the energy transducing periplasmic protein TonB mediates the active transport of ferric siderophores across the outer membrane (Ferguson *et al.,* 1998). These two proteins act as receptors for phage T1 (Letellier *et al.,* 2004). Phage T5 uses FhuA for the irreversible adsorption after binding reversibly to the LPS (Heller and Braun, 1979; Heller and Schwarz, 1985). *Salmonella* phages Gifsy-1 and -2 seem to target only the porin OmpC (Ho and Slauch, 2001), as well as the enterophages T-even that seem to adsorb only to the nonspecific diffusion-channel forming OmpA (Morona *et al.,* 1985).

Also important is the idea that polysaccharide capsule produced by some Gram-negative bacteria can act either as a barrier or as a receptor for phage adsorption (Scholl *et al.,* 2005).

Gram-positive hosts

Likewise, in Gram-positive hosts the most used targets for bacteriophage adsorption are the teichoic acids (WTA and LTA) and the exposed polysaccharides of the cell wall. The *Lactococcus lactis* exopolysaccharides (EPS) are heteropolysaccharides composed of repeating units of sugars in which galactose, glucose, and rhamnose are the most common carbohydrates. These components are used by phages like c2, bIL67 and several members of the 936-family i.e., ϕ67 and bIL10, to adsorb to their host (Monteville *et al.,* 1994; Dupont *et al.,* 2004; Deveau *et al.,* 2002). Phage PL-1, which infects *Lactobacillus casei*, also recognizes the L-rhamnose residues of the EPS (Watanabe and Takesue, 1975). Phages that adsorb to *Streptococcus thermophilus* do it by interacting with peptidoglycan-associated sugars such as glucosamine, rhamnose or ribose (Quiberoni *et al.,* 2000). Recently, Rodríguez *et al.* (2008) showed the importance of EPS and capsular polysaccharides (CPS) of *S. thermophilus* in the adsorption of phages.

The sugar side chain substituents of the WTA and the LTA are also major targets for bacteriophage adsorption. This is the case of most *B. subtilis* phages, i.e. ϕ29, ϕ25, SP01, SP10

(Young, 1967; Yasbin *et al.,* 1976) and of the *Lactobacillus delbrueckii* phage LL-H (Räisänen *et al.,* 2007) that recognize and adsorb to the WTA and LTA associated glucose, respectively. Adsorption of a few *Listeria monocytogenes* phages was also characterized, and these phages also target the TA sugar substituents such as glucose, for phages A500 and A511 (Cheng *et al.,* 2008), or the N-Acetylglucosamine and rhamnose residues, for phage A118 (Wendlinger *et al.,* 1996). Interestingly, this report also pointed the peptidoglycan itself as part of the receptor for *Listeria* phage A511. This kind of involvement of the peptidoglycan in phage adsorption was also reported for *Staphylococcus aureus*, in addition to the receptor activity of glucosylated TA (Coyette and Ghuysen, 1968). Although in *B. subtilis* most of the receptor activity is concentrated in the major WTA (poly-[glicerol-phosphate]), the minor WTA (poly-[glucosyl-N-acetylgalactosamine 1-phosphate]) was shown to act as the receptor for phage ϕ3T and other *B. subtilis* related phages (Estrela *et al.,* 1991).

Tailed phages can also use the flagellum to adsorb to their motile Gram-positive hosts. As an example, the *B. subtilis* phage PBS1, a *Myoviridae*, uses the filament of this structure to bind reversibly to the host, and the attachment becomes irreversible only after the phage has travelled along the filament to the basal body where it becomes irreversibly adsorbed (Lindeberg, 1973).

Unlike Gram-negative bacteria, involvement of proteins in adsorption to Gram-positive hosts was only recently characterized, although considered for a long time. This includes the *L. lactis* Pip (phage infection protein), the *B. subtilis* YueB and the *B. anthracis* GamR (gamma phage receptor) (Geller *et al.,* 1993; São-José *et al.,* 2004; Davison *et al.,* 2005), which act as receptors for phages c2, SPP1 and γ, respectively. Proteins Pip and YueB belong to the family of YueB-related membrane proteins that are characterized by the presence of five carboxyl-terminal transmembrane anchoring domains and a large elongated extracellular domain that faces the cell wall (São-José *et al.,* 2004). A more refined characterization of YueB was performed and the authors described that the ectodomain is an elongated dimer that forms a 36.5 nm long fiber that is exposed in the *B. subtilis* surface (São-José *et al.,* 2006). YueB differs structurally from protein receptors found in Gram-negative bacteria, making it a highly suitable receptor for phages infecting Gram-positive hosts. GamR is a cell wall-anchored protein that possesses the LPXTG motif at the C-terminus. Pip and YueB act as irreversible receptors while GamR was proposed by the authors to be strictly involved in the reversible step of adsorption.

Inclusion of proteins as receptors in Gram-positive hosts is for sure not so limited. A study performed with *B. subtilis* phage ϕ29 revealed that, although mutations in the glucosylation pathway of teichoic acids are sufficient to confer resistance to this phage, L-forms of *B. subtilis*, which are devoid of cell walls, are sensitive for ϕ29 infection, indicating a further step of adsorption and recognition at the cytoplasm membrane level (Jacobson and Landman, 1975; 1977).

It is also important to stress that when just one cell receptor is involved in phage adsorption, either to Gram-negative or -positive hosts, it is the phage itself that changes or rearranges structurally to accommodate the typical two step binding. This is especially true for phages with tail fibers and the best studied case is the coliphage T4 (see below).

I.5.1.2 – Phage receptor binding proteins

Phages too offer multiple structural proteins that can act as the receptor-binding proteins (RBP). Adsorption of myoviruses to their bacterial hosts is normally mediated by the long and short tail fibers attached to the contractile tail, whereas podoviruses bind to host cells usually through the stubby tail spikes attached to the short noncontractile tail. Siphoviruses use baseplates, straight tail fibers or tail spikes for host receptor recognition. A feature frequently included in tail spikes is the ability to, besides recognition, cleave the host cell receptor. This is believed to happen so that (i) it contributes to the mobility of phage particles on the bacterial surface, (ii) the phage can gain access to the receptor structure located deeper in the cell wall and (iii) phages can escape bacterial debris after cell lysis (Baxa *et al.,* 1999; Steinbacher *et al.,* 1996; Walter *et al.,* 2008). Tail spike enzymes degrading cell surface polysaccharides have long been known (Rieger-Hug and Stirm, 1981), the ones from podoviruses P22, Sf6 and HK620 being the best characterized (see I.5.1.2.1; Lieman and Molineux, 2008).

Phages of Gram-negative hosts

Phage λ uses the tip (C-terminal extremity) of its straight tail fiber, composed of protein gpJ, to bind to LamB (Wang *et al.,* 2000) and P22 employs the tail spike (gp9) to recognize and degrade the O-antigen of the LPS (Iwashita and Kanegasaki, 1976). When there are distal fibers, these are frequently used to sense the bacterial cell surface and interact in a reversible way with the receptors. Well studied examples of this case are the L-shaped fibers of phage T5, the tip of the T4 long tail fibers (gp37) and the T7 tail fibers, which interact in a reversible manner with the

LPS (Heller and Braun, 1982; Cerritelli *et al.,* 1996; Molineux, 2001). The later example also uses the same structures for irreversible adsorption. The protein pb5 of phage T5, located in the tip of the tail conical region, mediates irreversible binding to FhuA (Heller and Schwarz, 1985).

The mechanism of adsorption, the reversible and the irreversible step, is best studied in coliphage T4. The tip of the long tail fibers (LTF) is composed of the C-terminal regions of protein gp37 and, as stated above, the LTF tip interacts reversibly with the LPS on the host surface. This extremity can also recognize the porin OmpC although no further information on the subsequent steps is available. In the case of LPS receptors, after at least three LTFs are bound, the baseplate is brought closer to the cell surface and a signal is transmitted to the baseplate forcing the folded short tail fibers (STF) to unfold and reach the LPS. These STFs are composed of gp12 and are responsible for irreversible adsorption to LPS. The release of the STF from beneath the baseplate causes the baseplate to change from a hexagon shape to a star-like structure, having conformational implications on the tail structure. The tail sheath that is in direct contact with the baseplate expands, and this expansion is propagated throughout the length of the tail, leading to tail sheath contraction (~37% of the original size). The tail sheath contraction is accompanied by a rotation movement that culminates in the phage head, causing it to rotate almost a full turn relatively to the tail axis. The baseplate hub (gp29, gp27, gp5), which is connected to the capsid via the tail internal tube, protrudes from the tail as a consequence of the capsid rotation and tail sheath contraction. gp5, the needle-like cell-puncturing device (see below) is believed to penetrate the cell by combining pressure with rotation (Kostyuchenko *et al.,* 2005).

Phages from Gram-positive hosts

Work on the host-range determinants of *L. lactis* phages allowed the identification of virion tail proteins implicated in host recognition. These are frequently part of the baseplate structure in the 936- and P335-groups and of the tail and tail tip in the c2-group. The receptor-binding protein of phage p2, belonging to the 936-group, is the baseplate protein Orf18 and was the first RBP from a phage infecting a Gram-positive bacterium to be crystallized (Spinelli *et al.,* 2006). In other phages of the same group the RBPs have been identified by generation of chimeric phages with altered host range and adsorption specificity. This is the case of the RBPs Orf18 and Orf20 of sk1 and bIL170, respectively, which were shown to locate in the tip of the tail (Dupont *et al.,* 2004). Similar studies performed with phages from the c2- and P335-groups also identified tail proteins involved in host recognition. The Orfs 35 and 2 of the c2-group phages bIL67 and CHL92, respectively, where shown to be responsible for host determination at the step of

adsorption (Stuer-Lauridsen *et al.,* 2003). The equivalent RBP in phages c2 and 923 is L15 (Rakonjac *et al.,* 2005). Interestingly, a protein (L10) located at the tip of the tail of phage c2 that was initially suggested to be involved in host recognition (Lubbers *et al.,* 1995), is in fact more likely to play a role in phage DNA entry into the host cells (Rakonjac *et al.,* 2005). In the P335-group the double-disk baseplates of phages TP901-1 and Tuc2009 have been extensively studied. The RBPs are Orf49 in TP901-1 and Orf53 in Tuc2009, generally named BppL (lower baseplate protein). They are structurally organized as homotrimers that exist in 6 copies in the virion (Sciara *et al.,* 2008; Spinelli *et al.,* 2006) and have several homologous proteins in *L. lactis* phages (Vegge *et al.,* 2005).

Duplessis and Moineau (2001) identified Orf18 as the determinant of host specificity for the *S. thermophilus* phages DT1 and MD4 (and related phages by homology studies), also by the generation of chimeric phages with an altered host range. More recently, Duplessis *et al.* (2006) extended the involvement of two other virion proteins in phage-host interactions, specifically Orf15, which the authors speculate to be the tape measure protein, and Orf17, a putative tail protein.

In myoviruses infecting Gram-positive bacteria the participation of tail fibers in receptor recognition has not yet been reported, although bioinformatic analysis suggests these structures to be involved in the process (Uchiyama *et al.,* 2008; Klumpp *et al.,* 2008). In *Podoviridae* phage ϕ29, the neck appendages (gp12) of the virion are responsible for the reversible interaction with the *B. subtilis* cell wall (Villanueva and Salas, 1981; Guo *et al.,* 2003). Recently gp12 was also proposed to function as cell-wall degrading enzyme by Di Mauro *et al.* (2007) as it contains an N-terminal glycosidase domain.

I.5.1.2.1 – Common structures of phage RBPs

Most of the already determined RBP structures are elongated homotrimers with a fibrous morphology and β-sheet topologies with unusual repetitive folds, including the irregular triple β-helix, triple β-spiral and the unique fold of T4 short tail fibers (Mitraki *et al.,* 2006). Beta-helices are associations of β-strands arranged in a helical way with either two or three faces and β-spirals are α-helical coiled coil like structures composed of β-strands (see Mitraki *et al.,* 2002 for a review on these structures). Presently, triple β-spirals are known to exist in the human adenovirus type 2

(Ad2) adhesion protein and in the reovirus σ1 attachment protein and will not be discussed here (Weigele *et al.,* 2003; van Raaij *et al.,* 1999; Chappell *et al.,* 2002).

These proteins exhibit relatively high melting temperature, protease resistant domains, and stability in the presence of detergent at room temperature (Barbirz *et al.*, 2009). These features may reflect their function as external virion proteins that must remain intact in an extracellular environment in order to keep the phage infectious (Weigele *et al.,* 2003; Mitraki *et al.,* 2006). Their structure also catapults them for the specialized functions involved in virus attachment and infection as they can be multi-purpose proteins with an extended active site used to recognize, bind and cleave specific sequences on cell surface polysaccharides (Bradley *et al.,* 2001). The assessment of putative similar function proteins in other phages arrives from structural homologies.

RBPs from lactococcal phages share an overall structure comprised of three domains: the N-terminus or shoulders, the neck and the head (Fig. I.4A). The N-terminus is an α-helix, which may or may not be attached to a β-sandwich fold, and connects with the neck. The neck is a tight β-prism (formed by β-helices) and connects to the β-barrel head, constituted of a trimeric arrangement of double Greek key domains (Spinelli *et al.,* 2006a; Ricagno *et al.,* 2006; Sciara *et al.,* 2008). The head domain carries the saccharide binding site, and it was experimentally demonstrated that it binds glycerol molecules in a crevice between the monomers (Spinelli *et al.,* 2006b; Ricagno *et al.,* 2006). Amongst the phages that possess this type of RBP, the ones from phage sk1 and p2 share 97% of identity, not at all unexpected as they infect the same *L. lactis* strains. Between p2 and bIL170, although the shoulder domains share ~90% identity, this value resides in the 15% for the neck and head domains (Spinelli *et al.,* 2006a). Tremblay *et al.,* (2006) report a phylogenetic tree of the receptor-binding proteins of 936-like phages. The head domains of p2 and TP901-1 share ~28% identity while their N-terminal domains share no identity at all as the primary sequence of the TP901-1 RBP is significantly shorter (Spinelli *et al.,* 2006b). No high resolution structure of phage Tuc2009 baseplate is yet available, but a model of its organization was recently presented by Sciara *et al.* (2008). This phage shares ~78-97% amino-acid identity with TP901-1 in the individual protein components that make up the baseplate, but the BppL (RBP) N-terminal domains only have ~64% identity and surprisingly, the C-terminal domains share no identity.

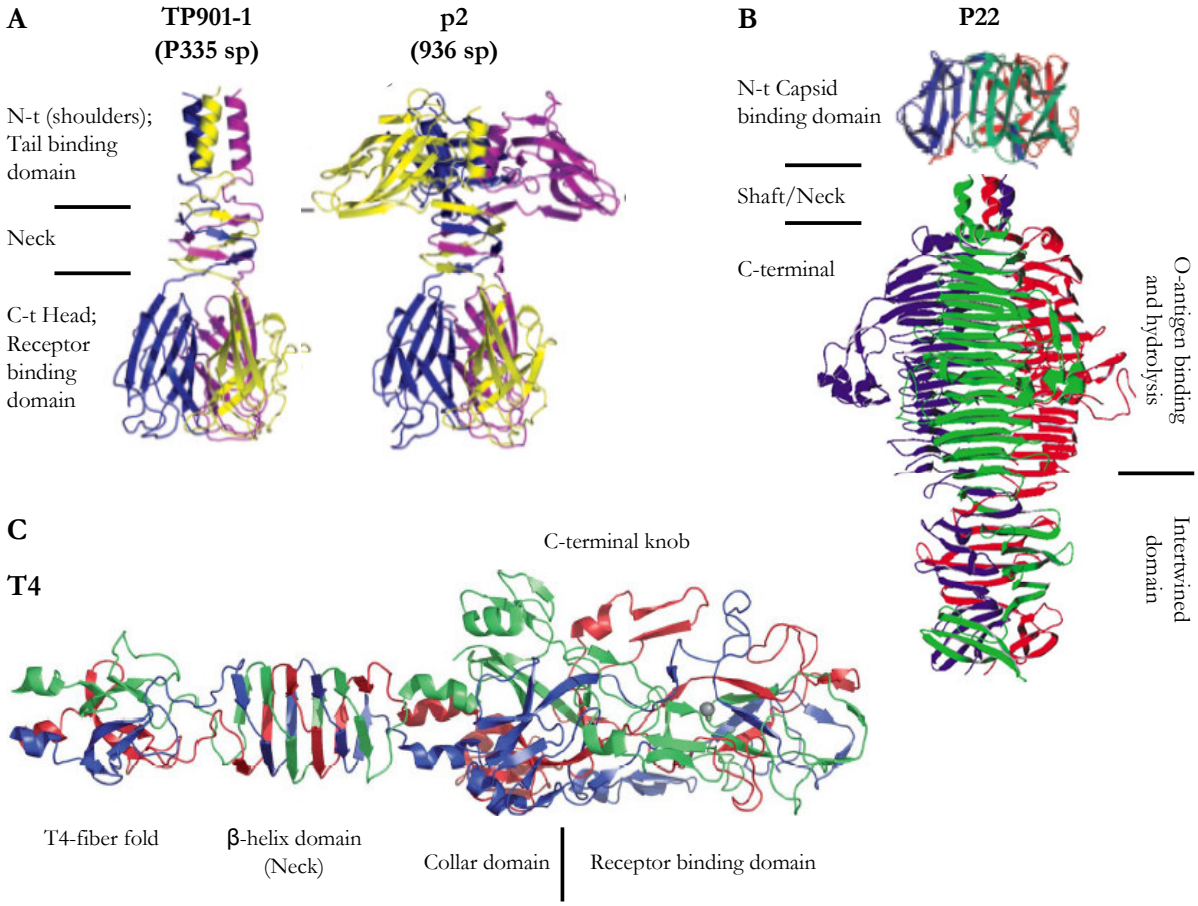


Figure I.4 | Structure comparison of three different types of RBPs featuring unusual β-folds found in phages. (**A**) The lactococcal phages RBPs are composed of an amino-terminal tail/baseplate binding domain, mainly α-helix, that can have a shoulder-like structure as in phage p2. The neck has the unusual β-helices forming a β-prism structure and the carboxyl-terminal β-barrel head comprises the receptor binding domain. Individual subunits are colored blue, yellow and purple. (**B**) The asymmetric tail spike protein of phage P22 (gp9; TSP). The N-terminal capsid binding domain (β-sheets) connects to the C-terminal half by the shaft (or neck). The two half's of the protein were crystallized separately due to the flexible nature of the α-helix shaft. The receptor binding and hydrolysis domain is a large β-helix that connects to a β-prism intertwined domain that stabilizes the entire TSP. (**C**) gp12 from phage T4 forms the short tail fiber, a flexible protein that is bend in the virion and unbends after reversible adsorption. The C-terminal half of the protein is shown (residues 246-527). The unusual T4-fiber fold is a β-spiral/β-helix like domain that connects to a β-helix neck. The C-terminal knob is a more globular head that connects to the neck via a short α-helical triple coiled coil. The collar domain that has β-strands and α-helix is followed by the receptor binding domain. It has a heavily intertwined, "knitted," fold, with a central zinc ion (shown as a gray sphere). In (**B**) and (**C**) individual subunits are colored red, blue and green. (**A**) was reproduced from Ricagno *et al.,* 2006; (**B**) was adapted from Steinbacher *et al.,* 1997 and Weigele *et al.,* 2005; (**C**) was reproduced from Mitraki *et al.,* 2006. Proteins are shown in different scales.

van Raaij *et al.* (2001a and b) solved the structure of the T4 virion protein that binds irreversibly to the LPS. The C-terminal half of bacteriophage T4 STF (gp12) contains an elongated shaft having a triple β-spiral like fold, a middle triple β-helix domain (similar to the neck, in lactococcal RBPs), and a globular C-terminal knob that binds the receptor, composed largely of β-sheets (β-sandwich). The N-terminal shaft (of the C-ter half) folds into an intertwined β-strand region

different from the β-spiral and previously described β-helices, and therefore is considered a different kind of unusual β-folds (Fig. I.4C; Mitraki *et al.,* 2006). Other T4 virion proteins are known to possess these irregular β-folds, an example is the lysozyme gp5. This protein can be divided also in three domains: a N-terminal oligosaccharide binding(OB)-fold domain, a middle lysozyme domain and a C-terminal triple β-helix. An interesting feature is that this protein undergoes a maturation cleavage between the lysozyme and the C-terminal domain but both the resultant parts, the N-terminal (gp5*) and the C-terminal (gp5C), remain associated and are found in the phage particle (Leiman *et al.,* 2003).

Tail spike proteins (TSPs) from podoviruses P22, Sf6, HK620, SP6 and myovirus Det7 are adhesion fibers essential for the phage to recognize its carbohydrate receptor during infection. These TSPs are asymmetric homotrimeric proteins consisting of an N-terminal virus-binding head, a central shaft or stalk domain and a C-terminal globular receptor-binding domain (Fig. I.4B; Mitraki *et al.,* 2006). The N-terminal trimeric dome-like structure is formed by β-sheets and caps a three-helix bundle observed in the structure of the C-terminal receptor binding and cleaving fragment (Steinbacher *et al.,* 1994; 1997). This three-helix bundle confers flexibility to the structure, probably to allow it to bend during the infection process. The receptor binding and cleaving domain is composed of 13-14 rungs of parallel β-helix domain, and ends in an intertwined β-prism domain where the three chains wrap around each other, which is vital for TSP trimerization and thermostability (Steinbacher *et al.,* 1997; Weigele *et al.,* 2005). The RBPs from *Salmonella* phages Det7 (Walter *et al.,* 2008) and SP6 (Scholl *et al.,* 2004) share a certain level of similarity with that of P22. Although the β-helix of Det7 and P22 are very similar (they share ~75% identity in the β-prism and their catalytic residues are almost 100% conserved), they share no significant sequence homology between their amino-terminal parts (Walter *et al.,* 2008). SP6 does not have a P22-like head binding domain either. On the contrary, the TSP from P22, Sf6 (*Shigella flexneri* phage; Chua *et al.,* 1999; Müller *et al.,* 2008) and HK620 (*E. coli* H phage; Barbirz *et al.,* 2008) have ~70% similarity in the N-terminal part, but no amino-acid sequence similarity in the carboxyl-end. Instead, the extreme C-terminal domain of HK620 tail spike forms a β-sandwich, as in Sf6 and unlike in the P22 tail spike (Barbirz *et al.,* 2008). Despite the different folds, the authors propose that the TSP genes of P22, Sf6 and HK620 have a common ancestor, based on the C-terminal motifs conserved between the three proteins revealed in sequence alignments. A summary covering the topic of how phages have evolved to adapt their tail spikes to the substrate diversity is well presented by Leiman and Molineux (2008). It is very interesting

to see that a secondary protein structure, obscure to biologists just a few years ago, is so widespread within the virus world.

I.5.2 – PHAGE DNA EJECTION AND TRANSFER TO HOST CELLS

Phage adsorption to the host cell surface may be considered as the first step of a process that culminates in the entry of the viral nucleic acid into the bacterial cytoplasm. The landmark experiment of Hershey and Chase (1952) with phage T2 led to the popular idea that phages inject their DNA via a syringe-like mechanism. However, over the years it became clear that the syringe model was insufficient to explain the rather complex phenomenon of phage genome entry (Grayson and Molineux, 2007). Although particular points of the process have been fairly characterized in some systems, globally it remains poorly studied for the vast majority of phages. The picture that emerges from the reported studies is that most probably each phage has its specific mechanism of genome transfer to host cells. However, one can elaborate a set of common events that govern tailed phages DNA translocation (reviewed in Vinga *et al.,* 2006a; Molineux, 2006; Grayson and Molineux, 2007) as follows: i) triggering of a DNA ejection signal resulting from the interaction of tail RBPs with bacterial receptors; ii) transmission of this signal through conformational changes of the virion structure; iii) creation of a pathway across the bacterial cell envelope to allow DNA passage; iv) opening of the capsid connector and action of forces that push out the DNA; and v) action of forces in the bacterial cytoplasm that actively pull the phage DNA inside.

I.5.2.1 – Virion rearrangements and crossing of cell barriers

The free energy released upon phage irreversible binding most likely induces tail structural rearrangements (Molineux, 2006). These are important to transmit the DNA ejection signal to the phage head (Plisson *et al.,* 2007; Boulanger *et al.,* 2008; Aksyuk *et al.*, 2009) and/or to allow the contact and penetration of virion components across the cell barriers (Kanamaru *et al.,* 2002; Leiman *et al.,* 2004; Kemp *et al.,* 2005; Boulanger *et al.,* 2008). These components may be part of the tail or be carried in the capsid, in the latter case being ejected before or along with the DNA (see below). It is not fully understood how phage proteins penetrate the host cell envelope, but it is clear that phage tails are too wide to penetrate the pore(s) of porins that are recognized as receptors, such as in the case of λ and T5. It seems therefore that phage structures should be able

to span the bacterial barriers (membranes and cell wall), which means that virion proteins must insert in the membranes and perform a local digestion of the cell wall.

Examples of such virion proteins, displaying one or both of the functions referred to above have been described for several phages. The tail protein pb2, a 121 kDa multimeric protein that forms the phage T5 straight fiber, was proposed to function as a multifunctional protein acting as a tape measure (N-terminus) and carrying membrane fusion and muralytic activities in its C-terminus (Boulanger *et al.,* 2008). It was also proposed that the pb2 N-terminal coiled coil region would serve as a sensor for triggering the opening of the head-tail connector. The C-terminus region of the tape measure protein from the mycobacteriophage TM4 also has peptidoglycan-hydrolysing activity (Piuri and Hatfull, 2006). While this activity is not required for phage viability, it facilitates efficient infection and DNA ejection into stationary phase cells. Phage T7 genome ejection into stationary phase *E. coli* cells seems also to be facilitated by the lytic transglycosylase activity of gp16, which is one of the components of the protein core carried in the capsid that is ejected before the phage DNA (Moak and Molineux, 2000).

The commitment of phage T4 to infection triggers major conformational changes of its tail (Leiman *et al.,* 2004; Rossmann *et al.,* 2004). One that is relevant for this discussion is the contraction of the tail sheath that causes the tail internal tube, carrying the baseplate hub that remains attached to its end, to protrude and penetrate the cell envelope. This hub terminates as a pointed needle, which if formed by the triple-stranded β-helical domain of gp5 C-terminus. The β-helical domain dissociates when it comes into contact with the periplasmic peptidoglycan layer, activating the lysozyme domains that are also carried in the gp5 trimer. Digestion of the peptidoglycan layer creates an opening through which the tail tube can reach the cytoplasmic membrane of the host cell.

For the T7 family of short-tailed phages, and likely many *Podoviridae* infecting Gram-negative bacteria, the hole in the peptidoglycan allows the assembly of a conduit, or extensible tail, formed by proteins ejected from the virion since the tail is too short to reach the cytoplasmic membrane (Molineux, 2006). The tail-spikes of some T7-like phages can even display glycosidase activities, allowing these phages to digest host capsules and reach the cell surface (Leiman *et al.,* 2007). As would be expected, virion proteins bearing muralytic activity were also found in many phages infecting Gram-positive hosts (Moak and Molinex, 2004; Kenny *et al.,* 2004).

Ejection of the tape measure from the lumen of the tail provides an obvious signal for genome ejection, simultaneously clearing the path for DNA passage. The tape measure protein of λ, gpH* (processed from gpH during morphogenesis), may have yet a third role in actually forming the channel across the cytoplasmic membrane (Roessner *et al.,* 1984). Additionally, phage λ DNA penetration seems to be facilitated by the interaction between gpH* and the cytoplasmic membrane component ManY of the mannose PTS system (Esquinas-Rychen and Erni, 2001 and references therein).

Quite interestingly, by monitoring by light microscopy the infection of phages tagged with quantum dots, Edgar *et al.* (2008) have shown that a number of phages preferentially target the bacterial poles for entry into host cells, particularly at low multiplicities of infection. In the case of phage λ, the authors found that ManY (see above) localized to the bacterial poles as well and labelling of λ DNA during infection revealed that it is injected and replicated at the polar region. The authors speculated that the polar region might be a preferred place for DNA injection to take place, as uptake of DNA by naturally competent bacteria also occurs at the pole.

I.5.2.2 – Phage DNA transfer

Opening of the capsid connector must be followed by the action of forces that expel the packaged DNA. The genome of most dsDNA bacteriophages is densely packed in the capsid (about 500 mg/mL; Earnshaw and Casjens, 1980). As result phage heads are under high pressure because of the strong repulsive forces between phosphate groups of closely packed double helices and of the bending stress that arises from confinement of DNA in the capsid (Tzlil *et al.,* 2003; Fuller *et al.*, 2007). Theory (Purohit *et al.,* 2005) and recent experiments (Evilevitch *et al.,* 2003; Grayson *et al.,* 2006; São-José *et al.,* 2007) support the hypothesis that the capsid pressure (tens of atmospheres corresponding to forces on a pN scale) should work as genome ejection force. However, this force has been shown to decrease monotonously as the DNA exits de capsid and, at best, it should only drive DNA ejection until the moment when the host cell turgor pressure balances that of the phage head. Complete DNA entry must therefore depend on other force sources.

Complete phage T7 DNA entry to *E. coli* cells has been shown to be intimately coupled to its transcription, first by the host RNA polymerase and then by the phage-encoded RNA polymerase (Molineux *et al.,* 2001). Phage N4 encodes a RNA polymerase (vRNAP) that is

packaged in mature virions. At the onset of infection, this enzyme is ejected into the host cell, presumably preceding the first ~500 bp of genomic DNA. Activity of vRNAP on this DNA extremity then pulls the next 10 to 40 kbp of the genome into the host. Transport of the remaining DNA requires RNA synthesis catalyzed by the N4-encoded RNA polymerase II, a product of early transcription (Choi *et al.,* 2008). It was proposed that phages T5 and ϕ29 internalize an initial part of their genomes through the capsid pressure (de Frutos *et al.,* 2005a; González-Huici *et al.,* 2006). Phage proteins synthesized from this initial portion were shown to be required for transfer of the remaining DNA (Lanni *et al.,* 1969; González-Huici *et al.,* 2004). Additional, general forces are likely generated by the condensation of the ejected DNA in the bacterial cytoplasm and/or by the ratcheting action of DNA binding proteins (de Frutos *et al.,* 2005b; Inamdar *et al.,* 2006; Jeembaeva *et al.,* 2008). Finally, it was recently proposed that the osmotic pressure difference between the cytoplasm and external fluids might provide a force for DNA ejection for many phages. According to this theory, it is the stream of water molecules moving along the osmotic gradient from the medium, through the phage virion, and into the cell, that drags the phage genome along with it (Molineux, 2006; Grayson and Molineux, 2007).

I.6. THE *BACILLUS SUBTILIS*/SPP1 SYSTEM

B. subtilis is a low G+C Gram-positive bacterium belonging to the *Firmicutes* phylum. It is commonly detected on the soil and has the ability to form spores. *B. subtilis subsp. subtilis* strain 168 has a genome of 4,214,810 base pairs with at least ten prophages (Kunst *et al.,* 1997, NC_000964) and has a considerable high genetic diversity (Earl *et al.,* 2007). It is one of the most studied bacteria and has been exhaustively used as a model for Gram-positive sporulating bacteria in various fields of Microbiology and Molecular Biology. Bacteriophages infecting *B. subtilis* are long known and have also been used as models for phages infecting Gram-positive bacteria. These include the podovirus ϕ29, the myovirus SP01 and the siphovirus SPP1. Study of phage adsorption has greatly contributed to understand the cell envelope of their hosts, and this is also true for *B. subtilis*. For example, pioneering studies involving the capacity of phage SP50 to adsorb to *B. subtilis* cell surface, contributed to understand the dynamics of teichoic/teichuronic acids replacement under phosphate limited conditions (Ellwood and Tempest, 1969; Archibald and Coapes, 1976). Important details on the unsettled topic of peptidoglycan biosynthesis and architecture have come to light recently with the work of Hayhurst *et al.,* 2008 on *B. subtilis*. They propose an innovate model of a coiled-coil shaped peptidoglycan architecture. It is for long known that teichoic acids in *B. subtilis* 168 are decorated with glucosyl and alanyl residues and that the former plays an important role in bacteriophage adsorption (Young, 1967). The implication of this sugar in the adsorption of phage SPP1 was first reported by Yasbin *et al.* (1976).

The *B. subtilis* phage SPP1 was first described by Riva *et al.* (1968) and has a 44,007 pb dsDNA genome (Alonso *et al.*, 1997) enclosed in an icosahedral capsid with a diameter of ~66 nm (Fig. I.5). The capsid is attached to a flexible non-contractile tail with ~193 nm that comprises a ~30 nm long tail fiber (Plisson *et al.,* 2007). It belongs to the lambda-like genus of the *Siphoviridae* family. The connector localized at one vertex is composed of the portal protein gp6, and of gp15 and gp16, which closes the channel for DNA exit (Orlova *et al.,* 2003). gp6 was the first protein to be described that associates to a 13-mer when isolated but acquires a typical 12-mer oligomerization in the virion connector (Lurz *et al.,* 2001). The nonessential minor head protein gp7 was found to be necessary for the correct routing of the virus genome to the host cytoplasm while remaining associated to the connector during ejection (Vinga *et al.,* 2006b). The product of gene *17* brings the head and tail together (Plisson *et al.,* 2007). The major head protein is gp13 and the mature capsid is decorated with gp12. The external helical tail tube, composed of gp17.1

and 17.1*, is arranged around an inner tube of the tape measure protein gp18 (Plisson *et al.,* 2007). The tail fiber will be discussed in detail in Part IV.

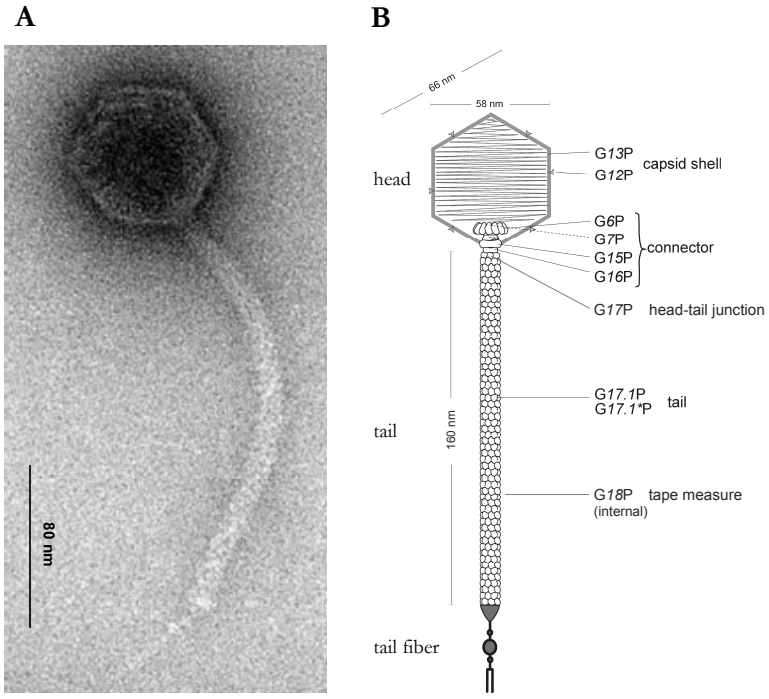


Figure I.5 | The SPP1 virion. (**A**) Electron microscopy and (**B**) scheme showing dimensions and structural proteins. TEM photograph by Rudi Lurz, Max-Planck-Institut für Molekulare Genetik, Berlin, Germany (magnification 67,000 x).

Phage SPP1 targets the *B. subtilis* membrane protein YueB for irreversible adsorption (São-José *et al.,* 2004). Under appropriate conditions, the interaction with the receptor YueB triggers SPP1 DNA ejection (São-José *et al.,* 2006; Raspaud *et al.,* 2007). The DNA ejection signal was suggested to be propagated from the tail tip to the capsid connector through structural rearrangements of the proteins composing the tail helical tube (Plisson *et al.,* 2007). This leads to opening of the connector and consequent release of DNA, which enters the host cell cytoplasm by a yet unknown mechanism. Entry of the first 17% of the SPP1 genome is believed to be powered by the pressure built in the virion capsid (São-José *et al.,* 2007). Once inside the cytoplasm DNA transcription is carried out by the *B. subtilis* RNA polymerase and host RNA synthesis continues throughout infection (Milanesi and Cassani, 1972; Esche, 1975; Esche *et al.,* 1975). Synthesis of SPP1 DNA is by theta and rolling circle replication and takes place in close association with the membrane, forming concatamers (Burger, 1980; Alonso *et al.,* 2006). Packaging is initiated by the binding of the small subunit (gp1) of the gp1-gp2 terminase complex, which is associated with the portal protein, to one of the *pac* sites of the DNA concatamer. SPP1 encapsidates its DNA by

a headful mechanism where the gp2 (ATPase) subunit terminates the packaging by an endonucleolytic cleavage. This efficient headful mechanism leads to the generation of terminally redundant and partially circularly permuted DNA molecules, meaning that the length of the packaged DNA is longer than one genome, approximately 44.9 kb (Raspaud *et al.*, 2007; Dröge and Tavares, 2000; Oliveira *et al.,* 2006 and references therein). Virions are released by phage mediated host lysis. The effectors are not yet identified but viral proteins gp25 and gp26 have homology with amidases and holins, respectively, pointing for a typical holin/lisin mediated lysis. An average of ~220 virions is released per infected cell.

I.7. THESIS GOALS

Given the accumulated knowledge in the molecular biology of SPP1 and its host and the availability of molecular tools for their genetic manipulations, we have selected this system as a model to study the initial steps of phage infection of Gram-positive bacteria. The background knowledge at the beginning of this work was that SPP1 requires the presence of a membrane protein, YueB, to irreversible adsorb and infect *B. subtilis* cells (São-José *et al.*, 2004). The genomic context of the YueB receptor gene suggested an operon organization with upstream genes. One of the goals was to confirm this operon structure and at the same time assess the possible involvement of operon products in SPP1 adsorption and/or infection (Part II). Taking our and previous work in consideration (Yasbin *et al.*, 1976) we hypothesized that SPP1 adsorption would involve two consecutive steps, one reversible that preceded irreversible binding. We aimed to confirm this hypothesis, to identify the cell wall receptors involved in reversible adsorption and to understand its relevance in the overall adsorption kinetics (Part III). With the knowledge and tools that emerged from this study we extended our goals to the identification of tail structural proteins involved in each type of adsorption (Part IV).

The work on this theme has been the result of a close collaboration between the laboratories headed by Professor Mário de Almeida Santos (Faculty of Sciences of the University of Lisbon, Portugal) and Dr. Paulo Tavares (Centre National de la Recherche Scientifique, Gif-sur-Yvette, France).

PART II

THE *B. SUBTILIS* YUKE OPERON AND PHAGE SPP1 INFECTION

II.1. INTRODUCTION

More than twenty years ago Santos *et al.* (1983) described a *B. subtilis* mutation (*pha-2*) that conferred resistance to phage SPP1 infection but allowed SPP1 adsorption. The *pha-2* mutant strains retained the capacity to adsorb phages at a wild-type rate, but phage-host complexes could be readily separated by dilution, showing that adsorption was fully reversible. The mutation was found to be specific for SPP1 and related phages producing no effect in the adsorption and infection of other *B. subtilis* phages such as ϕ29, SP01 and ϕ25 (Santos, 1991). Genome mapping by PBS1-mediated transduction suggested that *pha-2* was localized in the vicinity of the genetic marker *ald-1* since these mutations presented 95% of co-transduction. This preliminary work prompted us to search in the surroundings of *ald-1* for the gene(s) affected in *pha-2* strains, and which would be involved in the synthesis of the receptor responsible for SPP1 irreversible adsorption (São-José *et al.,* 2004).

II.1.1 – ESTABLISHING YUEB AS THE RECEPTOR FOR SPP1 IRREVERSIBLE ADSORPTION

Further studies by Santos *et al* showed that irreversible adsorption of SPP1 required a protease-sensitive, heat-labile component associated with protoplasts and membrane fractions (Santos, 1991). Taking this into account an after analyzing the genomic context of *ald* (coding for L-alanine dehydrogenase) we have selected *yueB* as a good candidate for a gene encoding a SPP1 membrane receptor. Its putative product is an orthologue of Pip (phage infection protein), a membrane protein previously shown to be required for the irreversible binding of phage c2 to its host, *Lactococcus lactis* (see Part I). The deduced *yueB* product (1,076 residues) was predicted to insert in the cytoplasmic membrane by one amino-terminal and five carboxyl-terminal transmembrane domains (TMD's), with most of its sequence (>800 residues) facing the wall compartment, a membrane topology that was also expected for Pip (São-José *et al.,* 2004).

Our work from 2004 demonstrated that in fact the *yueB* product was essential for SPP1 infection, acting as the receptor for irreversible adsorption (São-José *et al.,* 2004). In that work, we isolated a *pha-2* spontaneous mutant (CSJ1) that had a genomic deletion covering ca. $^{2}/_{3}$ of the *yueB* gene and the 3' end of the upstream gene *yukBA*. We restored sensitivity and receptor activity in CSJ1 strain by complementing the lack of YueB via ectopic expression of a xylose-inducible *yueB* copy

inserted in the *amyE* locus. Moreover, the rate of SPP1 irreversible adsorption was shown to correlate with the level of *yueB* expression.

II.1.2 – *YUEB* GENOMIC CONTEXT AND THE PUTATIVE OPERON STRUCTURE

The genomic organization of the region surrounding the receptor gene and the analysis of its nucleotide sequence suggested that *yueB* would be part of an operon structure, delimited by putative hairpin-like secondary structures, as annotated previously by Oudega *et al.* (1997). Assuming that that was the case, the complementation experiment mentioned above would not rule out the possibility of an involvement of the other genes of the operon in the SPP1 infection process.

This chapter presents the work we have conducted to confirm the operon organization of *yueB* and surrounding genes and the study of the possible involvement of the genes of the cluster, in addition to *yueB*, in the SPP1 adsorption/infection processes.

II.2. MATERIALS AND METHODS

II.2.1 – BACTERIAL STRAINS, BACTERIOPHAGES AND PLASMIDS

All biological material used in this work is listed in Table II.1.

Table II.1 | *E. coli* and *B. subtilis* strains, bacteriophages, vectors and plasmids used in this work.

Strains, phages and plasmids	Genotype or relevant features	Reference
***E. coli* strains**		
TG1	*SupE thi-1* Δ(*lac-proAB*) Δ(*mcrB-hsdSM)5* (r_K^- m_K^-) [*F' traD36 proAB lacI*^{q} *ZΔM15*]	Stratagene
XL1-Blue MRF'	Δ(*mcrA)183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [*F' proAB lacI*^{q}*ZΔM15* Tn*10 (Tet*^{r}*)*]	Stratagene
MC1061	*araD139* Δ(*ara leu)7697* Δ*lacX74 galU galK hsdR2 strA mcrA mcrB1*	Casadaban and Cohen, 1980
***B. subtilis* strains**		
L16601	*B. subtilis* 168; SPP1 indicator strain	Margot and Karamata, 1996
1A772	*pheA1 trpC2 amyE*::*cat86*; Cm^R	BGSC [a]
CSJ1	L16601 derivative resistant to SPP1; ΔJ*yueB*	São-José *et al.,* 2004
CSJ2	L16601 derivative resistant to SPP1; *pha-2*	São-José *et al.,* 2004
CSJ3	L16601 derivative overexpressing *yueB* upon IPTG induction; Ery^R	São-José *et al.,* 2004
CSJ4	CSJ1 derivative with gene *ery*^{R} inserted in the *amyE* locus; Ery^R	São-José *et al.,* 2004
CSJ6	CSJ1 derivative with *yueB* inserted in the *amyE* locus; overproduces YueB upon xylose induction; Neo^R	São-José *et al.,* 2004
CBM6	L16601 derivative, *yukE*ΩpCBM6; Ery^R	This work
CBM7	L16601 derivative, *yukD*ΩpCBM7; Ery^R	This work
CBM8	L16601 derivative, *yukC*ΩpCBM8; Ery^R	This work
CBM9	L16601 derivative, *yukBA*ΩpCBM9; Ery^R	This work
CBM10	L16601 derivative, *yueB*ΩpCBM10; Ery^R	This work
CBM11	L16601 derivative, *yueC*ΩpCBM11; *yueC* conditional mutant; Ery^R	This work
CBM12	L16601 derivative, *yukE*ΩpCBM12; *yukE* conditional mutant; Ery^R	This work
CBM14	L16601 derivative, *yukD*ΩpCBM14; *yukD* conditional mutant; Ery^R	This work
CBM6.40	CBM6 derivative with gene *cat86* inserted in the *amyE* locus; Cm^R	This work

CBM15	CBM6.40 derivative with *yukE* inserted in the *amyE* locus; overproduces YukE upon xylose induction; Neo^R	This work
L.cat86	L16601 derivative with *cat*86 inserted in the 5' end of *yukF*; Cm^R	This work
L.E1E2	L.cat86 derivative with point mutations in *yukE*; Cm^R	This work
L.D1D2	L.cat86 derivative with point mutations in *yukD*; Cm^R	This work
L.Del6	L.cat86 derivative deleted of the *yukE* operon, including its promoter region; Cm^R	This work
L.Del6-Pxyl-YueB	L.Del6 derivative with *yueB* inserted in the *amyE* locus; overproduces YueB upon xylose induction; Neo^R	This work
CBM-GFP	L16601 derivative, *yueB*ΩpCB-GFP; $Spec^R$	This work
CBM429	CSJ3 derivative, *yueB*ΩpCB-GFP; overproduces YueB-GFP upon IPTG induction; Ery^R, $Spec^R$	This work
CBM191	CSJ6 derivative, *yueB*ΩpCB-GFP; overproduces YueB-GFP upon xylose induction; Neo^R, $Spec^R$	This work
CBM-L191	L16601 derivative with *yueB-gfp* inserted in the *amyE* locus; overproduces YueB-GFP upon xylose induction; Neo^R, $Spec^R$	This work
***B. subtilis* phages**		
SPP1	Lytic phage	Riva *et al.,* 1968
SP01	Lytic phage; used to control specificity of *yueB* mutations	Okubo *et al.,* 1964
Vectors and plasmids		
pMutin-4	Integration vector used for gene inactivation; Amp^R, Ery^R	Vagner *et al.,* 1998
pGR40	Integration vector; allows the insertion of genes in the *amyE* locus under the P_{xyl} promoter; Amp^R; Neo^R	Real and Henriques, 2006
pUS19	Integration vector; used in the construction of *yueB-gfp* fusion; $Spec^R$	Benson and Haldenwang, 1993
pBluescript II KS/SK +	Cloning vector; Amp^R	Stratagene
pEA18	Harbors GFP variant *mut2* coding sequence; Cm^R; $Spec^R$	Quiesel *et al.,* 1999
pCBM6	pMutin 4 derivative carrying a PCR product internal to *yukE*	This work
pCBM7	pMutin 4 derivative carrying a PCR product internal to *yukD*	This work
pCBM8	pMutin 4 derivative carrying a PCR product internal to *yukC*	This work
pCBM9	pMutin 4 derivative carrying a PCR product internal to *yukBA*	This work
pCBM10	pMutin 4 derivative carrying a PCR product internal to *yueB*	This work
pCBM11	pMutin 4 derivative carrying a PCR product covering the RBS and 5' end of *yueC*	This work

pCBM12	pMutin 4 derivative carrying a PCR product covering the RBS and 5' end of *yukE*	This work
pCBM14	pMutin 4 derivative carrying a PCR product covering the RBS and 5' end of *yukD*	This work
pCBM15	pGR40 derivative carrying a PCR product covering the RBS and coding sequence of *yukE*	This work
pKS-del/point	pKS derivative carrying a PCR product covering from the 5' end of *yukF* to the 5' end of *yukC* (*yukF' - yukE - yukD - 'yukC*)	This work
pKS-del/point *cat*86	pKS-del/point derivative with *cat*86 (Cm^R) inserted in the 5' end of *yukF*	This work
pCB-E1E2	pKS-del/point *cat*86 derivative with a stop codon and a *Hind*III site in the 5' end of *yukE*	This work
pCB-D1D2	pKS-del/point *cat*86 derivative with a stop codon and a *Hind*III site in the 5' end of *yukD*	This work
pCB-1-6	pSK derivative with the *yukF' - cat*86 - *PR* and *yueC'* fused to a fragment spanning from the 3' end of *yueB* to the 5' end of *yueC*; (Cm^R)	This work
pCB-Del 1-6	pCB-1-6 derivative where the *PR* was removed by inverse PCR	This work
pCB-GFP	pUS19 derivative carrying the 3' end of *yueB* fused to *gfp*	This work

[a] – BGSC – Bacillus Genetic Stock Center

II.2.2 – OLIGONUCLEOTIDES

The oligonucleotides used in this work are presented in Table II.2. These primers were used in the construction of *yukE* operon mutants and of *yueB-gfp* fusions. Primers pMutin-1, lacZR-1, yukE-2, yukD-2, BS-11, yueB-33, yueC-2 and yukF-2 were used in the confirmation of clones DNA structure by PCR, as they flanked regions of interest (see section II.2.5 for details). Primer BS-9 was used for the first-strand cDNA synthesis and yukE-1 and yukC-2 were used for the PCR reactions on the produced cDNA (RT-PCR).

Table II.2 | Oligonucleotides used in this work.

Oligonucleotide	Sequence (5'→3')[a]
BS-9	CGATCACGCTTCATACGTTTC
BS-11	TCC**gaattc**TTTTATTATGATTGTTGTTGTTTGTG
yukE-1	ATGGCAGGATTAATTCGTGTCAC
yukE-2	AGTGTAGAACTTCGTTTAGAC

yukE-3	TGA**gaattc**TGGCAGGATTAATTCGTGTCAC
yukE-4	TG**ggatcc**GAAGGTTTGAGCTGCTCGTATT
yukE-7	GTG**gaattc**CCCATATTTAGGCTGATGGCA
yukD-2	AGTATTATAAAGTTCGGCCAG
yukD-3	ACA**gaattc**TTGAAACATTATAACGGCAGTG
yukD-4	CAC**ggatcc**CTTGTTCACCACTCTGATCC
yukD-5	TCT**gaattc**TGAAAAGCATGTGGGAAGGTG
yukD-6	TCA**ggatcc**CTGCCGTTATAATGTTTCAAATC
yukC-2	TAG**ggatcc**TATATTCTTTACGACTCCCAGAC
yukC-3	TCA**gaattc**AATCGATGTGACGGATGATGA
yukC-4	GTA**ggatcc**AAACGCAAAAGCGCCGTCAA
yukBA-3	TCA**gaattc**AAATCTGTCTAACCTCCCGTC
yukBA-7	GAG**ggatcc**ACAATTTCATCCCGATTTCCCG
yueB-23	GAA**ggatcc**ATTCGAGCGTTGCCTTTTGAG
yueB-27	TGC**gaattc**ACAGAACAACGAAAAAGCTTGAT
yueB-33	CAGTAA**ggatcc**AAAAAAGTATGGGTAAAGAACGTTACC
yueC-2	CAC**ggatcc**TACGCCTTCTTGTTTTCCTGC
yueC-3	TGA**ggattc**CAAATCTTTTATCAGCTCGGCG
yueC-4	GTA**gaattc**TGGGCACTTGTTTACAATGGG
pMutin-1	TTCTACATCCAGAACAACCTC
lacZR-1	GTGCTGCAAGGCGATTAAGTT
KE-Spe	TGA**actagt**AGGAGGTAATGAGGAATATGGC
KE-Eco	TTC**gaattc**TTATCCGCGGATTTGATTTGCG
Quick-E1	GATTAATTCGTGTCACACCCT**aagctt**TAAGAGCGATGGCGAAGC
Quick-E2	GCTTCGCCATCGCTCTTA**aagctt**AGGGTGTGACACGAATTAATC
Quick-D1	GTGTCTTTGATCTCAGATTGT**aagctta**ACCACCCGGTGAAAAAAG
Quick-D2	CTTTTTTCACCGGGTGGT**aagctta**ACAATCTGAGATCAAAGACAC
Cat86-Pst	CAC**ctgcag**GGTTTGGTTTATCCTCTGTTATG
Cat86-79	CCAAAG**gggcccc**GAAACCGTGTAATAGTAATG
yukF-2	ATG**ctgcag**GATAGGAGCGGGAGACAAAAG
yukF-3	GAAT**gaattc**TCGCCAGCAGCACGTAAACC
PE-yukC	AGC**ggatcc**CTGTTTTCTCGGCAACAGCTTC
QuickCM-KF1	CAAGTCTGTCAAAACTATATTTA**cccggg**TCATT**ctgcag**GGTCATAAATCT CCCTACTTT
QuickCM-KF2	AAAGTAGGGAGATTTATGACC**ctgcag**AATGA**cccggg**TAAATATAGTTT TGACAGACTTG
Delta-2	CCTAAATATGGGCTAATTCACTTA
Delta-5	AAAATAAGTGAATTAGCCCATATTTAGGTGGGCACTTGTTTACAATG GGAATAACAG

Delta PR cat86	AAA**atgcat**AGATTTATGACC**ctgcag**GGTTT
Delta PR yueC	TAT**atgcat**GGGCACTTGTTTACAATGGGAA
YB GFP-1	CGTTATTCG**gaattcctcgag**TCCTTGTCGGACTTA
YB GFP-2	CTCCAGTGAAAAGTTCTTCTCCTTTACTGTTGTTGTTGTTCGCTTCAT ACGTTTCATCGCTTTCTGCTGT
gfp 30D	AGTAAAGGAGAAGAACTTTTCACTGGAG
gfp 749R	GATC**ctcgaggaattc**TTATTTGTATAGTTCATCCATGC

[a] – Boldface lowercase letters indicate endonuclease restriction sites (*Eco*RI-**gaattc**; *Bam*HI-**ggatcc**; *Sma*I-**cccggg**; *Pst*I-**ctgcag**; *Nsi*I-**atgcat**; *Hind*III-**aagctt**; *Xho*I-**ctcgag**; *Spe*I-**actagt**).

II.2.3 – GROWTH CONDITIONS AND MICROBIOLOGICAL METHODS

II.2.3.1 – *Escherichia coli* growth conditions

E. coli strains were routinely grown in Luria-Bertani (LB) medium (Sambrook and Russel, 2001) at 37ºC with aeration. *E. coli* transformants harboring pMutin4 derivatives were selected and propagated at 30ºC. Agar was added to LB medium to a final concentration of 1.5% (wt/vol) in order to obtain bottom LB plates. When required, ampicillin (100 μg/mL), chloramphenicol (30 μg/mL), or spectinomycin (100 μg/mL) was added to the culture medium for plasmid selection. When possible, transformants carrying recombinant plasmids were selected by the β-galactosidase α-complementation assay in LB plates supplemented with X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) and IPTG (isopropyl-β-D-thiogalactopyranoside) (Sambrook and Russel, 2001).

II.2.3.2 – *Bacillus subtilis* and bacteriophage growth conditions

Unless stated otherwise, *B. subtilis* strains were grown overnight in LB medium at 30ºC, with aeration, re-inoculated 1:100 in fresh LB medium and grown at 37ºC, with agitation, until use. Agar was added to LB medium to a final concentration of 0.7 or 1.5% (wt/vol) in order to obtain top or bottom LB plates, respectively. *B. subtilis* transformants expressing β-galactosidase were selected on X-Gal-supplemented (0.02%, wt/vol) LB plates. Erythromycin (0.5 μg/mL), chloramphenicol (5 μg/mL), spectinomycin (100 μg/mL), or neomycin (7.5 μg/mL) were used to select *B. subtilis* transformants.

Propagation of phages SPP1 and SP01 was as described by Santos *et al.,* 1983. Phages were maintained at 4ºC and diluted in TBT buffer (100 mM Tris-HCl, 100 mM NaCl, 10 mM $MgSO_4$; pH 7.5). Determination of phage titers was routinely performed by incorporating 100 μL aliquots of phage serial dilutions and 100 μL of the indicator strain in 10 mL of top LB-Agar supplemented with 10 mM of $CaCl_2$.

II.2.3.3 – Phage purification

When necessary, phages from cleared lysates were PEG-precipitated and purified by a discontinuous CsCl gradient according to Sambrook and Russel, 2001. The CsCl solutions with densities of ρ=1.7, 1.5 and 1.45 were prepared in TBT. CsCl bands containing purified phages were double-dialyzed at 4ºC against a 100-fold excess of TBT.

Wild type SPP1 phages were also purified by FPLC using an anion exchange column (RESOURCE Q, 6 mL). Before loading the column, CsCl purified phages were dialyzed three times against 100 volumes of column buffer (100 mM Tris-HCl, 10 mM $MgCl_2$, pH 7.5) to remove NaCl. Phages were eluted with a gradient of column buffer supplemented with 2 M NaCl (0 to 25% of elution buffer in 5 minutes).

II.2.3.4 – Measurement of SPP1 irreversible adsorption to *B. subtilis* cells

SPP1 irreversible adsorption was measured with exponentially growing cultures of *B. subtilis* (A_{600}=0.8, about 10^8 colony forming units per mL). Specifically, one milliliter samples of *B. subtilis* cultures were supplemented with 15 mM $MnCl_2$ and maintained at 37ºC. Mn^{2+} ions fulfill the role of Ca^{2+} in phage adsorption with the advantage of inhibiting *B. subtilis* growth and SPP1 intracellular development (Santos, 1991). Phage SPP1 was added to the samples in order to obtain ~10^7 plaque forming units per mL (pfu/mL) and at defined time points 10 μL aliquots were diluted 100 fold in TBT buffer supplemented with 10% chloroform. After vigorous vortexing during 5 seconds the mixtures were allowed to equilibrate for 5 min at room temperature, centrifuged (15000g, 5 min) and the supernatants recovered for the enumeration of free phages (unadsorbed plus virions that were reversibly bound). Control mixtures without added cells were used to confirm the phage input in each experiment. The **irreversible adsorption constant** k_{ads} was calculated as the ratio between the irreversible adsorption rate and

the bacterial cell mass expressed as A_{600} ($k_{ads} = \ln(P0/P)/\Delta t.A_{600}$, where *P0* is the phage input and *P* the fraction of free phages after a Δt period).

II.2.3.5 – One-step growth curve and burst-size determination

One-step growth curves and determination of phage growth parameters were adapted from protocols described by Adams (1959). *B. subtilis* strains were grown in LB, supplemented with the required antibiotics, until the cultures reached an A_{600}=0.8. The cultures were supplemented with 15 mM of $CaCl_2$ just before addition of 10^6 pfu/mL of SPP1 to 1 mL of culture samples (i.m.=0.01). Phage adsorption was allowed to occur for 10 minutes at 37ºC. A first 100 µL sample was taken, cells were centrifuged and the supernatant was kept for enumeration of free phages. Adequate decimal dilutions of infected cells were prepared in fresh pre-warmed medium with $CaCl_2$. Phage release was followed by plating 100 µL aliquots of each dilution every 5 minutes until 100 minutes post-infection.

Average burst-size values were obtained from the one-step growth curves by dividing the total number of released phages by the number of infected cells in the latent period.

II.2.4 – MOLECULAR AND BIOCHEMICAL METHODS

II.2.4.1 – DNA extraction and separation by gel electrophoresis

Routine plasmid DNA extraction from *E. coli* strains was performed by the alkaline lysis method described in Sambrook and Russel, 2001. High quality mini and midi plasmid DNA preparations were carried out using JETquik miniprep extraction kit (Genomed) and the Wizard® *Plus* SV Minipreps DNA Purification System (Promega).

B. subtilis chromosomic DNA was obtained from 4-mL culture samples. Cells were recovered in 0.2 mL of TEG buffer (Sambrook and Russel, 2001) supplemented with 2 mg of fresh lysozyme/mL, and after an incubation period of 10 minutes at 37°C, the cells were lysed by the addition of 0.6 mL of GES-β (Pitcher *et al.,* 1989). The lysate was then extracted twice with a phenol-chloroform-isoamyl alcohol mixture (25:24:1) and once with a chloroform-isoamyl

alcohol solution (24:1) before DNA precipitation with 2-propanol. After washing with 70% ethanol, DNA was finally resuspended in TE buffer (Sambrook and Russel, 2001) supplemented with 10 μg of RNase A/mL.

DNA molecules were separated by standard agarose gel electrophoresis (Sambrook and Russel, 2001) using UltraPure agarose (Invitrogen) and 0.5X TBE buffer as gel/running buffer. Agarose gels from 0.5 to 2% (wt/vol), with 0.5 μg/mL ethidium bromide, were routinely used. The DNA molecular weight standard used was the 1 Kb*plus* DNA Ladder (Invitrogen). DNA bands were visualized under U.V. lightening and images captured by a digital camera.

II.2.4.2 – RNA extraction and separation by gel electrophoresis

All solutions and equipment used when working with RNA were pre-treated with 0.05 M NaOH and/or DEPC (Sigma) to prevent RNase activity (Sambrook and Russel, 2001).

B. subtilis total RNA was prepared from 20 mL culture samples (A_{600}=0.8) using the TRIzol reagent (Invitrogen) according to the supplier's instructions, except that the cells were recovered in TE buffer supplemented with 10 mg of fresh lysozyme/mL, incubated for 5 min at 37°C, and lysed by the addition of 5 mL of the reagent. RNA molecules were separated in 1% RNase-free agarose gels prepared in NBC buffer (50 mM boric acid, 1 mM sodium citrate, 5 mM NaOH, pH 7.5) containing 1% formaldehyde and 0.25 μg of ethidium bromide/mL and run in the same buffer. Before gel loading, RNA samples (10 to 15 μg in up to 5 μl) were supplemented with 2 μl of 10X NBC buffer, 3 μl of 37% formaldehyde, and 10 μl of formamide, followed by a denaturing step of 5 min at 65°C. Standard RNA molecular weight marker was from Invitrogen (0.24-9.5 Kb RNA Ladder). RNA bands were visualized as described above for DNA bands.

II.2.4.3 – Northern blot analysis

Northern blotting was carried out essentially as described by Sambrook and Russel, 2001. After checking RNA integrity by staining the RNA-carrying nylon membranes with a solution of 0.03% (wt/vol) methylene blue–0.3 M sodium acetate (Wilkinson *et al.,* 1990), the membranes were incubated in a prehybridization solution (5X Denhardt's solution, 6X SSC [1X SSC is 0.15 M NaCl, 0.015 M sodium citrate], 0.5% SDS, and 0.1 mg of salmon sperm single-stranded DNA/mL) for 60 min at 68°C. The prehybridization solution was then replaced by the

hybridization solution (2X Denhardt's solution, 6X SSC, and 0.1% SDS). ^{32}P-labeled probes prepared with the Ready-To-Go DNA Labelling Beads (minus dCTP) kit (Amersham Pharmacia Biotech) were added to the hybridization solution and incubated overnight at 60°C. Unattached probes were discarded with two washes of 10 min at 60°C with a 5X SSC–0.1% SDS solution, followed by two to four washes of 10 min with a 2X SSC–0.1% SDS mixture at room temperature. After a brief wash with 2X SSC, the membranes were autoradiographed.

II.2.4.4 – General recombinant DNA techniques

Restriction endonuclease digestions (New England Biolabs, MBI Fermentas, TaKaRa Biomedicals) and DNA ligations (T4 DNA Ligase, Invitrogen) were performed essentially as described by Sambrook and Russel, 2001, and according to the supplier's instructions. PCR amplification of DNA fragments was carried out in a RoboCycler Gradient 96 thermocycler (Stratagene) or in a Biometra thermocycler using *Taq* DNA polymerase (Invitrogen) or *Pfu* polymerase (Stratagene) as recommended by the suppliers. Introduction of mutations by PCR was performed with QuickChange XL Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. All the oligonucleotides used in this work were purchased from Invitrogen and are listed in Table II.2. When needed, DNA products resulting from restriction endonuclease cleavage or PCR amplification were extracted from agarose gels and purified using the JETquik gel extraction spin kit (Genomed). Highly specific PCR products were directly purified from the PCR mixtures using the Wizard® SV Gel and PCR Clean-up System (Promega). DNA sequencing services were purchased either from the ICAT Sequencing Unit, Portugal or from GATC Biotech AG, Germany.

RNA samples for RT-PCR were pre-treated with RNase free DNaseI (Invitrogen, 1U of DNaseI per μg of RNA in a final volume of 10 μL). After 15 minute incubation at room temperature DNaseI was inactivated by adding 1 μL of 25mM EDTA and incubating for 10 minutes at 65ºC. Synthesis of the first-strand cDNA and the PCR reaction were made according to the suppliers' instructions (SuperScript III RNase H$^-$ Reverse transcriptase, Invitrogen).

Development of *E. coli* and *B. subtilis* competence and transformation was as described by Chung *et al.,* 1989 and Yasbin *et al.,* 1973, respectively.

II.2.4.5 – Measurement of β-Galactosidase activity

β-Galactosidase activity was determined essentially as described by Miller, 1972, from 0.1-, 1- or 10 mL samples of *B. subtilis* liquid cultures, taken at different growth stages.

II.2.4.6 – Small-scale production of total protein extracts

To monitor protein production by *E. coli* strains, total protein extracts were prepared by pelleting cells from 1 mL culture samples and resuspending in 100 μL of 2X SDS-PAGE loading buffer (Laemmli, 1970).

To produce total protein extracts of *B. subtilis* strains, cells from 2 mL samples of exponentially growing cultures (A_{600}=0.8) were resuspended in 50 μL of a lysis buffer (50 mM Tris-HCl pH=8, 300 mM NaCl, 10 mM $MgCl_2$, 50 mM glucose) supplemented with 1 mg/mL lysozyme, 50 U/mL benzonase (Roche Applied Science) and Complete Mini EDTA-free Protease Inhibitor Cocktail (Roche Applied Science). Cells were treated for 5-10 minutes at 37ºC and, after 10 minutes on ice, 0.5 μL of 10% NP-40 was added. The homogenized extract was divided in small aliquots and kept at -20ºC until use. For protein analysis by SDS-PAGE, an equal volume of 2X SDS-PAGE loading buffer was added to the sample.

Protein concentration was determined by the method of Bradford, using a Bio-Rad kit, according to the instructions provided by the manufacturer.

II.2.4.7 – Protein analysis by SDS-PAGE and Western blotting

Adequate volumes of protein samples were analyzed by electrophoresis on 8-10% SDS-PAGE gels (Laemmli, 1970). Proteins were visualized by Coomassie Blue staining (Sambrook and Russel, 2001). Protein blotting onto nitrocellulose membranes (Bio-Rad) was routinely performed in a Trans-Blot SD Transfer Cell Apparatus (Bio-Rad). Blotting of high molecular weight proteins was carried out either with a Trans-blot Cell Tank Transfer system (Bio-Rad) or an XCell SureLock Mini-Cell (Invitrogen).

Immuno-detection of the receptor YueB780 was performed with rabbit polyclonal sera diluted 1:30,000. Green Fluorescence Protein (GFP) fusion proteins were detected using the Anti-GFP

(Clontech), following the suppliers' instructions. Antigen/antibody complexes were detected with the Chemioluminescence Western Blotting Kit (Roche Applied Science), according to the manufacturer's instructions.

II.2.5 – CONSTRUCTION OF *B. SUBTILIS* MUTANT STRAINS

II.2.5.1 – pMutin-4 based gene inactivation

Individual disruption of different genes of the *B. subtilis* chromosome was achieved using the integrative vector pMutin-4 (Vagner *et al.,* 1998). In this method, 5'end internal regions of the target genes were PCR-amplified from the genome of the wild type *B. subtilis* strain L16601. The PCR products were digested with *Eco*RI and *Bam*HI, cloned into similarly digested pMutin-4, and the recombinant plasmids recovered in *E. coli* strain MC1061. The constructs were then used to transform *B. subtilis* competent cells. The integrants resulting from the recombination events were selected in X-gal-erythromycin plates. The DNA structure obtained upon integration in the *yukBA* gene is shown in Figure II.1 as example.

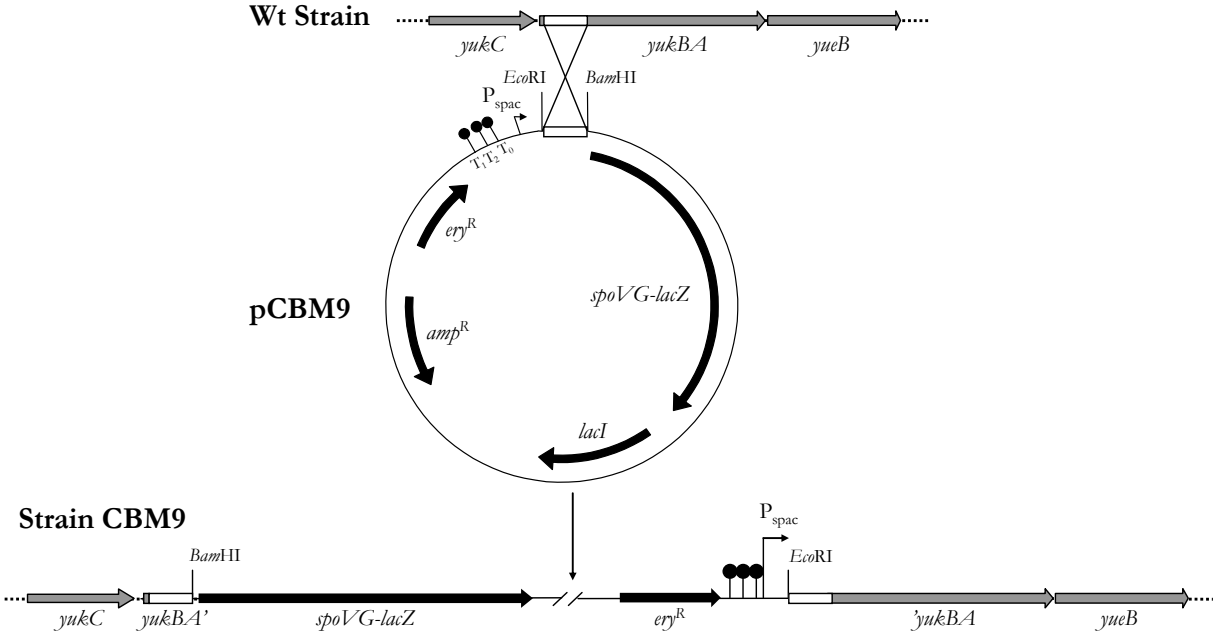


Figure II.1 | Schematic representation of the strategy used for disruption of gene *yukBA*. The internal region of the target gene is represented in white. The pMutin-4 derivative integrates by a single *crossing-over* event between homologous sequences.

When the cloned PCR products covered the translation signals and the 5' end of the target genes, the strains resulting from plasmid integration corresponded to conditional mutants, in which the target gene became separated from its natural transcriptional elements and placed under the control of the IPTG-inducible promoter, P_{spac}. In this type of mutants, the expression of the target gene is restored in the presence of IPTG (see as example the construction of the *yueC* conditional mutant, Fig. II.2).

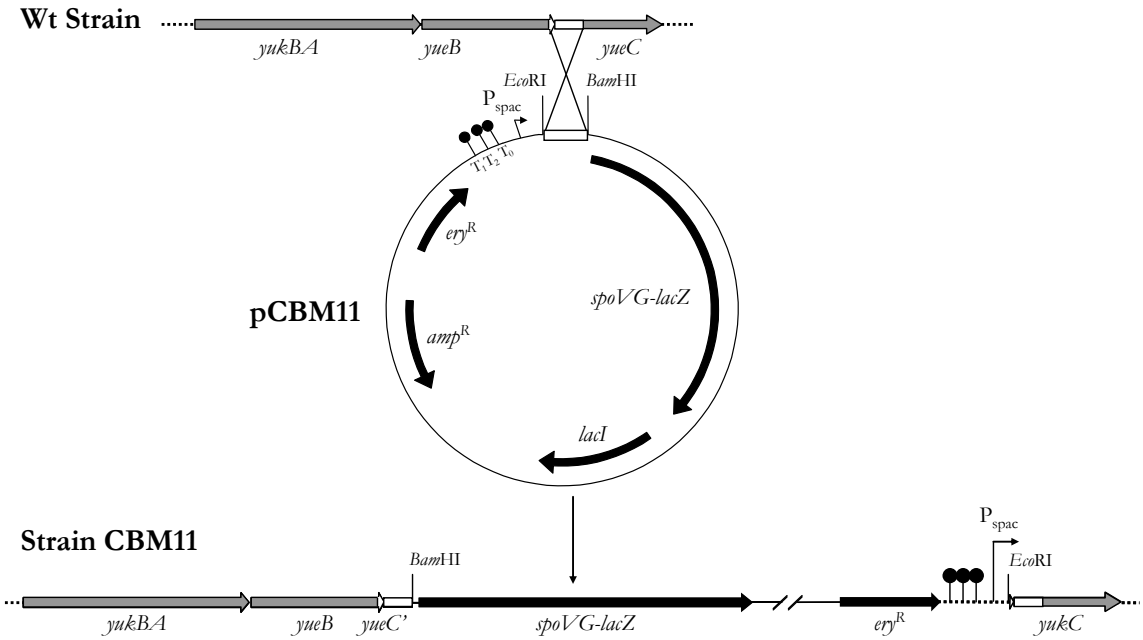


Figure II.2 | Schematic representation of the strategy used for the construction of a *yueC* conditional mutant. The target region is represented in white and comprehends from the 3' end of the upstream gene, *yueB*, to the 5' end of the target gene *yueC*, including its translational elements. Integration occurs by a single *crossing-over* event between homologous sequences.

II.2.5.2 – Construction of yukE operon integrants

The genes of the *yukE* operon were individually inactivated through the integration of pMutin-4 derivatives (Vagner *et al.,* 1998). For that, DNA fragments (~150 to 400 bp) corresponding to internal regions of the genes *yukE*, *yukD*, *yukC*, *yukBA*, and *yueB* were PCR amplified (using primer pairs yukE-3/yukE-4, yukD-3/yukD-4, yukC-3/yukC-4, yukBA-3/yukBA-7, and yueB-27/yueB-23, respectively, Table II.2), double digested with *Eco*RI and *Bam*HI, and ligated to similarly digested vector pMutin-4. The resulting recombinant plasmids were obtained in *E. coli* strain MC1061 and then used to transform *B. subtilis* L16601. The integrants (CBM6, CBM7, CBM8, CBM9, and CBM10, respectively) resulting from Campbell-type recombination between the chromosomal loci and the cloned sequences were selected for erythromycin resistance and

blue color in X-Gal–LB plates. Gene disruption resulting from each integration event was confirmed by PCR using pMutin-4-specific primers (pMutin-1 and lacZR-1) and primers flanking the target loci. A similar strategy was used to construct *yueC*, *yukE* and *yukD* conditional mutants. In these cases, the PCR products (obtained with primer pairs yueC-3/yueC-4, yukE-7/yukE-4 and yukD-5/yukD-6, respectively) covered the translation signals and the 5' end of each *orf*, resulting in strains CBM11, CBM12 and CBM14.

II.2.5.3 – Ectopic expression of *yukE*

Strain CBM6.40 was obtained via transformation of CBM6 with chromosomal DNA of strain 1A772 (*amyE*::*cat86*) and selecting for chloramphenicol resistant colonies. In parallel, a DNA fragment containing the *yukE* gene (primer pair KE-Spe and KE-Eco, Table II.2), double digested with *Spe*I and *Eco*RI, was joined to the equally digested vector pGR40, generating pCBM15 selected in *E. coli* strain TG1. In plasmid pCBM15, the *yukE* gene is under the control of the xylose inducible promoter (P_{xylA}) and is flanked by the 5' and 3' portions of the dispensable *amyE* gene. This plasmid was then used to transform strain CBM6.40 and the integrants resulting from the double cross-over between the *amyE* segments were selected by marker replacement (neomycin resistance and chloramphenicol sensitivity; strain CBM15).

II.2.5.4 – Introduction of nonsense mutations in *yukE* and *yukD*

First, a PCR product obtained with primers yukF-3 and PE-yukC, spanning from the 5' end of *yukF* to the 5' end of *yukC* (*i.e.*, covering *yukF'*, *promoter region*, *yukE*, *yukD* and *yukC'*; Fig. II.3A) was cloned in the *Eco*RI/*Bam*HI sites of pKS vector, generating plasmid pKS-del/point (Table II.1). Two restriction sites (*Pst*I and *Sma*I) were created in the 5' end of *yukF* by site-directed mutagenesis using mutagenic primers QuickCM-KF1 and QuickCM-KF2 and the QuickChange XL Site-Directed Mutagenesis Kit (Stratagene). The *cat86* gene was amplified from plasmid pSD009 (a gift from Susana Domingues) with primers Cat86-Pst and Cat86-79 and introduced in the *Pst*I/*Sma*I double digested pKS-del/point plasmid, originating plasmid pKS-del/point-*cat*86 (Fig. II.3B). This plasmid was then used to transform *B. subtilis* strain L16601, creating strain L.cat86 (control strain; Fig. II.4B).

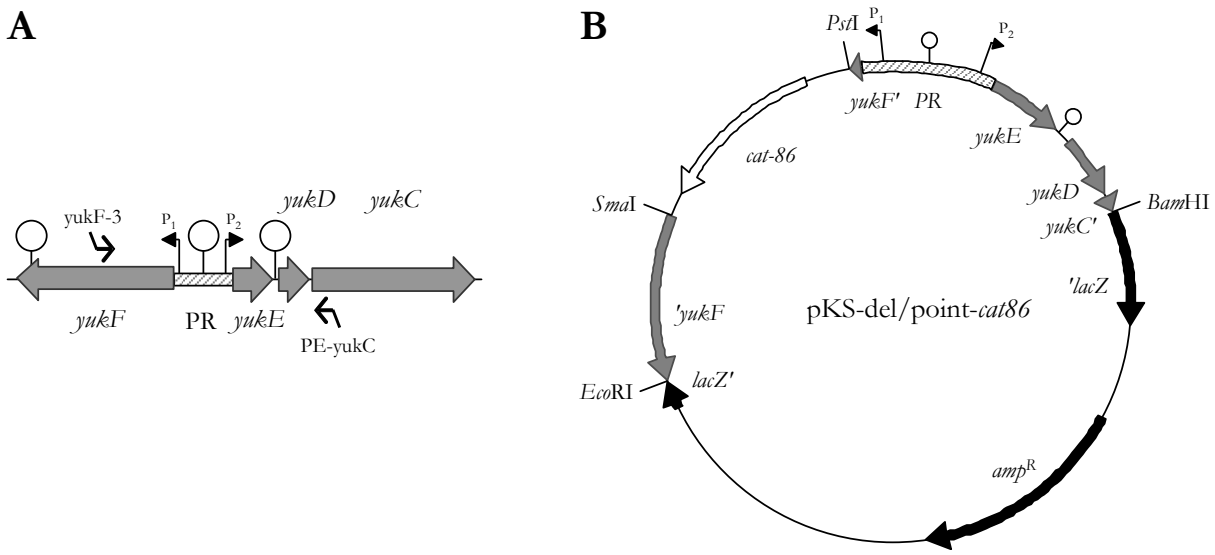


Figure II.3 | (**A**) Schematic representation of the genomic region and of the hybridization position of the oligonucleotides yukF-3 and PE-yukC, used to construct plasmid pKS-Del/point. (**B**) Final structure of plasmid pKS-Del/point -*cat86*. *B. subtilis* genes, pKS elements and *cat86* are represented in gray, black and white, respectively. The promoter region (PR) is represented as a striped box and the putative promoters, P_1 and P_2, are depicted as two black arrows. Inverted repeats are indicated ' ⚲ '.

To introduce a "TAA" stop codon and an overlapping *Hin*dIII site at the 5' end of *yukE* and *yukD* a QuickChange (Stratagene) protocol was performed using the pairs of oligonucleotides Quick-E1/Quick-E2 and Quick-D1/Quick-D2, respectively, and the plasmid pKS-del/point-*cat*86 as template. The resulting mutagenic plasmids pCB-E1E2 and pCB-D1D2 (Fig. II.4A), isolated in MRF', were then used to transform strain L16601 and recombinants selected for chloramphenicol resistance. The recombinants originated from double-crossover were identified by PCR using the flanking primer pair yukF-2/yukC4. The introduction of point mutations in the selected strains L.E1E2 and L.D1D2 was first confirmed by *Hin*dIII digestion of PCR products (L.E1E2 - BS11/yukE-2; L.D1D2 - yukD-5/yukD-2) and then by DNA sequencing (Table II.4, in section II.3.4).

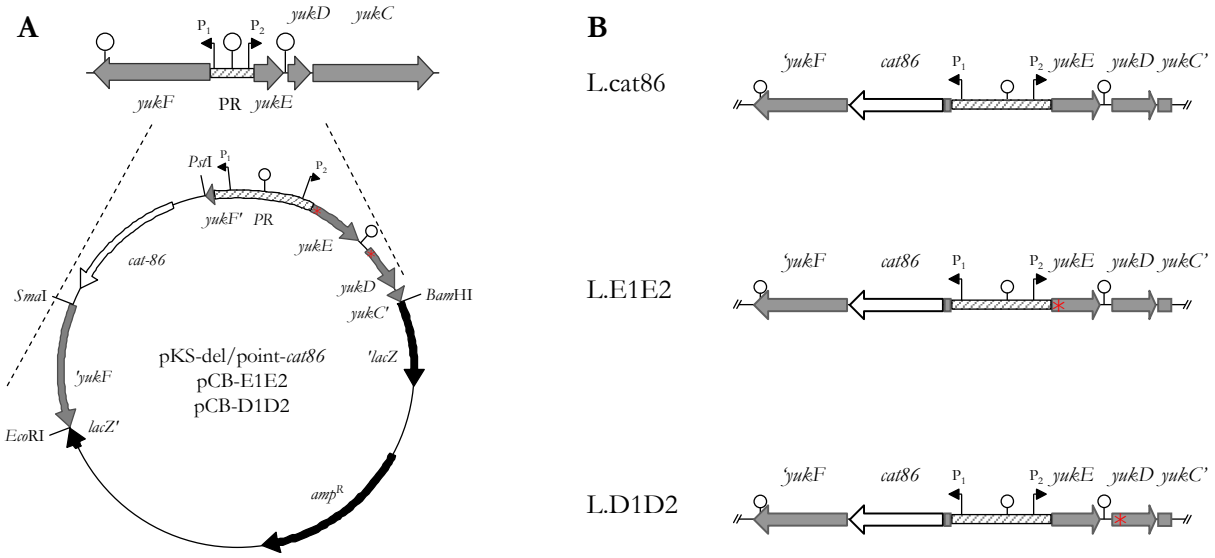


Figure II.4 | (**A**) Schematic representation of the strategy used to introduce non-sense mutations in *yukE* and *yukD* (red asterisks denote non-sense mutations). The double cross-over between the homologous regions (in gray) of the operon and of the plasmid allowed the introduction of the chloramphenicol resistance gene and of the *yukE* and *yukD* point mutations. (**B**) Final genomic structure upon recombination with plasmids pKS-del/point-*cat86*, pCB-E1E2 and pCB-D1D2. In gray are represented the *B. subtilis* genes, in black are the pKS elements and in white is *cat86*. Red asterisks denote non-sense mutations. The promoter region (PR) is represented as a striped box and the putative promoters, P_1 and P_2, are depicted as two black arrows. Inverted repeats are indicated ' ⚲ '.

II.2.5.5 – Construction of yukE operon deletion mutant

In a first step a PCR product was obtained with primers yukF-3 and Delta-2, which covered the region from the 5' end of *yukF* to just upstream of the *yukE* RBS carried in plasmid pKS-del/point-*cat*86 (Fig. II.4A). In a second PCR reaction the region of the L16601 genome covering the 3' end of *yueB* and the 5' end of *yueC* (primers Delta-5 and Delta-6) was amplified. The forward primer Delta-5 had a 5' extension complementary to the reverse primer Delta-2 used in the first PCR. This allowed the generation of a fusion PCR product using flanking oligos yukF-3 and Delta-6. This fusion product was then cloned in an *Eco*RI/*Bam*HI pSK digested vector, generating plasmid pCB-1-6 (carrying *yukF', cat86, promoter region* and *yueC'*), isolated and amplified in *E. coli* MRF'. After linearization, this plasmid was used to transform strain L16601 aiming for the recombination event that would lead to the *yukE* operon deletion. However, after several attempts we were only able to obtain mutant strains resulting from double recombination between the *yukF* and the intergenic promoter region. To overcome this problem the promoter region was removed from the constructed plasmid by inverse PCR using the pair of primers Delta PR cat86/Delta PR yueC, digesting with *Nsi*I and re-joining the plasmid. The resulting construct was renamed pCB-Del 1-6 (Fig. II.5). Strain L.Del6 is a L16601 derivative that was

obtained after transformation with pCB-Del 1-6, and which carried a deletion that joined promoter less *yukF* to *yueC*.

For the construction of strain L.Del6-Pxyl-YueB, competent cells of strain L.Del6 were transformed with chromosomic DNA from the xylose inducible YueB overproducing strain, CSJ6. The recombinants were selected for chloramphenicol and neomycin resistance, and the final genomic structure was confirmed by discriminatory PCR reactions, making sure that the *yukE* operon had not been re-introduced.

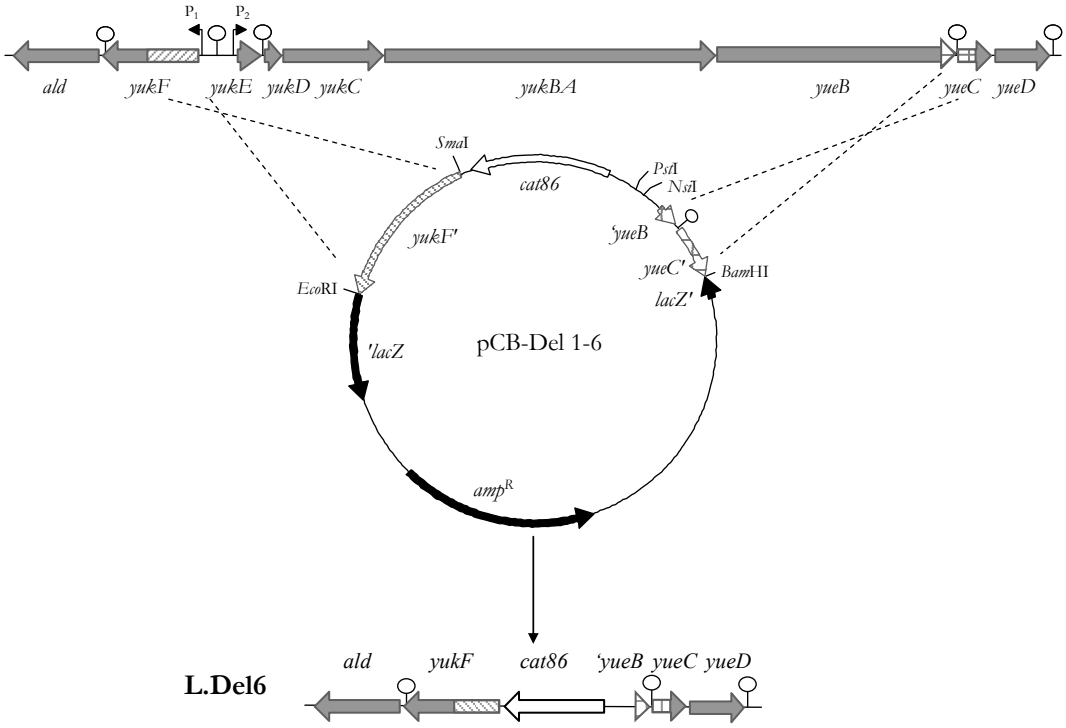


Figure II.5 | Schematic representation of the strategy used for the construction of deletion mutant of the *yukE* operon (genes not drawn to scale). The double cross-over between the homologous regions (signaled with different fills) of the operon and of the plasmid pCB-Del 1-6 allowed for the selection of chloramphenicol resistant recombinant strains carrying the deletion of the *yukE* operon (strain L. Del6). In gray are represented the *B. subtilis* genes, black arrows represent the pKSII+ elements and the chloramphenicol resistance gene (*cat86*) is depicted as a white arrow. The putative promoters, P_1 and P_2, are depicted as two black arrows and the inverted repeats are indicated ' ⚲ '.

II.2.5.6 – Expression of *yueB-gfp* fusions

For the construction and expression of the YueB-GFP fusion protein the *gfp mut2* gene was PCR amplified from vector pEA18 (primer pair gfp 30D/gfp 749R; Table II.2; this vector was a kind gift from Dr. A. O. Henriques) and fused, by PCR, to another PCR product that carried the 3'

end of *yueB* except for the stop codon (YB GFP-1/YB GFP-2). The PCR that fused these two PCR products with partial overlapping sequences was performed using the oligos YB GFP-1 and gfp 749R. This fragment was digested with *Eco*RI and cloned in the equally digested vector pUS19. The resulting plasmid isolated in *E. coli* strain MRF', pCB-GFP, was sequenced in the *yueB'-gfp* junction before *B. subtilis* transformation. A plasmid with the correct sequence was used to transform three different strains (Fig. II.6): L16601, where expression of the fusion protein is under the control of the *yueB* native promoter (CBM-GFP); CSJ3, in which the P_{spac} promoter controls the YueB-GFP expression (CBM429); CSJ6, in which promoter P_{xylA} drives the expression of YueB-GFP from P_{xylA} (CBM191). A fourth strain, CBM-L191, was constructed by transforming the wild type *B. subtilis* L16601 with genomic DNA of CBM191. In this case, the native YueB and the YueB-GFP fusion were produced in the same strain.

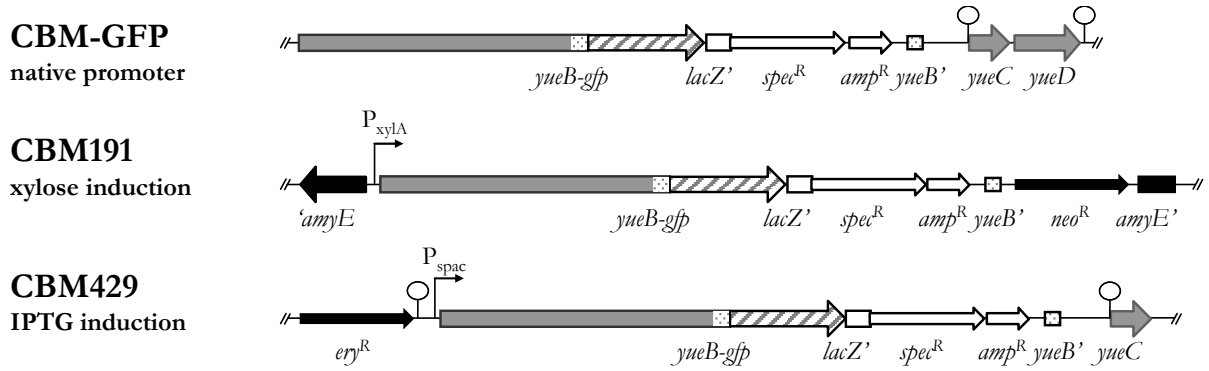


Figure II.6 | Schematic representation of the genomic structure resulting from pCB-GFP integration, in the three *B. subtilis* strains used. Native *B. subtilis* genes are represented in gray, in black are the elements of the previous integrated plasmid and in white the pCB-GFP genes. The *gfp* gene is represented with gray stripes and the duplicated *yueB* segments resulting from plasmid integration are as gray dots. Inverted repeats are indicated ' ⚲ '.

II.2.6 – BIOINFORMATICS ANALYSIS

Protein homology searches were carried out by BLAST and PSI-BLAST analysis (http://www.ncbi.nlm.nih.gov). Conserved and functional domains were identified with RPS-BLAST (http://www.ncbi.nlm.nih.gov). Transmembrane segments and membrane topology were inferred from analysis with TMHMM V2.0 (http://www.cbs.dtu.dk/services/TMHMM/). Multiple sequence alignments of protein sequences were performed with ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Prediction of Rho-independent terminators and secondary structures were performed with DBTBS (http://dbtbs.hgc.jp/) and Sci Ed Central Clone Manager. Predictions of SigmaA-dependent promoter regions performed with BPROM

tool
(http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb).

II.3. RESULTS

II.3.1 – BIOINFORMATICS ANALYSIS OF THE *YUEB* GENOMIC REGION

In the annotated *B. subtilis* genome region containing *yueB* (Z99120, section 17 of 21) the *yueB* gene is preceded by five putative genes (*yukE* to *yukA*) followed by two other *orf*s (*yueC-D*) all transcribed in the same direction. Genes *ald* and *yukF* have an opposite orientation relatively to this cluster (Fig. II.7A). As previously documented (Oudega *et al.,* 1997), divergent putative promoters were assigned in the 500 bp region between *yukF* and *yukE*, and two terminators, at the 3' end of *yukF* and *ald*, were also identified. Four additional Rho-independent terminators were predicted using the database of transcriptional regulation in *Bacillus subtilis* (DBTBS) and the Sci Ed Central Clone Manager tools. They locate at the 3'end of *yukE*, *yueB* and *yueD* and in the intergenic region between *yukF* and *yukE* (Fig. II.7A). Figure II.7B shows in detail the sequence between the *yukF* and *yukE* genes. The identified Rho-independent terminator is represented along with the two possible start codons for *yukF*. Shine-Dalgarno, -10 and -35 boxes were also predicted. The presence of putative upstream promoter (UP) elements, which are A+T rich regions that tend to modulate the strength of the downstream promoter sequences (Fredrick *et al.,* 1995; Dekhtyar *et al.,* 2008), is also indicated.

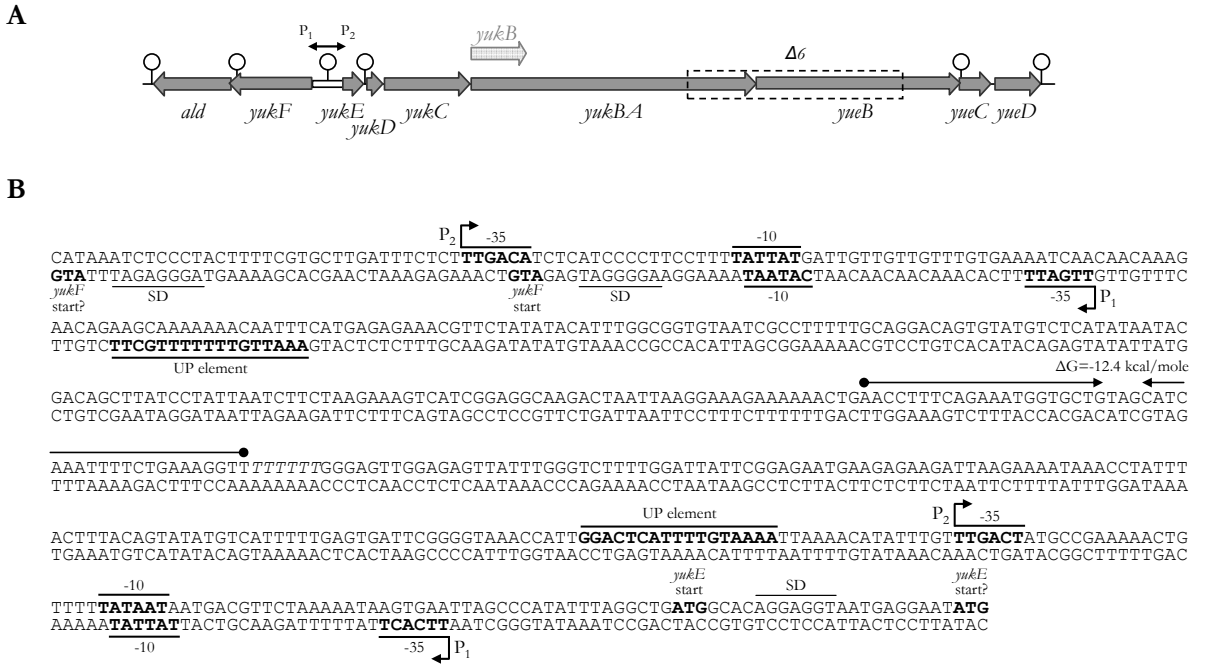


Figure II.7 | (**A**) Gene organization of the *B. subtilis* chromosome region that contains *yueB*. The DNA segment encompassed by the ⊿*6* deletion in CSJ1 strain is boxed. Rho-independent terminator structures identified by bioinformatic analysis (with ΔG ≤ -10.5 kcal/moles) are indicated by ' ⚲ '. The *yukF*/*yukE* intergenic region is represented by a white rectangle and the presence of putative SigmaA-dependent

divergent promoters is depicted as a double-headed arrow (P_1 and P_2 generically refer to all left and right oriented promoters, respectively; promoters predicted with the BPROM tool from SoftBerry). The previously assigned *yukB* gene is in light gray (see text for details). (**B**) DNA sequence of the *yukF*/*yukE* intergenic region with location of the putative sigmaA-dependent promoters, Shine-Dalgarno sequences (SD) and start codons. The putative -10 (consensus: TATAAT) and -35 (consensus: TTGACA) promoter sequences and the UP transcriptional element (consensus: AAAWWTWTTTTNNNAAA; W, A or T; N, any nucleotide) are indicated. Most likely alternative start codons are indicated as "start?" (see text). The identified Rho-independent terminator is represented as two convergent arrows with indication of the Gibbs free energy value and of the stretch of U's highlighted in italic.

The existence of divergent promoters in the region between *yukF* and *yukE* was verified by cloning it in a promoter probe vector (pRS415; Simons *et al.,* 1987). The region was cloned in the two orientations and high levels of β-galactosidase activity (8891.7 ± 3090.5 Miller Units for P_1 and 29664.9 ± 1597.4 MU for P_2) were measured in both *E. coli* constructed strains. The results confirmed the existence of two strong divergent promoters in this intergenic region (not shown).

During the analysis and search for similarities with the products of the genes mentioned above we noticed that genes *yukB* and *yukA* (annotated as two independent genes; Kunst *et al.,* 1997; Oudega *et al.,* 1997) could in fact correspond to a single open reading frame. This was suggested by the homology of *yukB* and *yukA* predicted products to the N- and C-terminal portions, respectively, of single proteins (e.g., the FtsK/SpoIIIE-like protein BH0975 of *Bacillus halodurans*; NP_241841). Sequencing of the *yukB* and *yukA* junction region revealed that a cytosine was missing in the *B. subtilis* genome sequence between positions 59287 and 59288, causing a frameshift and the occurrence of a premature stop codon. In this thesis and as published before, the complete gene resulting from the fusion of the previously reported *orf*s is designated *yukBA* (São-José *et al.,* 2004).

An error on the YukE coding sequence was also suggested by the same analysis. We found that the *yukE* product was consistently similar to only the first 50 amino-acids of any homologous protein (e.g., the YukE protein of *Bacillus amyloliquefaciens* FZB42; YP_001422458.1). Sequencing of the *yukE* gene exposed a missing cytosine in the database sequence between positions 61976 and 61977, originating an early stop codon. Also, the *yukE* start codon considered in the deposited genome sequence should not correspond to the correct one. An alignment with homologous proteins suggested that the third, in frame, ATG triplet should correspond to the start codon of *yukE*. In fact this ATG triplet is the only correctly positioned relatively to the identified SD consensus sequence (Fig. II.8). After correction of the frameshift which caused an early stop codon and alteration of the start codon, the YukE protein is predicted to have 97

amino acids, as is expected for a member of the ESAT-6/WXG superfamily (Fig. II.8 and see bellow). We also propose an alternative start codon for *yukF* (Fig. II.7) based on the position of the predicted -10 and -35 promoter sequences and presence of a more suitable SD sequence.

Figure II.8 | CLUSTAL W2 multiple sequence alignment of some YukE homologous proteins found in public databases. The corrected sequence proposed here is depicted as "yukE new" and the sequence present in databases is "B. subtilis Z99120". The *Staphylococcus aureus* EsxA protein and the *Streptococcus equi* Esx are known members of the ESAT-6 superfamily of proteins. The WXG conserved motif of the ESAT-6/WXG superfamily is boxed. The asterisks indicate conserved residues in all sequences. The single dot denotes conserved substitutions. The double dots indicate semi-conserved substitutions. Color codes: Red - small and hydrophobic (including aromatic -Y); Blue - acidic; Magenta - basic; Green - hydroxyl, amine, basic, Q; Gray - others. (CLUSTAL W version 2.0.10).

Bioinformatic analysis and similarity searches for each *orf* of the region surrounding *yueB* is summarized in Table II.3 and discussed below:

Ald. The *ald* gene codes for L-alanine dehydrogenase, an enzyme which catalyzes the NAD-dependent reversible reductive amination of pyruvate into alanine. The product of this gene plays an important role in the sporulation of *B. subtilis* (Siranosian *et al.,* 1993).

YukF. The product of *yukF* showed significant homology to proteins that function as transcript regulators. Although YukF does not seem to be the regulator of the *yukE* and downstream genes expression (not shown), the identification of three conserved domains specific of transcription regulators clearly points that way.

YukE. Searches on the databases indicated that this product would be a member of the ESAT-6/WXG-100 family of proteins (early secreted antigenic target of 6 kDa) that contain the Tryptophan-X-Glycine domain (W-X-G domain present in proteins of 100 amino acids). It has a high degree of similarity with EsxA from *Staphylococcus aureus*, a known ESAT-6/WXG-100 protein. Although they differ in their degree of sequence similarity (Fig. II.8), some rules are completely fulfilled by YukE to be a family member: i) it is a ~100 amino acid protein, ii) it has a W-X-G domain aligned with other known members of this family and iii) is clustered with genes that are also present in vicinity of the already described members of ESAT-6/WXG-100 family. ESAT-6/WXG-100 are small proteins that are secreted by a recently described secretion system, proposed as the Type VII secretion system (Abdallah *et al.,* 2007; see below). They play a fundamental role in virulence of *Mycobacterium tuberculosis* (Brodin *et al.,* 2004), haemolysis, conjugation and protective immunity of other Mycobacterium (see Abdallah *et al.,* 2007 for a review) and in the pathogenesis, dissemination and colonization of *S. aureus* (Burts *et al.,* 2005). The existence of these proteins in low G+C non-pathogenic bacteria such as *B. subtilis* and the relevance of the recently proposed type VII secretion system are addressed in Discussion (II.4).

YukD. The product of the *yukD* gene showed a weak sequence similarity with another member of the proposed *S. aureus* Ess-Esx secretion system, the EsaB protein (see Discussion). Besides that, the YukD protein has been described as a new type of prokaryotic ubiquitin-like proteins (Iyer *et al.,* 2006). In fact, it was recently crystallized and the authors showed that YukD and ubiquitin have the strongest resemblance among the single-domain ubiquitin-like proteins in the protein database (van den Ent and Löwe, 2005). Although it adopts a fold that is most closely related to ubiquitin, YukD has the shortest C-terminal tail of all known ubiquitin-like proteins.

YukC. The YukC protein is predicted to have one transmembrane domain and a membrane protein conserved domain. It also showed weak similarity with the putative membrane protein EssB of the aforementioned *S. aureus* secretion system.

YukBA. Sequence homology searches over the corrected amino acid sequence of the product of *yukBA* revealed a high probability of belonging to the FtsK/SpoIIIE family of proteins. It contains an ATPase conserved domain that, in the case of the homologue FtsK is involved in cell division in *E. coli* and, for the case of SpoIIIE in chromosome partitioning in *B. subtilis*. Another member of this family of proteins is the DNA segregation ATPase involved in the abovementioned Ess-Esx secretion system of *S. aureus*, the EssC protein, with which YukBA

shows a high degree of similarity. This protein in *S. aureus* was proved to be essential for the secretion of the ESAT-6 proteins, EsxA and EsxB.

YueB. YueB sequence homology searches also evidenced homology with another member of the *S. aureus* Ess-Esx secretion system, the membrane protein EsaA. In addition to that, YueB shares the same membrane topology with the Pip (phage infection protein) protein from *Lactococcus lactis* and with the product of *yhgE*, a *B. subtilis yueB* paralogue (São-José *et al.,* 2004). YueB-like proteins seem to be exclusive of gram-positive bacteria and are ubiquitously distributed with some bacteria having up to five copies of these genes (São-José *et al.,* 2004). Some of these copies (eg. *S. aureus* EsaA) have already been assigned as members of the new type VII secretion system, although they do not seem to be essential for the transport of ESAT-6/WXG-100 substrates.

YueC. The protein encoded by *yueC* showed no conserved domains and no sequence similarities with proteins present in the databases other than putative YueC-like proteins in *Bacillus sp.*

YueD. This putative protein shows sequence similarity with dehydrogenase/reductase proteins. Two conserved domains were detected related to the putative function, the PRK06924 short chain dehydrogenase domain and the Rossmann-fold NAD(P)(+)-binding proteins domain. Maruyama and colleagues described a benzyl reductase activity in *Bacillus cereus* (Maruyama *et al.,* 2002) attributed to locus BCE_G9241_3493, with which YueD shares a high degree of similarity.

Given the organization of the region containing *yueB* and upstream genes, the homologies found for some of the encoded products and the position of the putative promoter and terminator sequences, we raised the hypothesis that these genes could form an operon structure. The hypothesis that it may represent a novel secretion system potentially driven by the activity of the YukBA protein is presented in Discussion.

Table II.3 | Database sequence homologies and features of the *orf*s in the *yueB* genomic region.

Product	Gene length (bp)	Protein size (aa)	Mol. Mass (kDa)	Relevant similar proteins[1]	Conserved domains[2]	TMD's[3]	Putative function/feature[4]
YukF	1265	434	50.3	Putative transcriptional regulator (gi: 171325270) of the PucR family from *Geobacillus* sp. WCH70. E value: $5e^{-124}$; DNA-binding transcriptional regulator (gi:212640606, contains C-terminal Fis domain) from *Anoxybacillus flavithermus* WK1. E value: $2e^{-112}$; Regulator of polyketide synthase expression (gi: 56962540) from *Bacillus clausii* KSM-K16. E value: $1e^{-52}$.	COG2508, Regulator of polyketide synthase expression (Signal transduction mechanisms); COG3835, CdaR, Sugar diacid utilization regulator (Transcription/Signal transduction mechanisms); PRK11477, carbohydrate diacid transcriptional activator CdaR.	0	Transcription regulator.
YukE	294	97	11	ESAT-6/Esx family secreted protein EsxA/YukE (gi: 195977671) from *Streptococcus equi* subsp. *zooepidemicus* MGCS10565. E value: $1e^{-13}$; WXG domain-containing protein (gi: 116871469) from *Listeria welshimeri* serovar 6b str. SLCC5334. E value: $7e^{-13}$; Chain A of homodimeric *Staphylococcus aureus* EsxA (same as SAV0282, gi:197304995). E value: $1e^{-9}$.	COG4842, Uncharacterized protein conserved in bacteria; pfam06013, WXG100, Proteins of 100 residues with WXG motif.	0	Secreted protein; Chaperone of a ESX-1-like secretion system (proposed as Type VII secretion system).
YukD	240	79	9.1	YukD (gi:154687296) from *Bacillus amyloliquefaciens* FZB42. E value: $1e^{-33}$; Ubiqutin-like protein YukD (gi:52081687) from *Bacillus licheniformis* ATCC 14580. E value: $8e^{-32}$; Hypothetical protein EsaB (same as SAV0285, gi:15923275) from *Staphylococcus aureus*. E value: 0.18.	pfam08817, YukD, Bacterial protein that adopts an ubiquitin-like fold.	0	Crystal structure of YukD revealed a ubiquitin-like fold; Component of a ESX-1-like secretion system (proposed as Type VII secretion system).
YukC	1356	451	52.2	YukC (gi:154687295) from *Bacillus amyloliquefaciens* FZB42. E value: $2e^{-154}$; Hypothetical protein BL02621 (gi:52081686) from *Bacillus licheniformis* ATCC 14580. E value: $3e^{-133}$; Putative membrane protein EssB (same as SAV0286, gi:15923276) from *Staphylococcus aureus*. E value: $7e^{-17}$.	COG4499, Predicted membrane protein.	1	Membrane protein; Component of a ESX-1-like secretion system (proposed as Type VII secretion system).

Table II.3 | Database sequence homologies and feature searches of the ORFs in the *yueB* genomic region (cont.).

Product	Gene length (bp)	Protein size (aa)	Mol. Mass (kDa)	Relevant similar proteins[1]	Conserved domains[2]	TMD's[3]	Putative function/feature[4]
YukBA	4500	1499	172.2	FtsK/SpoIIIE family cell division protein YukA (gi: 157693610) from *Bacillus pumilus* SAFR-032. E value: 0; Cell division FtsK/SpoIIIE (gi: 196250239) from *Geobacillus* sp. G11MC16. E value: 0; DNA segregation ATPase and related proteins, EssC (same as SAV0287, gi:57634614) from *Staphylococcus aureus*. E value: 0.	pfam01580, FtsK/SpoIIIE, FtsK/SpoIIIE family; COG1674, FtsK, DNA segregation ATPase FtsK/SpoIIIE and related proteins.	2	DNA transport/ translocation; Component of a ESX-1-like secretion system (proposed as Type VII secretion system).
YueB	3231	1076	120	YueB (gi: 154687293) from *Bacillus amyloliquefaciens* FZB42. E value: 0; YhgE/Pip N-terminal domain protein (gi: 192811414) from *Geobacillus* sp. Y412MC10. E value: $1e^{-8}$; Putative protein EsaA (same as SAV0283, gi:15923273) from *Staphylococcus aureus*. E value: $8e^{-8}$.	COG1511, Predicted membrane protein; COG1196, Smc, Chromosome segregation ATPases. (only partial; E value: $2e^{-4}$)	6	Membrane protein; Component of a ESX-1-like secretion system (proposed as Type VII secretion system); Bacteriophage SPP1 irreversible adsorption receptor.
YueC	456	151	16.9	YueC (gi:154687292) from *Bacillus amyloliquefaciens* FZB42. E value: $7e^{-40}$; YueC (gi:52787069) from *Bacillus licheniformis* ATCC 14580. E value: $2e^{-21}$; Hypothetical protein BPUM_2855 (gi:157693608) from *Bacillus pumilus* SAFR-032. E value: $8e^{-15}$.	No putative conserved domains have been detected.	1	Unknown.
YueD	732	243	27.1	Short chain dehydrogenase (gi:154687291) from *Bacillus amyloliquefaciens* FZB42. E value: $1e^{-95}$; Short-chain dehydrogenase/reductase (gi:194015438) from *Bacillus pumilus* ATCC 7061. E value: $3e^{-64}$; Benzil reductase (gi:19386574) from *Bacillus cereus*. E value: $3e^{-49}$.	PRK06924, short chain dehydrogenase; Rossmann-fold NAD(P)(+)-binding proteins.	0	Benzil reductase.

[1]- Similarity searches performed with BLASTP 2.2.19+ and 2.2.18+ (Altschul *et al.,* 1997, 2005). The indicated proteins may be those that point for a putative function and not necessarily the most similar ones;
[2]- Conserved domain (CD) analysis with CDD (Marchler-Bauer *et al.,* 2004, 2005, 2007); tool available at http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi;
[3]- Transmenbrane domains predicted by TMHMM (available at http://www.cbs.dtu.dk/services/TMHMM/);
[4]- Putative function based on similarity and CD searches, membrane topology and literature.

II.3.2 – *YUEB* AND UPSTREAM GENES FORM AN OPERON: THE YUKE OPERON

II.3.2.1 – Transcription and functional analysis by the pMutin-4 strategy

We used the pMutin-4 integration vector (Vagner *et al.,* 1998) to study the possible operon organization of *yueB* and surrounding genes and at the same time to assess their possible role in SPP1 infection. This vector has three major properties that made it suitable for our purposes. First, upon integration, the 5' portion of the disrupted gene becomes transcriptionally fused to *spoVG-lacZ*, allowing quantification of its native expression. Second, strong transcriptional terminators disable read-through of the transcription initiated upstream the integration site. Finally, genes located downstream the inactivated gene can be expressed from the IPTG-inducible promoter P_{spac}, thus bypassing eventual polar effects resulting from the integration event. On one hand by disrupting the genes upstream *yueB* with this strategy we could test the operon organization in the absence of IPTG, by evaluating the direct effect in *yueB* expression and consequently in SPP1 infection. On the other hand, the IPTG driven expression of the genes located downstream the integration site allowed us to isolate and study the effect of each gene inactivation in SPP1 infection.

Individual disruption of genes *yukE* to *yueC* was accomplished as described in section II.2.5.1 and II.2.5.2. In none of the constructs, except for the one disrupting *yueC*, could SPP1 plate efficiently in the absence of IPTG (Fig. II.9). However, SPP1 efficiency of plating (E.o.P) was restored to wild type levels in the presence of IPTG, except for the strain carrying integration in *yueB*. This clearly indicates that *yueB* is the sole gene in the examined cluster that is indispensable for plaque formation. The fact that disrupting the *yueB* upstream genes caused a polar effect on SPP1 efficiency of plating ascertains that at least genes from *yukE* to *yueB* are organized as an operon. Measurements of β-gal activity in the different constructs indicated a high level of expression of gene *yukE* compared to the remaining genes (Fig. II.9).

Analysis by western-blot of YueB production by the integration mutants in the absence of IPTG confirmed that the deficiency in SPP1 plating correlated with a severe decrease in YueB production (Fig. II.10). Addition of IPTG drove *yueB* production from the P_{spac} promoter and, consequently, restored SPP1 ability to plate. As noted before (São-José *et al*., 2004), frequently the only YueB polypeptides that are clearly detected are those that probably accumulate after proteolytic cleavage of the full length, 120 kDa, YueB. These are detected as two to four bands

ranging from 60 to 70 kDa. As expected, the IPTG dependence of YueB production in the integrants upstream of *yueB* produced a direct impact on SPP1 irreversible adsorption (section II.3.3). In the western blot presented in Figure II.10 a ~60 kDa band is visible in almost all protein extracts, independently of IPTG induction, including the CSJ1 extract (a spontaneous resistant mutant strains that carries a deletion of *yueB*; see Table II.1 and Fig. II.7A). This band corresponds to a polypeptide that is reactive with the YueB antibodies although not being YueB. The presence/detection of this protein was stochastic, varying with each individual extract production.

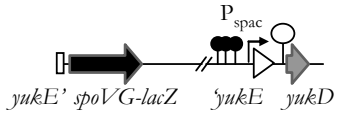
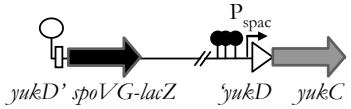
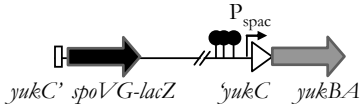
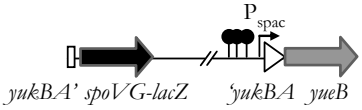
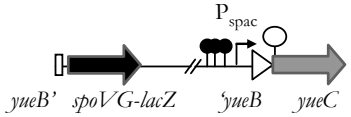
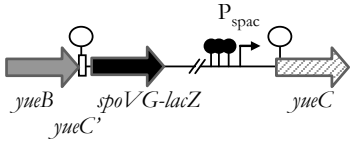
Strain	DNA structure	E.o.P		β-gal (Miller Units)
		- IPTG	+ IPTG	- IPTG
L16601		1	1	0.45 ± 0.04
CBM6: *yukE*ΩpCBM6		(*)	0.64	277.97 ± 22.58
CBM7: *yukD*ΩpCBM7		< 3.5 x 10^{-5}	0.67	4.44 ± 0.47
CBM8: *yukC*ΩpCBM8		< 3.5 x 10^{-5}	0.76	3.83 ± 0.36
CBM9: *yukBA*ΩpCBM9		< 3.5 x 10^{-5}	0.75	6.79 ± 0.56
CBM10: *yueB*ΩpCBM10		< 3.5 x 10^{-5}	< 3.5 x 10^{-5}	2.46 ± 0.22
CBM11: *yueC*ΩpCBM11		0.96	0.94	3.13 ± 0.25

Figure II.9 | SPP1 efficiency of plating (E.o.P.)and β-galactosidase activity measured in the wild type strain (L16601) and in the different integration mutants in the absence or presence of 1mM IPTG. A schematic representation of the relevant DNA structure of each integrant is provided (genes not drawn to scale). Disrupted *orf*s are represented by interrupted white arrows while intact genes are depicted as gray arrows. The reconstituted copy of *yueC* in the conditional mutant is represented as a dashed gray arrow. pMutin4-derived elements (*spoVG-lacZ*, transcriptional terminators and P_{spac} promoter) are in black.

Symbol ' ⚲ ' is as in Fig. II.7. The indicated β-gal activity values are the average from at least three independent experiments ± standard-deviation of the mean. E.o.P values are expressed as the ratio between the phage titer obtained in each mutant strain and that obtained in the wild type strain (L16601). (*) Minute phage plaques could be observed suggesting an E.o.P almost identical to that observed in the presence of IPTG, but their reduced size did not allow accurate plaque counting.

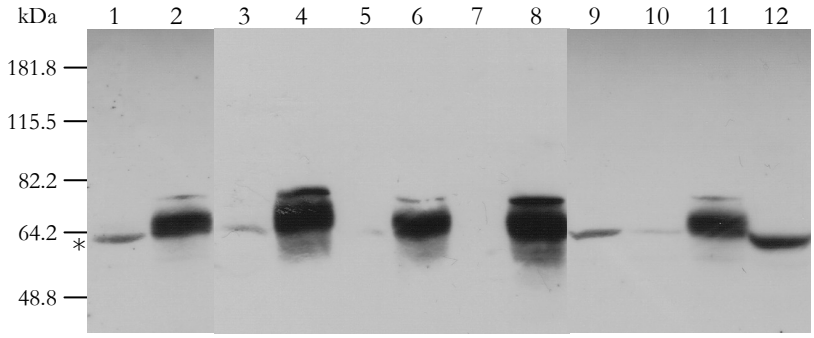


Figure II.10 | Western blot analysis of YueB polypeptides in the protein extracts obtained from the different *B. subtilis* integrant strains grown in the absence or in the presence of the inducer. Strains are designated accordingly to Figure II.9. Lanes: 1, CBM6; 2, CBM6+IPTG; 3, CBM7; 4, CBM7+IPTG; 5, CBM8; 6, CBM8+IPTG; 7, CBM9; 8, CBM9+IPTG; 9, CBM10; 10, CBM10+IPTG; 11, L16601; 12, CSJ1. L16601 is the wild type strain; CSJ1 is a *yueB* non-producing strain. Asterisk denotes a non-specific reacting band (see text for details).

II.3.2.2 – Transcription analysis by Northern-blot and RT-PCR

As referred to above, β-galactosidase activity values measured in the mutant strains revealed a much higher expression of the first gene of the operon (*yukE*) as compared to the others (Fig. II.9). Northern-blot analysis using a probe internal to *yukE* evidenced a major transcript with an estimated size of ~600 nucleotides. The same analysis with probes further down the operon only detected a smear, starting at the top of the gels, suggesting degradation of high-molecular mass mRNA species (Fig. II.11A). Existence of a polycistronic mRNA was evidenced by Reverse Transcriptase-PRC (RT-PCR). A cDNA generated with primer BS-9 (complementary to the 3' end of *yueB*) was used as template for a PCR reaction with primer pairs yukE-1/BS-9 and yukE-1/yukC-2. As can be seen in Figure II.11B, a ~10 kbp PCR product, covering the five genes of the operon, could be amplified from the BS-9 cDNA with primers yukE/BS-9 (lane 4). The size of this product is compatible with the size of the amplicon obtained with the same primers but using L16601 DNA as template (lane 3; the apparent smaller size of this product may be due to electrophoretic distortion caused by an higher load of DNA). Amplification of a smaller product corresponding to genes *yukE*, *yukD* and the 5' half of *yukC* was also obtained using the same

cDNA as template and primers yukE-1/yukC-2. These amplifications confirm that polycistronic mRNA species exist, although with relative instability.

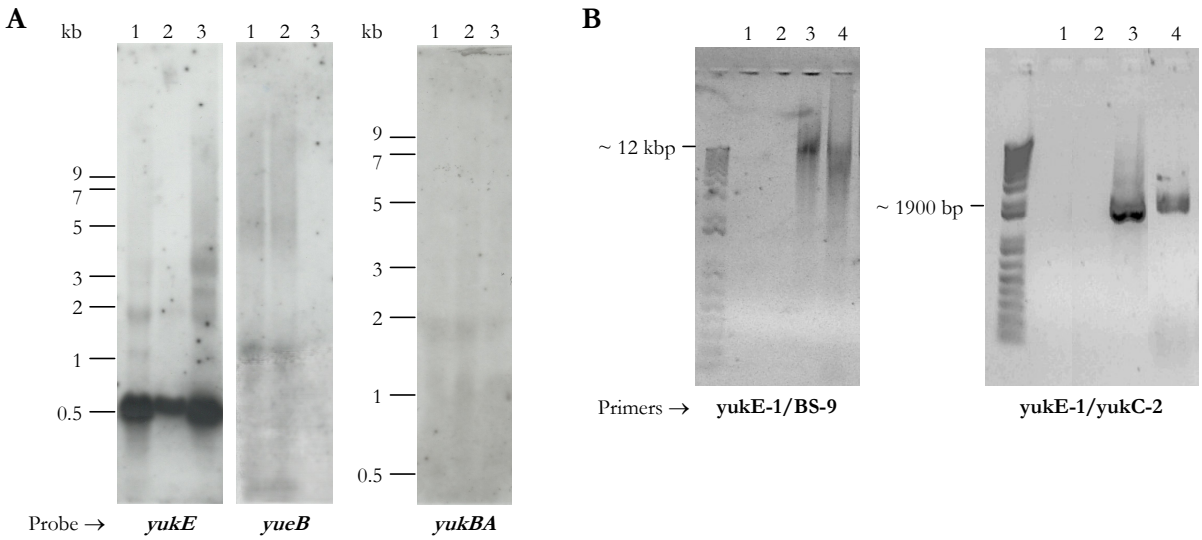


Figure II.11 | (**A**) Northern blot analysis of RNA extracted from *B. subtilis* strains. Lanes: 1, L16601 (wild type); 2, CSJ2 (*pha-2* strain); 3, CSJ1 (⊿*yueB*). Radiolabeled probes correspond to internal PCR products of the indicated genes. The *yueB* probe is included in the DNA segment deleted in CSJ1. (**B**) RT-PCR analysis of a cDNA generated with primer BS-9 and using L16601 extracted RNA. Two distinct PCR amplifications were performed with the same cDNA (primers pair yukE-1/BS-9 and yukE-1/yukC-2). Lanes: 1, negative control PCR without template; 2, negative control PCR with RNA as template; 3, positive control PCR with L16601 DNA as template; 4, PCR reaction with generated cDNA as template.

II.3.3 –SPP1 PLAQUE MORPHOLOGY IN THE YUKE OPERON INTEGRATION MUTANTS

Even though SPP1 could plate efficiently in all but *yueB* integrants in the presence of IPTG, phage plaques were smaller than those obtained in the wild type strain, a phenotype that was particularly noticeable when *yukE* was the disrupted gene (Fig. II.12). The smaller plaques could be explained if, in addition to *yueB*, other genes of the operon played a role in SPP1 binding, namely in irreversible adsorption. However, when irreversible adsorption constants (k_{ads}) were evaluated in liquid medium supplemented with IPTG, all knockouts in genes upstream *yueB* resulted in similar k_{ads} values, about 2-3 fold higher than that obtained for the wild-type strain (Fig. II.12). Smaller plaques in these cases are thus not correlated with a decrease in irreversible binding.

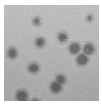
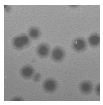
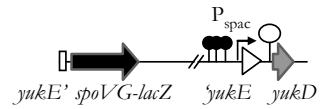
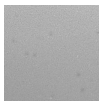
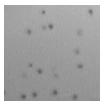
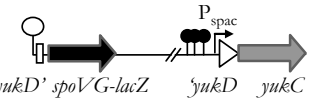
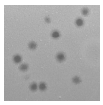
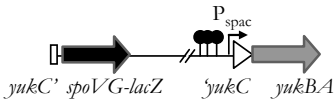
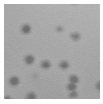
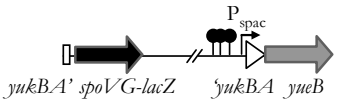
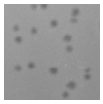
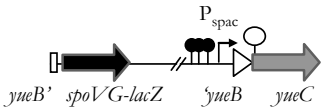
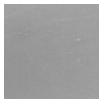
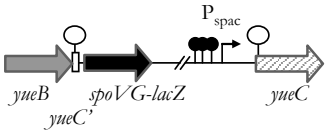
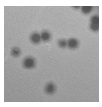
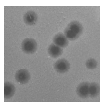
Strain	DNA structure	Plaque morphology		k_{ads}($min^{-1}.A_{600}^{-1}$)	
		- IPTG	+ IPTG	- IPTG	+ IPTG
L16601				1.21 ± 0.17	1.07 ± 0.14
CBM6: *yukE*ΩpCBM6				≤ 0.01	2.28 ± 0.52
CBM7: *yukD*ΩpCBM7				≤ 0.01	2.93 ± 0.78
CBM8: *yukC*ΩpCBM8				≤ 0.01	2.00 ± 0.51
CBM9: *yukBA*ΩpCBM9				≤ 0.01	2.88 ± 0.32
CBM10: *yueB*ΩpCBM10				≤ 0.01	≤ 0.01
CBM11: *yueC*ΩpCBM11				1.11 ± 0.18	0.95 ± 0.27

Figure II.12 | SPP1 plaque morphology and k_{ads} values obtained in the wild type strain (L16601) and in the different integration mutants in the absence or presence of 1 mM IPTG. A schematic representation of the relevant DNA structure of each integrant is provided (as in Fig. II.9). The indicated k_{ads} values are the average from at least three independent experiments ± standard-deviation of the mean, except under conditions where no accurate k_{ads} measurements (≤ 0.01) could be performed.

It is known that phage plaque size can be influenced by the host metabolism and growth rate. We could not find however, at least in liquid medium, significant differences in the mutant strains growth rates that could explain the plaque size phenotype (Fig. II.13). Furthermore, the unrelated phage SP01 formed undistinguishable plaques in all integrants, both in IPTG-supplemented or non-supplemented medium (not shown). The reported effect on plaque size is thus specific for SPP1.

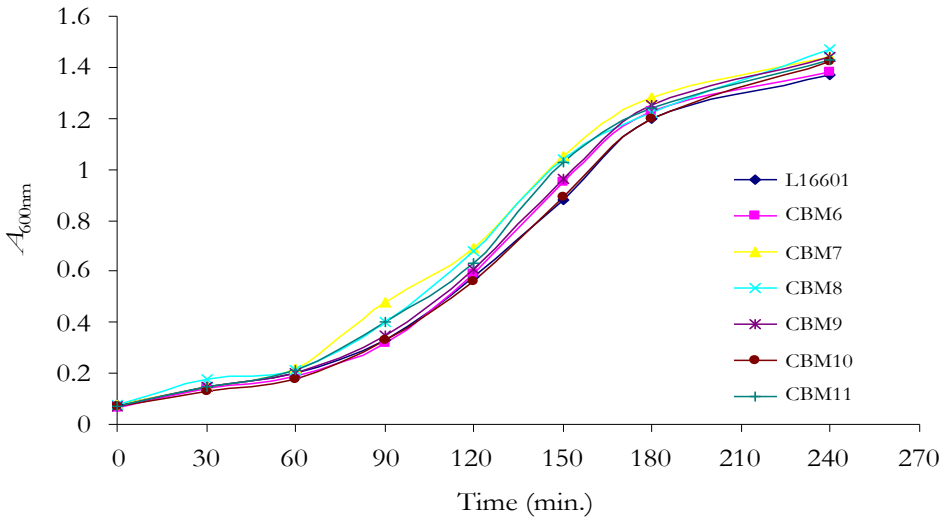


Figure II.13 | Growth curves of wild type strain and of the integration mutants of the yukE operon.

II.3.3.1 – The minute phage plaque enigma: *yukE* and *yukD* conditional mutants

The results presented in the previous section seemed to rule out a role of the operon genes upstream of *yueB* in SPP1 irreversible adsorption. The product of these genes could however be important for an infection step occurring downstream of irreversible binding, like for example in SPP1 DNA penetration into host cells. The perturbation of this step(s) could be on the basis of the observed small plaque phenotype. In this scenario, the *yukE* product would be of major importance since its disruption led to the most significant decrease of SPP1 plaque size. Hence, we tried next to ascertain if the lack of YukE was in fact responsible for the observed phenotype. It should be noted that minute plaques have also been observed in the absence of inducer in the *yukE* disrupted strain suggesting the presence of a weak promoter located upstream of *yukD*. The idea of this promoter became more substantial after analyzing the *yukE-yukD* region (Fig. II.14). A putative promoter could be found downstream of the Rho-independent terminator of *yukE*, which might account for the initiation of transcription downstream of the integration point in strain CBM6. We decided to address these questions by constructing conditional mutants of *yukE* and *yukD*. In these cases, instead of disrupting the gene and annul its expression, their coding sequence and translation signals are placed under the control of P_{spac}, meaning that we have an IPTG-dependent expression of the target and downstream genes. For the *yukE* conditional mutant this implies that transcription of the entire operon is inducible (section II.2.5.1).

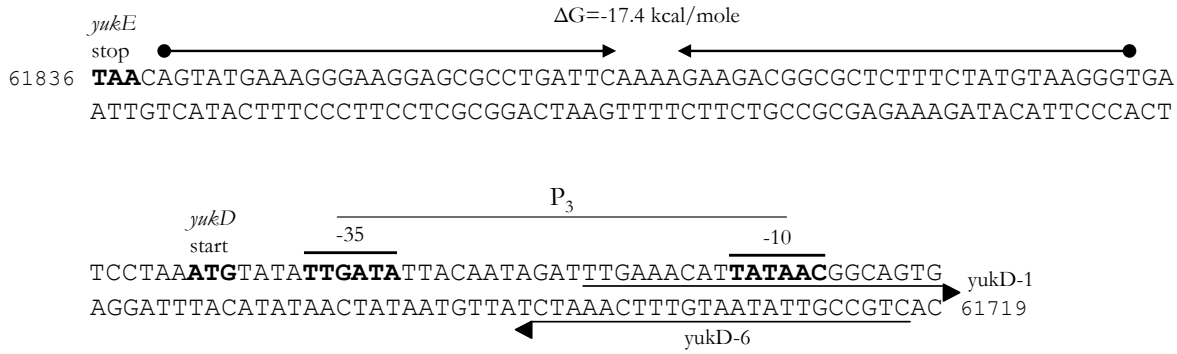


Figure II.14 | Schematic representation of the putative secondary promoter, P_3. The putative -10 and -35 consensus regions are represented. The identified rho-independent terminator is represented as two convergent arrows and the Gibbs free energy value obtained at the DBTBS is indicated. *yukE* stop codon and *yukD* start codon are in bold face. Hybridization position of primers yukD-1 and yukD-6 is also represented.

Somehow surprisingly, SPP1 produced minute phage plaques in the absence of the inducer, in both conditional mutants (Fig. II.15). Two possible, non-exclusive explanations fit this result. First, if we compare the genome organization after the integration events, we see that in all mutants that show minute plaques in the absence of IPTG (the two conditional mutants and the *yukE* integrant CBM6), the secondary structure that is predicted between *yukE* and *yukD* is always present downstream of P_{spac}. It is well known that secondary structures in the 5' end of transcripts often tend to stabilize them (Condon, 2003; 2007; Kozak, 2005), and this could be the case if we consider some leakiness of the P_{spac} promoter. Even in the absence of inducer, the small amount of transcript produced could be stabilized by its 5' end secondary structure, allowing some expression of YueB. In the cases where this structure is not present (i.e., non-conditional integration mutants of *yukD*, *yukC* and *yukBA*) the transcripts resulting from the P_{spac} leakiness might be rapidly degraded and no sufficient YueB is produced to allow infection. A second possible explanation arises if we assume that there is a secondary promoter (P_3; Fig. II.14) upstream of *yukD*. As schematized in Figure II.15, this putative weak promoter is located downstream of the integration site in the *yukE* and *yukD* conditional mutants (and also in the non-conditional *yukE* mutant, see above) and could be accounting for the minute plaques observed. In the integrations downstream of this promoter region (as is the non-conditional *yukD*, *yukC* and *yukBA*) transcription initiated from it stops at the pMutin-4 terminators.

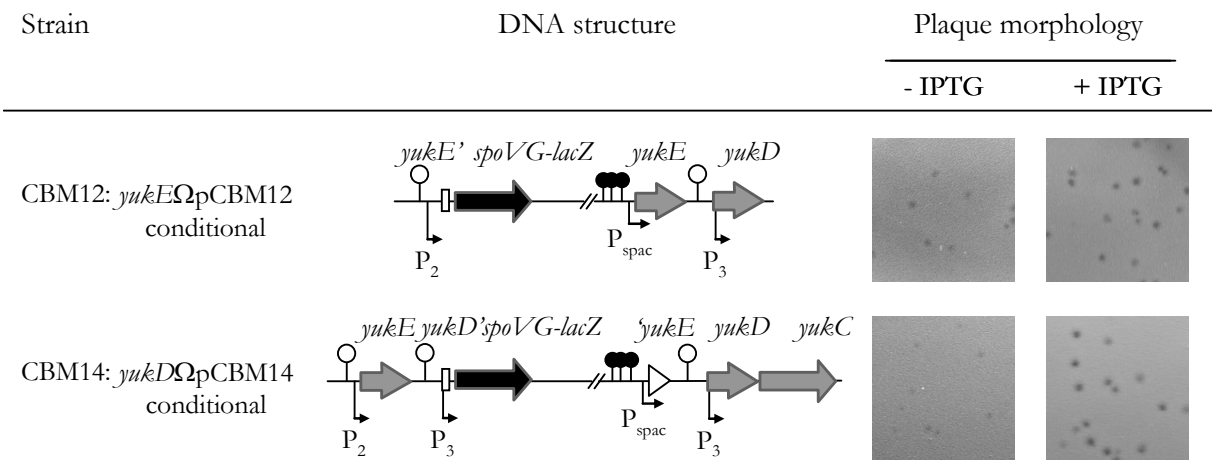


Figure II.15 | SPP1 plaque morphology in the *yukE* and *yukD* conditional mutants in the absence or presence of 1mM IPTG. A schematic representation of the relevant DNA structure of each integrant is provided (as in Fig. II.9). Predicted positions of promoters P_2 and P_3 are shown.

Another surprising result was that SPP1 plated equally in both conditional mutants in the presence of the inducer (Fig. II.15), maintaining the small plaque phenotype as observed previously for the non-conditional *yukE* mutant. This obviously means that the plaque phenotype observed can not be explained by the absence of the *yukE* and *yukD* gene products. Instead, the justification for this phenotype may rather rely on the altered transcription levels of the operon, and consequently on the stoichiometric relation of its products (see Discussion).

In conclusion, in one hand we speculate that in the integrants allowing formation of minute plaques in the absence of IPTG there is a certain amount of YueB produced, which does not occur in the others integrants. Unfortunately, such increased level of YueB in the absence of inducer is not obvious in western-blot analysis (Fig. II.10). On the other hand, the strains where expression of the entire operon or of the genes *yukD* to *yueB* is driven from the P_{spac} promoter only allow formation of very small SPP1 plaques.

II.3.3.2 – Ectopic expression of *yukE* does not complement the small plaque phenotype

Although the results presented above have indicated that the observed plaque phenotype could not be explained by the absence of the *yukE* product we decided to complement the mutation and see if we could restore the normal plaque morphology. For that we inserted a *yukE* copy in the *amyE* locus of strain CBM6 (see section II.2.5.3 for details), under the control of the xylose-inducible promoter P_{xylA}.

In none of the tested strains could SPP1 plate efficiently in the presence of both inducers, IPTG and xylose. Instead, the constructed strains behaved exactly as the parental strain, CBM6 (not shown). Thus, the results further indicated that the absence of the *yukE* product is not what is causing the small plaque phenotype.

II.3.4 – SPP1 PRODUCES WILD TYPE PLAQUE MORPHOLOGY IN *YUKE* AND *YUKD* NONSENSE MUTANTS

To further support the hypothesis that it is the altered genomic structure of the operon and corresponding expression profile that is responsible for the small plaque phenotype, we engineered strains carrying nonsense mutations in *yukE* and *yukD* genes. This caused silencing of the expression of these genes keeping intact the natural transcription and DNA structure of the operon (see section II.2.5.4). The final constructs were sequenced (Table II.4) and tested for SPP1 plaque morphology and efficiency of plating (Fig. II.16).

Table II.4 | *yukE* and *yukD* relevant sequences in the constructed strains L.cat86, L.E1E2 and L.D1D2.

	***yukE* (5'→3')**[a, b, c, d, e]	***yukD* (5'→3')** [a, b, c, d, e]
L.cat86 (WT)	atgGCAGGATTAATTCGTGTC ACACCCGAAGAGCTAAGAGCG	atgTATATTGATATTACAATAGAT TTGAAACATTATAACGGCAGTGTC TTTGATCTCAGATTGTCAGATTACCACCCGG
L.E1E2 (*yukE*-ochre)	atgGCAGGATTAATTCGTGTC ACACCC**TAA**GCTTTAAGAGCG	WT
L.D1D2 (*yukD*-ochre)	WT	atgTATATTGATATTACAATAGAT TTGAAACATTATAACGGCAGTGTC TTTGATCTCAGATTGT**TAA**GCTTACCACCCGG

[a] – Lowercase letters indicate the start codon of each gene;
[b] – Underlined depict the mutated nucleotides in each construct;
[c] – Highlight letters designate the engineered *Hin*dIII restriction site;
[d] – Boldface letters indicate the engineered stop codon;
[e] – WT = wild type.

Strain L.cat86, which is isogenic to L.E1E2 and L.D1D2 but with wild type *yukE* and *yukD*, behaves exactly as the wild type L16601 strain regarding growth and SPP1 E.o.P., plaque size and irreversible adsorption (not shown). The strains carrying the nonsense mutations in *yukE* and *yukD* also behaved as the wild type strain. SPP1 produced wild type plaques in these mutants and the E.o.P. was also indistinguishable (Fig. II.16). These results further indicate that *yukE* and *yukD* are not required for normal SPP1 plating.

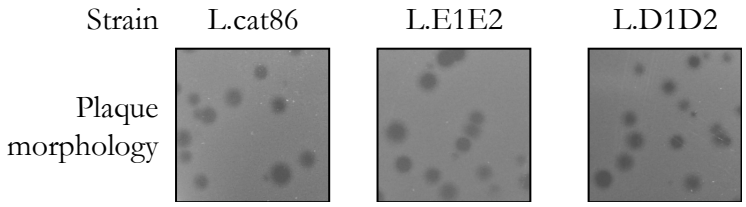


Figure II.16 | SPP1 plaque morphology in the *yukE* and *yukD* nonsense mutants (L.E1E2 and L.D1D2). Strain L.cat86 is used as the control.

II.3.5 – *YUEB* IS THE UNIQUE ESSENTIAL GENE FOR SPP1 INFECTION

With the aim of elucidating any preponderant role of some operon products besides YueB in SPP1 infection we decided to delete the entire yukE operon, including *yueB*, and try to restore phage sensitivity with an ectopic expression of *yueB*. For that, strains L.Del6 and L.Del6-Pxyl-YueB were constructed (section II.2.5.5). Strain L.Del6 has a deletion spanning from the 5' end of *yukF* to the 3' end of *yueB*, while L.Del6-Pxyl-YueB is a L.Del6 derivative with a *yueB* copy inserted in the *amyE* locus under inducible expression from the P_{xylA} promoter. As expected strain L.Del6 is resistant to SPP1 since it does not carry *yueB*. The ectopic expression of *yueB* restores the sensitivity to SPP1 infection and allows formation of plaques of wild type morphology (Fig. II.17A). The absence of YueB production in strain L.Del6 and its synthesis in L.Del6-Pxyl-YueB in the presence of xylose was confirmed by western blot (Fig. II.17B). Therefore, the only gene whose expression is essential for normal SPP1 infection and plating is indeed *yueB*.

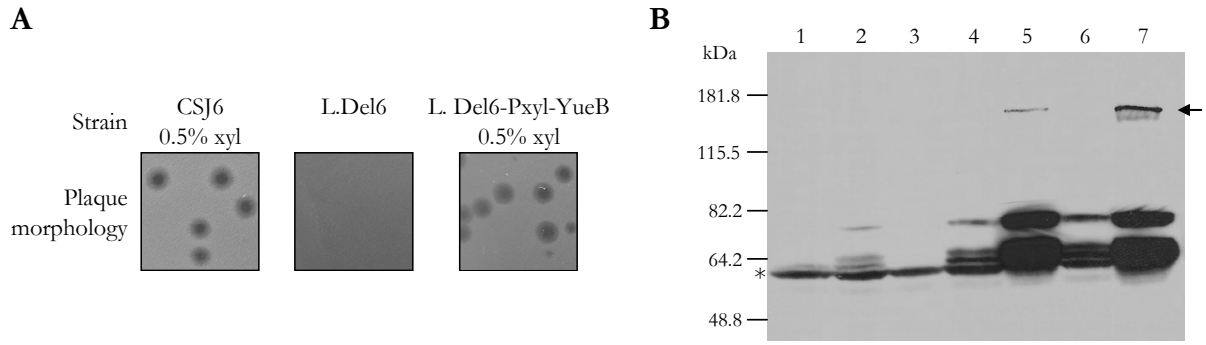


Figure II.17 | (**A**) SPP1 plaque morphology. Strain L.Del6 is resistant to phage infection. CSJ6 serves as a control strain for L.Del6-Pxyl-YueB as it is a CSJ1 derivative with *yueB* inserted in the *amyE* locus. (**B**) Western blot analysis of YueB in protein extracts obtained from strains L16601 (lane 1), L.cat86 (lane 2), L.Del6 (lane 3), CSJ6 (lane 4), CSJ6+xylose (lane 5), L.Del6-Pxyl-YueB (lane 6) and L.Del6-Pxyl-YueB+xylose (lane 7). Asterisk denotes the non-specific anti-YueB reacting band referred to above, as in Fig. II.10. When YueB is overexpressed from the P_{xylA} promoter it is possible to detect a polypeptide that corresponds to the full length YueB (arrow).

II.3.6 – CELLULAR LOCALIZATION OF YUEB BY GFP FUSION

The YueB topology prediction indicated the presence of one N-terminal and five C-terminal transmembrane domains (TMD). The product of the *yueB* gene is expected to be anchored in the membrane at least by its five C-terminal TMDs whereas the N-terminal domain may alternatively function as a signal peptide. Previously (São-José *et al.,* 2004) we raised the hypothesis that the products of the yukE operon might form a membrane complex implicated in SPP1 DNA translocation into the host cell. At the same time this hypothesis was inspired by reports claiming that YukBA-like proteins in *Mycobacteria* were the ATPase components of a new membrane transport system (Pallen, 2002, and references therein). Back then, of course, we were not aware that in fact *yueB*, at least when overexpressed, is sufficient to support normal SPP1 infection. Anyway, this hypothesis would imply that YueB co-localized with YukBA. To accomplish their function the FtsK-like proteins locate specifically at the cell middle region (FtsK) or at the cell poles (SpoIIIE) (Yu *et al*., 1998; Wu and Errington, 1997). To test if YueB had a specific localization at the *B. subtilis* cellular surface, we constructed a *yueB-gfp* translational fusion and placed it under the control of three different promoters: i) the native *yueB* promoter(s) (strain CBM-GFP), ii) the IPTG inducible promoter P_{spac} (strain CBM-429) and iii) the xylose inducible promoter P_{xylA} (strain CBM-191). These strains are derivatives of strains L16601, CSJ3 and CSJ6, respectively (see section II.2.5.6 for construction details).

The constructed strains were initially challenged with SPP1 infection and they proved to behave exactly as the parental regarding the efficiency of plating and plaque morphology. In order to confirm that the YueB-GFP fusion was being produced, total protein extracts of the three strains were analysed by western blot using anti-YueB and anti-GFP sera. The detected polypeptides corresponding to the full length and to the C-terminal cleavage products presented the expected molecular mass increment of ~30 kDa (due to the fusion with GFP) relative to the native YueB (compare the detection profiles of the parental strains with those of the corresponding derivative strains, Fig. II.18A). The predominant C-terminal cleavage product is also detected in the anti-GFP western blot, together with a smaller product that is not revealed by the anti-YueB serum (Fig. II.18B). This serum probably fails to detect this small product because it was raised against a truncated YueB that only includes the amino acids from positions 30-797 (São-José *et al*., 2004), and therefore products resulting from cleavage events downstream this position will not be detected.

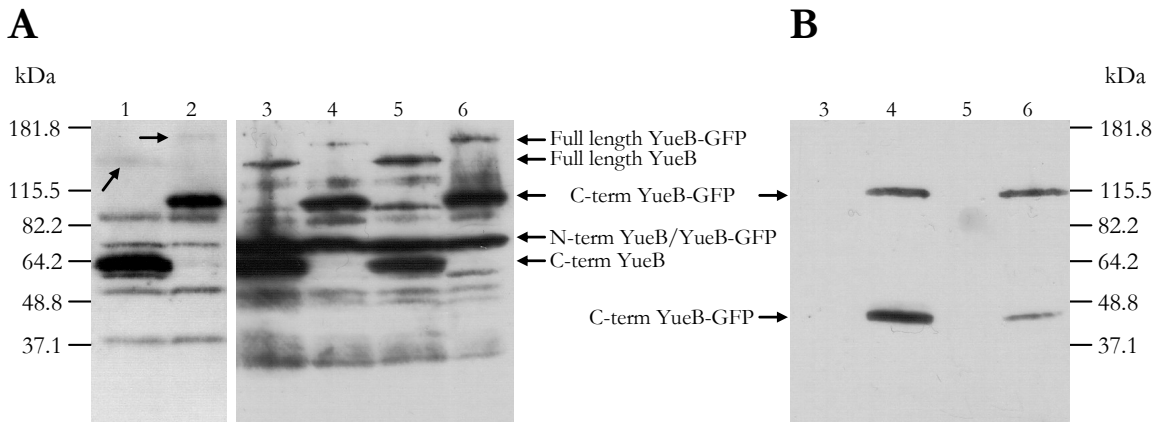


Figure II.18 | Western blot analysis of YueB and YueB-GFP polypeptides in protein extracts obtained from strains L16601 (lane 1), CBM-GFP (lane 2), CSJ3 (lane 3), CBM-429 (lane 4), CSJ6 (lane 5) and CBM-191 (lane 6). The proteins were immunodetected using polyclonal anti-YueB780 (**A**) and anti–GFP (**B**) antibodies. Arrows in lanes 1 and 2 point the full length YueB and YueB-GFP polypeptides, respectively, which are often difficult to detect in western blot (São-José *et al.,* 2004). Lanes 1 and 2 are overexposed comparatively to the remaining.

After confirmation of YueB-GFP production, we next analyzed its cellular localization by fluorescence microscopy (Fig. II.19). This work was carried out by Dr. Gonçalo Real in Dr. Adriano Henriques' lab at the ITQB, UNL, Lisbon. As expected, the YueB-GFP fusion localizes at the cell envelope. Overall, the analysis of the fluorescence patterns in the studied strains indicated that the SPP1 receptor is basically present all over the cell envelope (Fig. II.19A). A higher fluorescence signal is observed in the strains carrying an inducible expression of the fusion (stains CBM-429 and CBM-191). We should note however that the observed GFP signal indicates the localization of the YueB C-terminal product, which may or may not co-localize with the N-terminal portion. Interestingly, the deconvolution analysis suggested that, at least in some cells, the GFP fusion localizes preferentially at the cell poles (Fig. II.19B, sub-panel b). A *B. subtilis* strain expressing at the same time the native YueB and the YueB-GFP fusion (CBM-L191) gave identical results (data not shown).

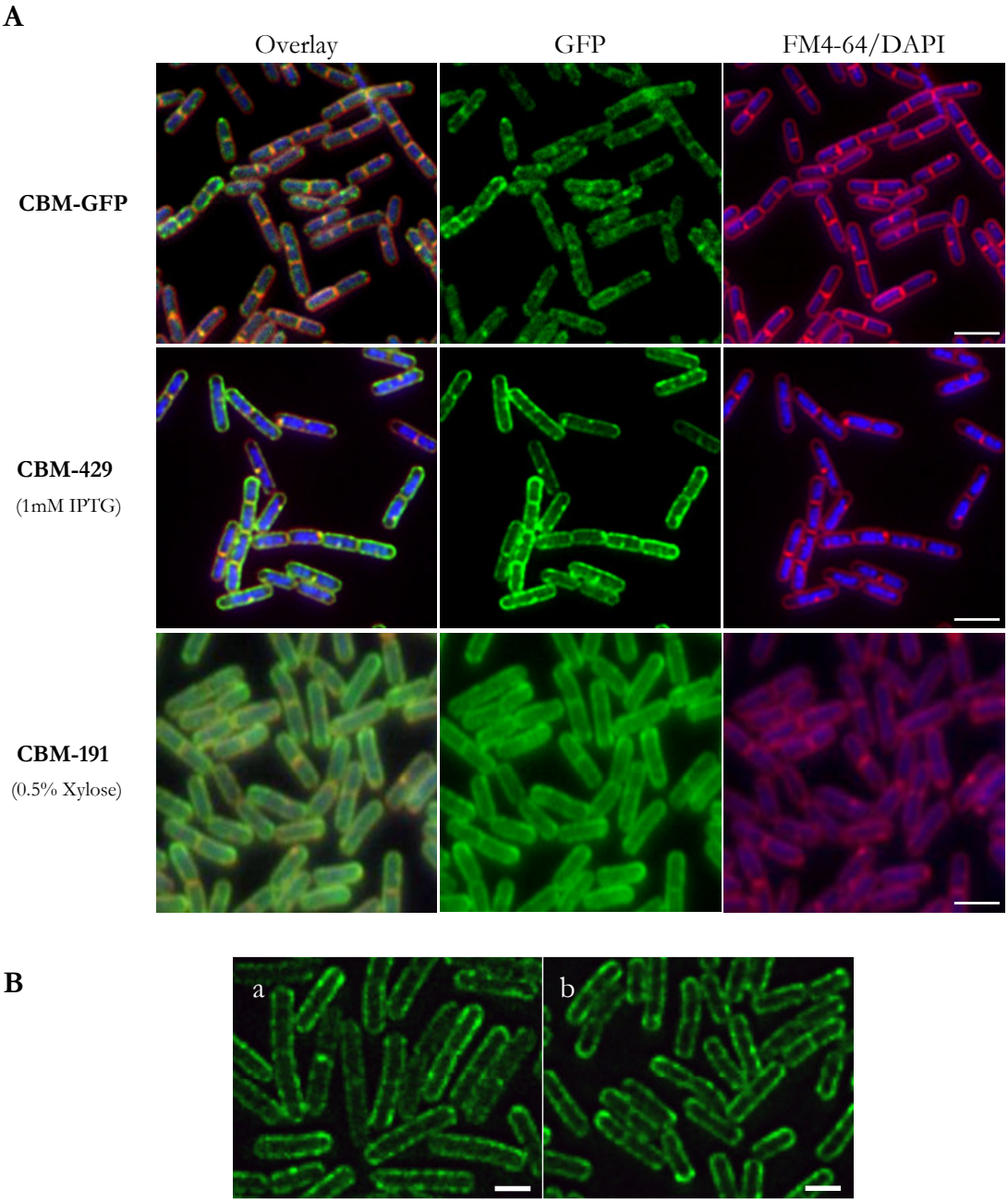


Figure II.19 | Localization of the YueB-GFP functional protein in *B. subtilis*. (**A**) Strains expressing the YueB-GFP fusion were grown in LB medium in the absence or in the presence of xylose (0.5 %) or IPTG (1 mM). Samples (600 µl) were collected during exponential growth (A_{600}=0.3) and prepared as previously described (Real *et al.,* 2005). Localization of YueB-GFP is shown in the central column. The rigth column shows the nucleoids (in blue) and the membranes (in red) stained with 4,6-diamidino-2-phenylindole dihydrochloride (DAPI) at 5 µg.ml^{-1} and with FM4-64 at 10 µg.ml^{-1}, respectively. A merge between membranes, nucleoids and GFP is shown in the left column. Microscopy was made on a Leica DMRA2 optical sectioning microscope coupled with a CoolSNAPTM HQ Phoptometrics camera (Roper Scientific, Tucson, Arizona). Scale bar, 2 µm. (**B**) Deconvolved images showing the localization of YueB-GFP in cells of strain CBM-GFP (a) and strain CBM-191 (b) collected during exponential growth in LB medium. For deconvolution analysis, 12 optical sections were collected from each field, and the images were deconvolved using the adaptive blind method with 20 interations of the Leica Deblur 1.03 software. Scale bar, 1 µm. The photographs were taken by Gonçalo Real at ITQB, UNL.

II.4. DISCUSSION

The work presented in this chapter derived from our search on the bacterial cell receptor responsible for irreversible adsorption of phage SPP1. Based on previous results with *B. subtilis pha-2* mutants, we were able to narrow this search to proteins encoded by genes in the vicinity of *ald*. Given its similarity with the *L. lactis* phage receptor Pip, we selected YueB as the most likely SPP1 receptor, a role that was confirmed experimentally by our group (São-José *et al.,* 2004). The YueB protein is an integral membrane protein and spans the entire cell envelope to expose a region recognized by SPP1. The interaction with this receptor motif triggers phage genome ejection (São-José *et al.,* 2006).

The analysis of the *yueB* genomic context suggested that it could be part of an operon structure in conjunction with the upstream genes *yukE*, *yukD*, *yukC* and *yukBA*. Several studies presented in this chapter support this operon organization. A transcriptional analysis by RT-PCR evidenced the presence of a large mRNA corresponding to the transcription of genes *yukE* to *yueB*. Northern-blot experiments indicated that this transcript is unstable and subjected to extensive processing. A small transcript encompassing the *yukE* gene produced a strong signal in Northern-blots. Inspection of the region between *yukE* and *yukD* showed an inverted-repeat sequence that could be a Rho-independent terminator, as suggested by bioinformatics analysis. However, predictions can highlight false-positive terminator structures that indeed may play different functions (de Hoon *et al.,* 2005). Based on our transcription analysis and molecular data we believe that this structure may attenuate and limit read-through of transcription initiated at the *yukF*/*yukE* intergenic region. Moreover, the secondary structures flanking *yukE* can also be involved in its transcript stabilization as shown in other *B. subtilis* operons (Ludwig *et al.,* 2001; Condon, 2007).

In another approach to study the transcriptional pattern of the yukE operon and, at the same time, to assess its possible role in SPP1 infection we have employed the pMutin-4 strategy to individually disrupt each gene of the operon. All integration mutants displayed phage resistance in the absence of IPTG, indicating that *yueB* transcription depends on the promoter sequences localized upstream *yukE*, confirming therefore the operon organization. Measurements of β-galactosidase activity agreed with a higher level of expression of the *yukE* gene when compared to the other members of the operon. Restoration of phage sensitivity was achieved in the

presence of IPTG for all the integrants, except for that disrupting *yueB*. The product of *yueB* was thus considered to be the only gene of the operon essential for SPP1 infection.

II.4.1 – THE YUKE OPERON AND SPP1 INFECTION

Regarding the other members of the operon, despite not being determinant for SPP1 infection, their disruption caused a contradictory effect on SPP1 plaque phenotype and irreversible adsorption constants (k_{ads}) (Fig. II.12). In the presence of the inducer that warranted YueB production at levels greater than the wild type, the integrants showed higher values of k_{ads} but smaller plaques. We know from our studies (São-José *et al.,* 2004) that increments on the native amount of YueB present in one bacterial cell has a direct and corresponding effect on the irreversible adsorption constant value. So, overexpression of YueB from a strong promoter such as P_{spac} is plausible to induce such an increment on k_{ads}. If the higher k_{ads} values are of simple explanation, the plaque phenotype is not. This raised the hypothesis that some operon genes could be participating in a step of phage infection, subsequent to adsorption. This was particularly true for *yukE* as its disruption had the most drastic effect on plaque size.

Any possible post-adsorption role played by the products encoded in this operon seems to be ruled out by three main experiments. First, ectopic expression of YukE in the *yukE* disrupted strain could not complement the effect on plaque size. In fact, this strain behaved exactly as the parental, signifying that the presence or absence of this specific product was irrelevant for SPP1 infection. Second, generation of nonsense mutants of genes *yukE* and *yukD* evidenced the expendable function of these products in phage infection. Finally, we managed to engineer a yukE operon deletion mutant strain that, as expected, was resistant to phage infection. However, overexpression of the receptor gene from the *amyE* locus rendered phage sensitivity and allowed normal sized plaques.

Without further experiments we can only speculate that the small plaque phenotype observed in the operon integration mutants (especially in the non-conditional *yukE* and the conditional *yukE* and *yukD* mutants), rather than resulting from the lack of the inactivated products, is the consequence of an imbalance in the stoichiometric relation between the operon products. This imbalance is likely to be created by the strong promoter activity of P_{spac} (in the presence of IPTG) and by the alteration of the native structure of the operon mRNA(s). Somehow this imbalance

has a negative effect in SPP1 infection in a step downstream of phage irreversible adsorption. In conclusion, none of the yukE operon products other than YueB seem to play a crucial role in SPP1 infection. However, if their native expression is altered, namely augmented, that seems to cause a negative effect in SPP1 infection after the step of irreversible binding.

We would expect an obvious impact in the growth parameters of SPP1 when propagated in the *yukE* integrant given the marked small sized plaques obtained in this strain. Curious enough, a SPP1 one-step growth curve in the *yukE* mutant did not varied significantly from that obtained in the wild type strain (not shown). However, we neither followed the lysis of cultures synchronously infected at high input multiplicities nor performed single-burst experiments, two procedures which allow a more accurate estimation of the lysis timing and burst-size. It is well known that a few minutes delay in lysis timing or a reduction of the burst size may result in a decrease of phage plaque size (Young, 1992; Abedon and Culler, 2007).

The hypothesis of the presence of a secondary operon promoter in the 5' end of the *yukD* coding sequence arise from the observation of minute plaques in the *yukE* (conditional and non-conditional) and *yukD* (conditional) integrants, in the absence of the inducer. A weak promoter here located could account for transcription of downstream genes after transcription from the upstream signals came to a stall in the secondary structure identified downstream of *yukE*. Also, the presence of this secondary structure immediately downstream of P_{spac} in those integrants could stabilize any transcript that initiated even in the absence of IPTG (Kozak, 2005), allowing for phage sensitivity.

II.4.2 – CELLULAR FUNCTION OF THE YUKE OPERON

In this work neither the cellular role of YueB nor of the operon products was addressed. Theoretically no major cellular function was affected as all integrants had the same growth rate as the wild type strain in laboratory conditions. However, all products shared similarities with members of the *S. aureus* Ess-Esx locus. This gene cluster encodes a new type of Gram-positive secretion system (Type VII; T7SS) that was first identified in *M. tuberculosis* (see Brodin *et al.,* 2004 and Abdallah *et al.,* 2007 for a review). T7SSs are specialized apparatus for the transport of extracellular proteins and are used by pathogenic bacteria, such as *M. tuberculosis* and *S. aureus* to actively secrete virulence factors. These belong to the ESAT-6 (early secreted antigenic target of 6

kDa) and CFP-10 (culture filtrate protein of 10 kDa) families. The ESX-1 and Ess-Esx locus of *M. tuberculosis* and *S. aureus*, respectively, are the most studied of this kind and each possesses two ESAT-6-like proteins (EsxA and EsxB, also called ESAT-6 and CFP-10, respectively, in *M. tuberculosis*; Fig. II.20). ESAT-6 proteins possess a distinguishable WXG motif, a signature sequence of ESAT-6-like proteins (Pallen, 2002), that is also present in the YukE protein of *B. subtilis* (Fig. II.8). These two antigens were found to associate and make a tight 1:1 complex (Renshaw *et al.,* 2005) secreted by the apparatus. Several independent studies have demonstrated that the genes that surround the ESAT-6-encoding genes are involved in the production of such specialized secretion system (for a review see Abdallah *et al.,* 2007). These include the FtsK/SpoIIIE ATPase -like proteins that are encoded by two genes in *M. tuberculosis* (Rv3870 and 3871) and only one in *S. aureus* (*essC*), which are homologous to *B. subtilis yukBA* (Fig. II.20). It was proposed that secretion is powered by the ATPase activity of these proteins. Vital for EsxA and EsxB secretion in *S. aureus* is also the EssB protein, which shares sequence similarity with the *B. subtilis* YukC. EsaA, EsaB and EsaC (ESAT-6 secretion accessory) are dispensable for the substrate secretion but, interestingly, in a recent report by Burts *et al.* (2008) it was shown that EsaC is also a secretion substrate under conditions that occur when bacteria enter host tissues. Production of EsaC is somehow regulated by EsaB, which shares sequence similarity with YukD. Interestingly, this protein of *B. subtilis* was crystallized and shown to adopt a fold that is closely related to ubiquitin (van den Ent and Löwe, 2005). However, conjugation assays performed by the authors indicate that *B. subtilis* YukD lacks the capacity for covalent bond formation with other proteins. The other accessory protein, EsaA, shares similarities with the SPP1 irreversible receptor, YueB, and no parallel gene was found in the genome of *M. tuberculosis*.

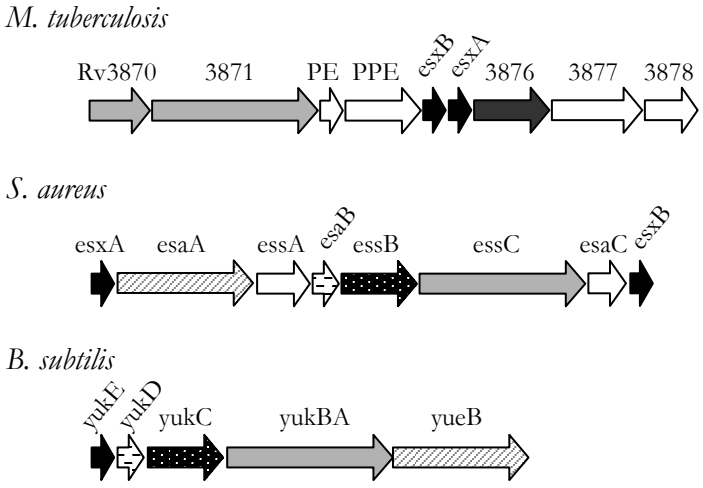


Figure II.20 | *M. tuberculosis*, *S. aureus* and *B. subtilis* genomic regions encoding ESAT-6/WXG100 (black arrows) and FtsK/SpoIIIE (light and dark gray arrows) family members. All members of the yukE operon have sequence correspondence to the other two represented gene clusters. Sequence similarities between

gene products are depicted by the same arrow pattern, except for white arrows that are region-specific genes. In *M. tuberculosis* the RD1 (ESX-1) locus corresponds to genes Rv3871 to Rv3878 plus Rv3879c (not represented). In *S. aureus* the Ess-Esx cluster is also incomplete. Genes not drawn to scale.

Putative T7SSs are present in several non-pathogenic bacteria and in pathogenic bacteria without being used for virulence (*Listeria monocytogenes*; Way and Wilson, 2005). Furthermore, it was recently established that the ESX-1 locus is essential for conjugal DNA transfer in *Mycobacterium smegmatis*, indicating that this system is not designed to act as a secretion system of virulence *per se*, rather it was acquired by evolution of function (Coros *et al.,* 2008; Abdallah *et al.,* 2007). The ESX-1 and Ess-Esx loci are dispensable for laboratory growth but essential for the pathogenesis of some bacteria (Pym *et al.,* 2002; Hsu *et al.,* 2003; Stanley *et al.,* 2003; Burts *et al.,* 2005). In *B. subtilis* this is most likely not the case, but no environment other than the laboratory growth in rich media was tested. The existence of an effective T7SS and its function in this bacterium remains elusive. However, the important related gastrointestinal pathogen *Bacillus anthracis* has six WXG100 proteins and seems to activate the T7SS during infection (Garufi *et al.,* 2008). The variations of T7SSs found between Actinobacteria (e.g. *M. tuberculosis*) and Firmicutes (e.g. *S. aureus*, *Bacillus* sp, and *L. monocytogenes*) impelled Abdallah *et al.* (2007) to propose the designation of a subtype for the Firmicutes systems, type VIIb.

II.4.3 – YUEB SURFACE LOCALIZATION

The surface localization of YueB was predicted by bioinformatics methods and assessed in this work by GFP fusions. Three independent YueB-GFP polypeptides were expressed from three different promoters in the *B. subtilis* genome, one from the native promoter sequences and two from strong, inducible promoters (P_{spac}; and P_{xylA}). In all three constructs YueB localized at the cell envelope, distributed in a rather ubiquitously way, except for small patches that were mainly observed at the cell poles or at mid-cell (future polar regions). To be confirmed, this targeted localization at the cell poles may not be completely unexpected. Recently it was demonstrated for a number of *E. coli*, *Yersinia pseudotuberculosis* and *Vibrio cholerae* phages that they adsorbed preferentially at the host polar regions. Further, for phage λ it was demonstrated that not only adsorption was greater at the poles but also that this bacterial region harbours fundamental phage developing steps, such as genome ejection, replication and fate of infection cycle (lytic *vs* lysogenic) (Edgar *et al.,* 2008).

It will be shown in Part III of this thesis that SPP1 has a reversible step of adsorption to an over represented cell wall polymer. Evidences will be presented indicating that this reversible interaction accelerates the recognition of the irreversible receptor YueB, which should be present in a much less number in the *B. subtilis* surface. If we assume that YueB is localized in a preferential zone, reversible adsorption at low multiplicity of infection (m.o.i.) gains extra importance for the phage, as a mean to "drive" it to the sub cellular location where ejection can take place. In conformity, Edgar *et al.* (2008) postulated a model for λ adsorption that begins with a first encounter with one of the many LamB receptors on the cell surface. LamB has been shown to be distributed in spirals that extend from pole to pole along the length of the cell and to move laterally in the outer membrane along these spirals (Gibbs *et al.*, 2004). Then, LamB-bound λ phage can migrate along the cell surface until it reaches the poles where it stops its movement. The DNA is injected at the poles, with the help of ManY, which localizes specifically at this region, where phage replication occurs. Replication of SPP1 DNA in close association with the membrane has been previously reported (Burger, 1980).

Although no information on the cellular localization of the Mycobacteria and Staphylococcal T7SSs is available, the fact that such systems comprise an FtsK/SpoIIIE-like member might suggest a preferable subcellular localization of this complex at the bacterial cell surface. The proteins belonging to this family operate in intercellular and intracellular DNA transfer (Pogliano *et al.,* 2003). FtsK is an essential and conserved cell division protein from *E. coli* that coordinates chromosome segregation, and is a homolog of *B. subtilis* SpoIIIE. The later is a chromosome partition protein that acts at the asymmetric sporulation septum by a conjugation-like mechanism (Wu and Errington, 1997). To accomplish their function FtsK proteins specifically localize to the septum at the cell middle region (Yu *et al.,* 1998). Presently, the existence of such FtsK like component in the T7SS is justified as the motor that provides ATP for the subtract transport.

PART III

PHAGE SPP1 REVERSIBLE ADSORPTION

III.1. INTRODUCTION

The nature of receptors and their interaction with virion components has been well characterized at the molecular level for a few tailed phages (section I.5). However, a complete integrated view of molecular and adsorption kinetics data is lacking in most cases. This is particularly true for phages infecting gram-positive bacteria where current knowledge on the adsorption process hardly extends beyond the identification of a few phage receptor-binding proteins and cellular receptors.

On the previous chapter we mentioned the fact that SPP1 could still adsorb, in a reversible way, to *pha-2 B. subtilis* resistant strains. The goal of the present study was to identify the cellular receptors involved in the SPP1 reversible interaction with the *B. subtilis* surface and to understand its impact in the whole process of SPP1 adsorption. We begin by introducing the underlying concepts of adsorption and its study and present the *B. subtilis* cell wall mutations most commonly associated with bacteriophage resistance.

III.1.1 – REVERSIBLE AND IRREVERSIBLE ADSORPTION: CONCEPTS AND THEORIES

Studies on phage adsorption have early suggested that this process could proceed in two distinct steps, one reversible step followed by irreversible binding (Adams, 1959). In general, reversible adsorption corresponds to binding to a non-saturable receptor whereas irreversible adsorption is typically characterized by binding to a specific saturable receptor. In practical terms, phages can be released from cells as infectious particles by dilution if they are reversibly adsorbed, but are not recoverable after irreversible binding as they become committed to infection, and hence inactivated. Reversible and irreversible interactions can occur with the same receptor as suggested for phage λ (Moldovan *et al.,* 2007) or, more frequently, involve different surface components (section I.5). The first step has been interpreted as a way to keep phages at the cell surface before the irreversible binding takes place. The division between irreversible adsorption and DNA ejection is not obvious and some authors believe that they are different descriptions of the same phenomenon. However, our group has shown that SPP1 efficiently binds irreversibly to purified YueB at 0ºC without leading to significant DNA ejection (São-José *et al.,* 2006). Therefore, at least in vitro, irreversible adsorption of SPP1 to YueB can be separated from DNA ejection. The unlocking of the metastable state of virions and the consequent DNA ejection occurs in

particular conditions of phage-receptor interaction, where the released free energy drives the required structural rearrangements leading to phage DNA ejection (Rossmann *et al.,* 2004; Kemp *et al.,* 2005; Plisson *et al.,* 2007).

The classical adsorption theory proposed individually by Schlesinger and Delbruck in the 30ths, that evolved from the early studies on adsorption kinetics, is called the collision theory and can be represented by the equation: P + B → PB* where P and B represent free phages and bacteria, respectively and PB* is the irreversible phage-bacteria complexes or infected bacteria (Adams, 1959). What they stated and was later tested by Schwartz (Schwartz, 1976), was that phages are captured by the host receptors at each collision. Under an excess of bacteria the process would follow a pseudo first-order kinetics being the rate of adsorption $d[P]/dt = -k[B][P]$, where k is the adsorption rate constant and P the concentration of virus remaining unattached to cells at the time t in the presence of a bacterial concentration B (Adams, 1959; Gamow, 1969). To explain observed deviations from the simple collision theory several authors have proposed models for bacteriophage adsorption:

Sequential model. In this model suggested by Puck, Garen, and Cline (1951) the reversible binding plays the role of a preliminary step essential for the establishment of the second step, the irreversible binding. The rate-limiting step in the whole process would be irreversible adsorption: P + B ↔ PB → PB* (being PB the reversible complex between phage and bacteria).

Competitive model. Stent and Wollman (1952) suggested an alternative model in which the reversible step, instead of leading to the irreversible one, competes with it. So, each collision can be "good" and irreversible or "bad" and reversible. The occurrence of reversible complexes would be therefore the rate-limiting step: PB ↔ P + B → PB*.

Modified sequential theory. From his experiments on bacteriophage T1, Christensen (1965) concluded that a combination of the two models would be the most accurate manner to explain the results. The existence of two types of reversible complexes, one that guides to the irreversible binding (a) and the other that leads to phage release (b) would be on the basis of the adsorption process and equilibrium: PB(b) ↔ P + B ↔ PB(a) → PB*.

The sequential model was the one generally adapted by researchers to explain phage adsorption and apparently is the model that responds best to the results obtained experimentally in several studied phages. So, the phenomenon of phage adsorption is typically described by the equation

$$P + B \underset{k_2}{\overset{k_1}{\leftrightarrow}} PB \xrightarrow{k_3} PB^* \qquad (1)$$

where k_1, k_2 and k_3 denote the reversible adsorption rate constant, the dissociation rate constant and the irreversible adsorption rate constant, respectively, and P, B, PB and PB* as stated above. The rate constants correspond to the ratio between the experimentally measured velocities of adsorption/dissociation and the number of host cells (section III.2.3.2).

In 1968 Adam and Delbrück conjectured that the advantage of a two-stage process during ligand/ligate interactions could be attributed simply to a reduction from a three-dimensional (3D) to a two-dimensional (2D) diffusion process, and they called this phenomenon the reduction of dimensionality (RD). This RD theory would explain the great sensitivity of some cells to low concentrations of ligands and, citing Eigen (1974), reduction of dimensionality is the "nature's trick to overcome the barrier of diffusion control, and make multimolecular reaction processes at low concentrations more efficient." This idea was applied by inumerous authors to explain actual biochemical and biological systems (see Axelrod and Wang, 1994 for references). Berg and Purcell in 1977 extended this concept to bacteriophage adsorption. Their initial idea assumed that the receptors were "perfect-sinks" that can irreversibly absorb with infinite capacity any ligands that collide with them, also known as the "diffusion-limited" reaction, thus creating a local concentration depletion zone (ligand concentration right at the target is held equal to zero). Several theoretical updates were added regarding condition limiting steps, e.g, single-capacity receptors (Berg, 1978; Collins and Kimball, 1949) and "reaction-limited" bindings in which the speed of the reaction is limited by the low binding probability rather than the depleted local concentration of the ligand (for more details see Axelrod and Wang, 1994).

III.1.2 – *BACILLUS SUBTILIS* CELL WALL RECEPTORS

The major cell wall teichoic acid (WTA) in *B. subtilis* 168 is D-alanyl-[α-D-glucosylated poly (1,3-glycerol-phosphate)], with a chain length of 45 to 60 glycerol-phosphate residues (Fig. III.1A; Neuhaus and Baddiley, 2003). WTA is attached to the peptidoglycan (C-6 of the MurNAc) via the linkage unit $(Gro\text{-}P)_{2\ or\ 3}$ ManNAc(β1-4)GlcNAc-P (abbreviation: Glc, glucose; Gro-P, glycerol-phosphate; MurNAc, N-acetylmuramic acid; ManNAc, N-Acetyl-D-mannosamine; GlcNAc-P, N-acetylglucosamine-phosphate). Lipoteichoic acids (LTA) in *B. subtilis* have a chain

length of 14-33 residues of Gro-P attached to C-6 of a glicolipid anchor composed of Glc(β1-6)Glc(β1-3)(gentiobiosyl)diacyl-Gro. However, the Gro-P residues of WTA and LTA are of the opposite sterochemical series, deriving from different sources and different biosynthetic pathways (Ward, 1981; for reviews see Neuhaus and Baddiley, 2003; Formstone *et al.,* 2008; Gründling and Schneewind, 2007). α-D-glucose and D-alanyl esters are both substituents of the WTA chain and the latter also decorates LTA. In fact, the D-alanyl esters of LTA are the precursors of those in WTA (Perego *et al.,* 1995). WTA substitution can only occur in C-2, meaning that the level of glucose/alanyl substituents is subjected to adjustments depending on the environmental and cellular conditions (Neuhaus and Baddiley, 2003).

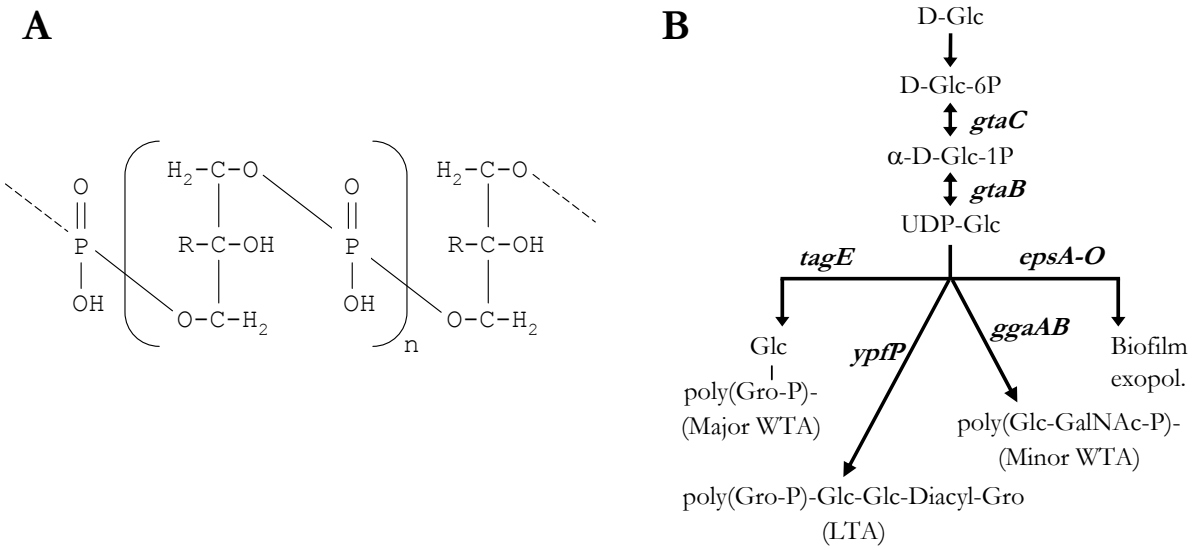


Figure III.1 | (**A**) Glycerol-phosphate WTA of *B. subtilis*. The substituent R can be hydrogen, α-D-glucose or D-alanyl ester. (**B**) Schematic representation of the pathway of UDP-glucose as a precursor for the synthesis and modification of the major *B. subtilis* cell wall polymers. Relevant genes involved in the different biosynthetic steps are depicted in boldface. *gtaC* is also known as *pgcA*. *ypfP* is also known as *ugtP*. *tagE* is also known as *gtaA*.

Mutations affecting the pathway of poly(Gro-P) glucosylation have been associated with *B. subtilis* resistance to several phages such as ϕ25, ϕ29, and SP01 (Yasbin *et al.,* 1976; Young, 1967). The nucleotide sugar UDP-Glucose (UDP-Glc) is a pivotal precursor in the biosynthesis of the major cell wall associated polymers in *B. subtilis* (Fig. III.1B). This molecule is a substrate of several glucosyltranferases (TagE, YpfP and glucosyltranferases from the ggaAB and epsA-O operons), which transfer glucosyl groups from UDP-Glc to constituents of the major cell WTA, LTA, minor WTA and of an exopolysaccharide involved in biofilm formation, respectively (Mauël *et al.,* 1991; Jorasch *et al.,* 1998; Branda *et al.,* 2004; Kearns *et al.,* 2005; Lazarevic *et al.,* 2005; Freymond *et al.,* 2006). Mutations impairing the formation of UDP-Glc, like those affecting α-

phosphoglucomutase and UTP:α-glucose-1-phosphate uridylyltransferase activities (*gtaC* and *gtaB*, respectively, Soldo *et al.,* 1993; Lazarevic *et al.,* 2005) have thus a pleiotropic effect by changing the composition of different polymers of the cell wall.

To our knowledge, the only report of the involvement of glucosylated teichoic acids in SPP1 adsorption to *B. subtilis* cells is more than 30 years old (Yasbin *et al.,* 1976). In this report the authors described that adsorption of SPP1 to strains carrying *gtaC*, *gtaB*, or *tagE* mutations was reduced relative to the wild type. However, the authors also observed that SPP1 could still plate on these mutants, whereas the same mutations conferred resistance to ϕ29 and SP01. This chapter presents the results that allowed us to explain this apparent contradictory behaviour of SPP1 and to have a global understanding of the SPP1 adsorption process to *B. subtilis*.

III.2. MATERIALS AND METHODS

III.2.1 – BACTERIAL STRAINS, BACTERIOPHAGES AND PLASMIDS

All biological material used in this work is listed in Table III.1.

Table III.1 | *E. coli* and *B. subtilis* strains, bacteriophages, vectors and plasmids used in this work.

Strains, phages and plasmids	Genotype or relevant features	Reference
***E. coli* strains**		
XL1-Blue MRF'	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI*^{q}Z⊿*M15* Tn*10 (Tet*^{r}*)*]	Stratagene
MC1061	*araD139* _(*ara leu*)*7697* _*lacX74 galU galK hsd*R*2 strA mcrA mcrB1*	Casadaban and Cohen, 1980
***B. subtilis* strains**		
L16601	*B. subtilis* 168; SPP1 indicator strain	Margot and Karamata, 1996
L5703	168/W23 interstrain hybrid	Karamata *et al.,* 1987
CSJ1	L16601 derivative resistant to SPP1; ⊿*yueB*	São-José *et al.,* 2004
CSJ3	L16601 derivative overexpressing *yueB* upon IPTG induction; Ery^{R}	São-José *et al.,* 2004
EB1451	Δ*tagO*::Ery^{R}	D'Elia *et al.,* 2006
HB0048	Δ*dltA*::Spec^{R}	Cao and Helman, 2004
L16131	*gtaC*ΩpBS640; Cm^{R}	V. Lazarevic, unpublished data
L16114	*gtaB*Ωp624; Cm^{R}	Lazarevic *et al.,* 2005
KP261	*ΔypfP*; Kan^{R}	Price *et al.,* 1997
CSJ1.*gtaC*	CSJ1 derivative, *gtaC*ΩpBS640; Cm^{R}	This work
CSJ1.*gtaB*	CSJ1 derivative, *gtaB*Ωp624; Cm^{R}	This work
CSJ1.*ypfP*	CSJ1 derivative, *ΔypfP*; Kan^{R}	This work
CSJ1.*tagE*	CSJ1 derivative, *tagE*ΩpCBM30; Ery^{R}	This work
CSJ1.*tagE.ypfP*	CSJ1 derivative, *tagE*ΩpCBM30 *ΔypfP*; Ery^{R}; Kan^{R}	This work
CSJ1.*dltA*	CSJ1 derivative, *ΔdltA*; Spec^{R}	This work
L16601.*gtaC*	L16601 derivative, *gtaC*ΩpBS640; Cm^{R}	This work
L16601.*gtaB*	L16601 derivative, *gtaB*Ωp624; Cm^{R}	This work
L16601.*ypfP*	L16601 derivative, *ΔypfP*; Kan^{R}	This work
L16601.*tagE*	L16601 derivative, *tagE*ΩpCBM30; Ery^{R}	This work
L16601.*tagE.ypfP*	L16601 derivative, *tagE*ΩpCBM30 *ΔypfP*; Ery^{R}; Kan^{R}	This work
CSJ3.*gtaC*	CSJ3 derivative, *gtaC*ΩpBS640; Ery^{R}; Cm^{R}	This work
CSJ1.EPS1	CSJ1 derivative, *epsDEF*ΩpSK-EPS1; Cm^{R}	This work

CSJ1.EPS2	CSJ1 derivative, *epsDEF*ΩpSK-EPS2; Cm^R	This work
***B. subtilis* phages**		
SPP1	Lytic phage	Riva *et al.,* 1968
SP01	Lytic phage; used to control transfer of cell wall mutations	Okubo *et al.,* 1964
Vectors and plasmids		
pMutin-4	Integration vector used for gene inactivation; Amp^R, Ery^R	Vagner *et al.,* 1998
pBluescript II SK +	Cloning vector; Amp^R	Stratagene
pCBM30	pMutin-4 derivative carrying a PCR product internal to *tagE*	This work
pSK-EPS1	pSK derivative carrying flanking regions of *epsDEF* separated by *cat86* (Cm^R)	This work
pSK-EPS2	pSK derivative carrying flanking regions of *epsHIJ* separated by *cat86* (Cm^R)	This work

III.2.2 – OLIGONUCLEOTIDES

The oligonucleotides used in this work are presented in Table III.2. These primers were used in the construction of mutant strains with modified cell wall polymers, except for primers epsC Fwd and epsK Rev, which were used for confirmation of mutants' genome organization by PCR, as they flanked the regions of interest.

Table III.2 | Oligonucleotides used in this work.

Oligonucleotide	Sequence (5'→3')[a]
tagE-1	TTG**gaatcc**CATGCGGTGAGTGAATCTAATAT
tagE-2	TTT**ggatcc**GTCAAGGCCTATCCTTTCTTC
EC Fwd Pst	CGG**ctgcag**GCGGATCAATCGGCTCGGA
EC Rev tagBam	TCA**ggatcc**CTAATGAACGCTGGCAGCCG
EG Fwd tagBam	ATGACGGCTGCCAGCGTTCATTAG**ggatcc**TGAGGAAAAGGACCATAAC CGATGATTGTATATG
EG Rev Xba	AAT**tctaga**CTAGGATCATTTGATACAGACC
EG Fwd Pst	CAG**ctgcag**GCTATCGAGTCGGCACGGAT
EG Rev tagBam	GTG**ggatcc**TTACCGGGAAAAAATCGTTCTGT
EK Fwd tagBam	CTTACAGAACGATTTTTTCCCGGTAA**ggatcc**CACGCAGTGCGAAGCAG GCATGAAATTCACG
EK Rev Xba	TCC**tctaga**TAAGGCAGCAGGAAGTGCAAG

epsC Fwd	CTCGACACTAGCGAAATTTCG
epsK Rev	AGGACGGTCCGAGCCAGATT

[a] Boldface lowercase letters indicate endonuclease restriction sites (*Eco*RI-**gaattc**; *Bam*HI-**ggatcc**; *Sma*I-**cccggg**; *Pst*I-**ctgcag**; *Xba*I-**tctaga**).

III.2.3 – GROWTH CONDITIONS AND MICROBIOLOGICAL METHODS

III.2.3.1 – *E. coli*, *B. subtilis* and bacteriophage growth conditions

E. coli and *B. subtilis* growth conditions, media supplements and antibiotic concentrations were as described in sections II.2.3.1 and II.2.3.2. Kanamycin was used (7.5 μg/mL) in *B. subtilis* for construct selection. Phage propagation and purification was also as described previously (II.2.3.2 and II.2.3.3).

III.2.3.2 – Measurement of SPP1 adsorption to *B. subtilis* cells

In all adsorption assays *B. subtilis* cells were collected from exponentially growing cultures (A_{600}=0.8, about 10^8 colony forming units per mL) and adjusted to different A_{600} in LB medium (ranging from 0.4 to 16). Unless otherwise indicated, cells were supplemented with 15mM $CaCl_2$ and 50μg/mL of chloramphenicol. Chloramphenicol ensured the arrestment of cellular growth and phage development throughout the experiments. From here on cell incubation was without shaking. Cells were equilibrated for 10 minutes at the different tested temperatures (from 0 to 37ºC) and then SPP1 was added in order to obtain ~10^7 plaque forming units per mL (pfu/mL) and mixed by brief vortexing. Control mixtures without added cells were used to confirm the phage input in each experiment.

SPP1 total adsorption was measured at defined times of cell/phage contact by rapidly sedimenting cells from 100 μL mixture samples (0.5 to 1 minute centrifugation, 15000g, room temperature) and titrating the free phages present in the supernatant (pfu/mL).

The initial rate of reversible adsorption (R_{ads}) to strain CSJ1 (unable to adsorb irreversibly) was calculated at time points 0.25, 0.5 and/or 1 min (depending on the tested temperature) after phage addition using the formula $R_{ads} = ln(P0/P)/\Delta t$, where $P0$ is the phage input and P the

fraction of free phages after a Δt period. The **adsorption rate constant** (k_1) is the ratio between R_{ads} and the bacterial cell mass expressed as A_{600} (equal to 0.8 in k_1 determinations).

SPP1 reversible adsorption was also quantitatively studied by determining the fraction of free phages at the equilibrium state (time points 40, 50 and 60 min after phage addition unless stated otherwise). The results obtained allowed the calculation of the **equilibrium constant**: $K_{eq} = k_1/k_2 = (P0\text{-}P)/(P.A_{600})$, being $P0$, P and A_{600} as described above. Unless indicated, in the assays with *B. subtilis* cell wall mutants the value of K_{eq} was determined with cultures adjusted to at least three different A_{600} values (ranging from 0.4 to 16), with the lowest one tested ensuring a minimum of 50% of SPP1 adsorption. The **dissociation rate constant** (k_2) was determined using the values obtained experimentally for K_{eq} and k_1 and the formula $K_{eq} = k_1/k_2$.

SPP1 irreversible adsorption was measured with *B. subtilis* cultures set at A_{600} = 0.8 or 1.6 as described in II.2.3.4.

III.2.3.3 – SPP1 adsorption in the presence of Concanavalin A and α-methyl-glucoside

Concanavalin A (ConA) and α-methyl-glucoside (α-M-G) were purchased from Sigma Aldrich. CSJ1 cells were concentrated to an A_{600} of 1.6 in TBT buffer supplemented with 0.01% gelatin, 0.5 mg/mL ConA, 15 mM $CaCl_2$ and 50 μg/mL chloramphenicol. After an incubation period of 10 min at 37ºC, phage SPP1 was added (~10^7 pfu/mL) and the mixtures were placed on a rotating wheel with gentle agitation at room temperature (~24.5ºC). The presence of gelatin ensured phage stability during permanent agitation. Reversible adsorption was evaluated by determining the fraction of free phages at the equilibrium state. Sixty minutes after phage addition, 25mM of α-M-G was added and the fraction of free phages was determined at times 75, 90 and 105 min. A control sample was prepared as the test tube but ConA dilution buffer (10mM Tris, 10mM $MgCl_2$, pH 7) was added instead of ConA. Phage input was controlled as described above.

III.2.4 – MOLECULAR AND BIOCHEMICAL METHODS

DNA extraction and separation by gel electrophoresis was as described in II.2.4.1. PCR amplifications, restriction endonuclease digestions, DNA ligations, and DNA purification and

extraction from gels were as described in II.2.4.4. All the oligonucleotides used in this work were purchased from Invitrogen and are listed in Table III.2. Development of *E. coli* and *B. subtilis* competence and transformation was as described by Chung *et al.,* 1989 and Yasbin *et al.,* 1973, respectively.

Production of protein extracts and analysis by SDS-PAGE and Western blotting was as described previously (II.2.4.6 and II.2.4.7).

III.2.5 – CONSTRUCTION OF *B. SUBTILIS* STRAINS

III.2.5.1 – Mutants affected in WTA and LTA composition/modification

A *B. subtilis tagE* integrant was obtained by the pMutin-4 strategy described in section II.2.5.1. The *tagE* internal region cloned in pMutin-4 was amplified with primers pair tagE-1/tagE-2. This integrant was selected and grown in the presence of 0.5 mM IPTG to prevent transcriptional polar effects on the downstream indispensable *tagF* gene.

B. subtilis strains carrying *gtaC*, *gtaB*, Δ*ypfP* and Δ*dltA* mutations (L16131, L16114, KP261 and HB0048, respectively) were kindly provided by V. Lazarevic (*gta* mutants), R. Losick (*ypfP*) and J. D. Helmann (*dltA*). The *B. subtilis* 168/W23 interstrain hybrid was also a gift of V. Lazarevic.

Transfer of mutations *gtaC*, *gtaB*, Δ*ypfP*, *tagE* and Δ*dltA* to L16601, CSJ1, and CSJ3 genetic backgrounds was performed by transforming the recipient strains with donor chromosomal DNA. The mutations corresponded to gene disruptions after plasmid integration (*gtaB*, *gtaC* and *tagE*) and to gene replacement by an antibiotic resistance cassette (*ypfP* and *dltA*). Transformants were selected for resistance markers, and mutations were confirmed by PCR and by the SP01 resistance phenotype (in the case of *gta* and *tagE* mutations [Yasbin *et al.,* 1976]). *B. subtilis gtaC*, *gtaB*, and *ypfP* mutant strains were grown in the presence of 1 mM $MgCl_2$ to ensure normal growth and prevent abnormal cell morphologies (Lazarevic *et al.,* 2005).

III.2.5.2 – Mutants of the exopolysaccharide operon

The gene clusters *epsDEF* and *epsHIJ* of the exopolysaccharide operon (*eps* operon; Fig. III.2) were deleted through recombination with plasmids carrying flanking regions separated by the *cat86* gene. PCR products obtained with primers pairs EC Fwd Pst/EC Rev tagBam and EG Fwd tagBam/EG Rev Xba, were used to construct plasmid pSK-EPS1. PCR products using primers pair EG Fwd Pst/EG Rev tagBam and EK Fwd tagBam/EK Rev Xba, were used to construct plasmid pSK-EPS2. The structure of the plasmids was completed by introducing *cat86* gene in the *Bam*HI site created in between the cloned flanking regions. Transformation of *B. subtilis* CSJ1 strain with these plasmids allowed for gene replacement by double-crossover recombination, and the recombinants were selected in the presence of chloramphenicol. PCR reactions were used to confirm the desired genomic structure of the recombinants.

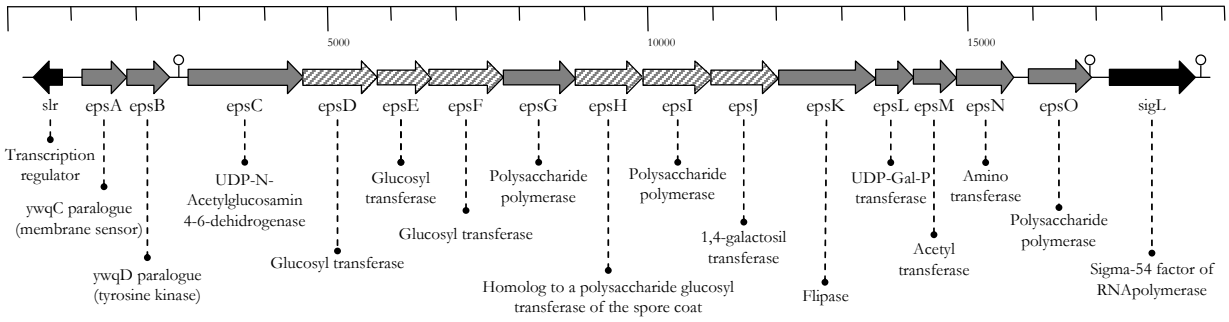


Figure III.2 | Gene organization of the *B. subtilis* chromosome region that contains the e*ps* operon. Putative function of the gene products are indicated. Genes that are not a part of the operon are depicted in black. The genes represented as slashed arrows were replaced for the chloramphenicol resistance gene (*cat86*). The remaining genes of the operon are in gray. Inverted repeats with the potential to form stem-loop type structures are indicated ' ⚲ '.

III.3. RESULTS

III.3.1 – KINETICS OF SPP1 REVERSIBLE ADSORPTION TO *YUEB* MUTANT CELLS

We started this work by studying the SPP1 adsorption profile to *B. subtilis* strain CSJ1. As mentioned before this strain does not synthesize the YueB receptor because it carries an extensive deletion of the corresponding gene (São-José *et al.,* 2004). The idea was to check if CSJ1 behaved as the previously described *pha-2* mutants (section II.1), namely with respect to the effect of temperature in SPP1 adsorption (Santos, 1991). Cells of the wild type strain L16601 and of the SPP1 resistant derivative CSJ1, were pre-equilibrated at 0ºC and initially challenged for SPP1 total adsorption. Total adsorption refers to the sum of phages adsorbed either reversibly or irreversibly (section III.2.3.2). Twenty min after phage addition each strain had already bound about 99% of the added virus particles (pfu/mL) (1% of free phages), being the rate of SPP1 adsorption very similar in both strains (Fig. III.3A). The fraction of adsorbed phages remained basically unchanged for an additional period of 20 min, meaning that an equilibrium between adsorbed and free phages had been reached. The mixtures were then shifted to 37ºC and scored for total and irreversible adsorption during 40 min (filled and open symbols in Fig. III.3A, respectively). We evaluated irreversible adsorption by diluting 100-fold samples of the bacteria/phage mixtures prior to their centrifugation, a procedure which releases reversibly adsorbed phages (Adams, 1959; section III.2.3.2). The measurement of these two types of adsorption 10 min after shifting the temperature indicated that phages had been released from both strains. However, whereas SPP1 efficiently re-adsorbed to the wild type strain with the continuous incubation at 37ºC, it completely failed to re-adsorb to strain CSJ1 to the level observed at 0ºC. The total fraction of phages adsorbed to strain CSJ1 quickly stabilized around 80% (filled squares curve in Fig. III.3A). Moreover, these adsorbed phages were shown to be bound reversibly to CSJ1 since dilution of the bacteria/virus complexes allowed complete recovery of the initial phage input (open squares curve in Fig. III.3A). Dilution of CSJ1/SPP1 complexes maintained at 0ºC gave the same results as long as enough time was allowed to phage release (data not shown). Note that the continuous incubation of SPP1 with the wild type strain at 37ºC resulted in higher percentages of both total and irreversible adsorption when compared to total adsorption observed at 0ºC. This most probably results from the fact that at the lower temperature the contribution of the irreversible binding to total adsorption is much reduced (see below). Overall, the results presented above indicated that SPP1 adsorbs reversibly to CSJ1 strain

(but not irreversibly), apparently with the same efficiency as observed for the wild type strain, as previously observed for the *pha-2* mutants.

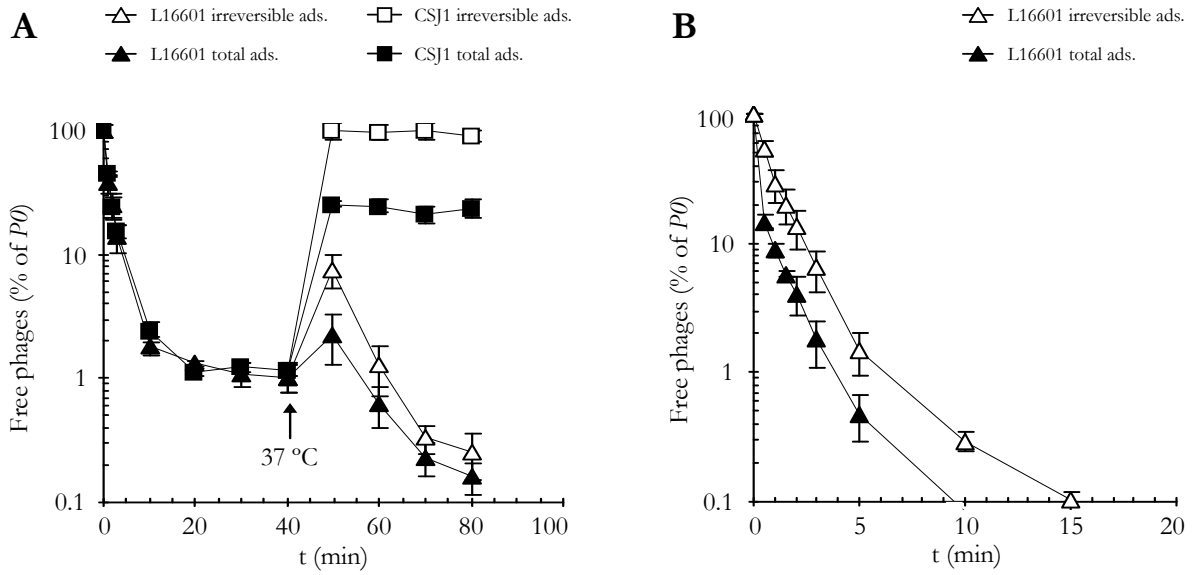


Figure III.3 | SPP1 adsorption profile to the wild type strain L16601 (triangles) and to the *yueB* strain CSJ1 (squares). Free phages are represented as a normalized concentration (*P*x100)/*P0*, where *P* is the titer of free phages (plaque forming units per ml (pfu/ml)) at time *t* and *P0* the initial input. The plotted values correspond to the average from three independent experiments. Standard deviation bars are shown. (**A**) Cultures of both strains were adjusted to an optical density at 600 nm (A_{600}) of 1.6 and equilibrated at 0ºC. SPP1 was then added to each culture and phage total adsorption measured during a period of 40 min. At this point (arrow) mixtures were transferred to 37ºC and both total (filled symbols) and irreversible adsorption (open symbols) evaluated for an additional period of 40 min. (**B**) SPP1 total (filled symbols) and irreversible adsorption (open symbols) to strain L16601 set to A_{600} = 0.8 and pre-equilibrated at 37ºC before phage addition.

SPP1 reversible interaction with the *B. subtilis* surface is further illustrated in Figure III.3B, which shows the total and irreversible adsorption profiles to the wild type strain, at 37ºC. SPP1 total adsorption to cells is a very fast process, with less than 1% free phages recovered from the supernatant after 5 min incubation. At each scored time point the percentage of total adsorption is always higher than the percentage of irreversibly bound phages, indicating the presence of a fraction of reversibly adsorbed phages in the first case. In the irreversible binding curve we could observe an initial exponential decay of free phages and extract an adsorption rate constant (k_{ads} = 1.13 ± 0.18 min^{-1} . A_{600}^{-1}; section III.2.3.2). In contrast, the total adsorption curve revealed a strong sink of free phages at the beginning (t = 0.5 min), which preceded an exponential decay. Although being most probably an underestimation, the initial rate of total adsorption determined at the single point t = 0.5 min is 4.8 ± 0.3 min^{-1} . A_{600}^{-1}.

III.3.2 – EFFECT OF TEMPERATURE ON SPP1 ADSORPTION: PHAGE RELEASE FOLLOWS ARRHENIUS KINETICS

In a wild type background the phenomenon of phage adsorption is typically described by equation (1) (section III.1). For CSJ1 strain, to which SPP1 is unable to adsorb irreversibly this is simplified to

$$P + B \underset{k_2}{\overset{k_1}{\leftrightarrow}} PB \qquad (2)$$

where a specific equilibrium between the fraction of free and adsorbed phages depends on the relative contribution of k_1 and k_2. The results presented above (Fig. III.3A) suggested that at least the release of reversibly adsorbed SPP1 virus particles is a temperature(T)-dependent process, in a way that high temperatures would favour phage dissociation (k_2) from cells. To test this hypothesis we have measured the fraction of unbound phages at different temperatures, after reaching the equilibrium state with strain CSJ1. As shown in Figure III.4A, the results supported our hypothesis as the fraction of unadsorbed phages increased exponentially with T. The increment of free phages with T could not be attributed to a negative effect of the high temperature on the adsorption rate constant (k_1) as shown below.

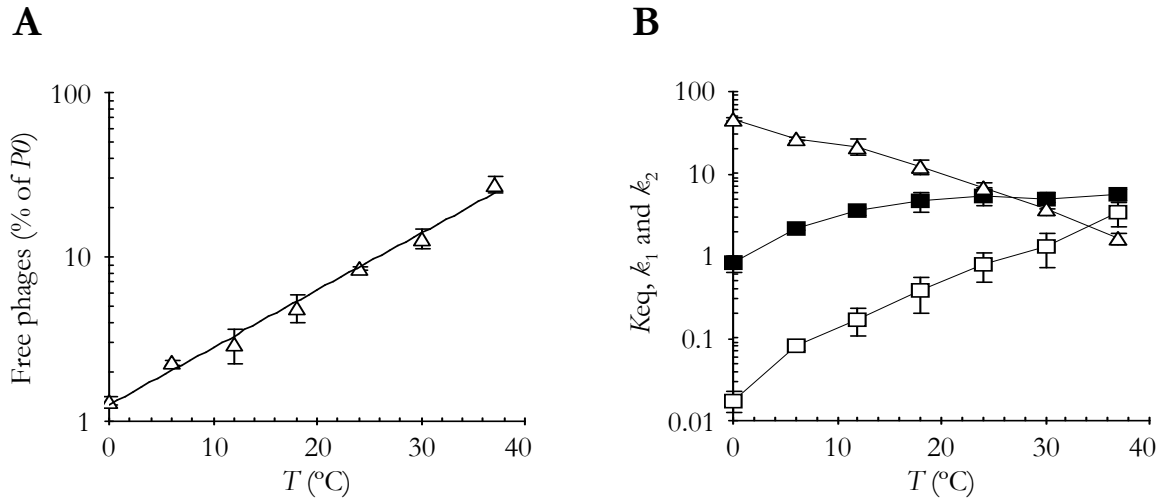


Figure III.4 | Effect of temperature on SPP1 reversible adsorption to strain CSJ1. (**A**) A CSJ1 culture was adjusted to A_{600} = 1.6 and samples equilibrated at 0, 12, 18, 24, 30 and 37ºC before phage addition. For each temperature the fraction of free phages in the equilibrium state was measured (plotted values as in Fig. III.3). Correlation coefficient for exponential dependence is 0.9919. (**B**) Variation of the equilibrium (K_{eq} (A_{600}^{-1}), triangles), adsorption (k_1 (A_{600}^{-1} . min^{-1}), filled squares) and dissociation (k_2 (min^{-1}), open squares) constants with temperature (averaged values from at least three independent measurements with indication of standard deviation).

The equilibrium constant (K_{eq}), which provides a measure of the number phage/bacteria complexes and therefore of the relative contribution of k_1 and k_2 (expression (2)), is determined by

$$K_{eq} (A_{600}^{-1}) = (P0-P)/(P.A_{600}) = k_1/k_2 \qquad (3)$$

where *P0* is the initial phage input, *P* the fraction of free phages at the equilibrium state and A_{600} the cellular mass (here expressed as the absorbance at 600nm). Equation (3) tells us that increasing the fraction of adsorbed phages (low *P* values) results in higher K_{eq} values.

As the increment of free phages observed in Figure III.4A could be attributed either to a negative effect of temperature rising on k_1 or to a positive effect on k_2, we decided to gain more quantitative insight by experimentally determining K_{eq} and k_1 for a range of temperatures. SPP1 reversible adsorption to strain CSJ1 is characterized by a fast sink of free phages in the first instances after phage addition and by the rapid establishment of an equilibrium state (not shown). We have thus taken the adsorption rate calculated at the shortest time point possible to estimate the reversible adsorption rate constant k_1 at the different temperatures (section III.2.3.2). We observed that k_1 increased moderately with temperature levelling at values ranging between 4.6 and 5.6 min^{-1} . A_{600}^{-1} for $T \geq 18$ºC (Fig. III.4B). On the contrary, K_{eq} showed a sharp decrease with *T* (Fig. III.4B), in agreement with the temperature-dependence increment of free phages shown in Figure III.4A. The extraction of k_2 (equation (3)) from the experimentally obtained values of K_{eq} and k_1 at the different temperatures indicated that the dissociation rate constant increased essentially in an exponential way with *T* (Fig. III.4B). Thus, temperature seems to have a major impact on the desorption rate of SPP1 from *B. subtilis* cells (k_2 increases 200 fold from 0 to 37ºC) and a much less pronounced effect on the adsorption rate constant (an average 6 fold increase of k_1 at temperatures ≥ 18ºC when compared to the value obtained at 0ºC).

Table III.3 summarizes all the equilibrium and rate constants values obtained for SPP1 adsorption to *B. subtilis* strain CSJ1. From k_2 values a forth parameter can be extracted ($1/k_2$), which is an estimation of time that phages remain attached to *B. subtilis* cells at the different temperatures before being released.

Although what is relevant here is the relative variation of the constant values, it may happen that the absolute values of k_2 are underestimated. This is because they derive from k_1/K_{eq} where k_1 is probably underestimated because of practical limitations (see above). In this case, the retention time ($1/k_2$) should correspond to an overestimation.

Table III.3 | Equilibrium and rate constants values measured for every temperature tested.

Temp (°C)	K_{eq} (A_{600}^{-1})	k_1 (min^{-1}. A_{600}^{-1})	k_2 (min^{-1})	$1/k_2$ (sec)
0	46.843 ± 2.475	0.815 ± 0.192	0.017	3447.534
6	26.956 ± 1.146	2.210 ± 0.178	0.082	731.698
12	21.465 ± 4.915	3.603 ± 0.410	0.168	357.477
18	12.342 ± 2.392	4.659 ± 1.193	0.377	158.942
24	6.808 ± 1.051	5.445 ± 1.260	0.800	75.023
30	3.777 ± 0.688	4.892 ± 1.157	1.295	46.325
37	1.646 ± 0.262	5.604 ± 0.890	3.404	17.625

Figure III.5 shows that $ln(k_2)$ varies linearly as a function of the inverse of T (T in Kelvin). The temperature dependence of the dissociation rate constant is therefore consistent with the Arrhenius law $ln(K) = -Ea/R \cdot 1/T + \log(A)$, where Ea is the activation energy, R the gas constant (8.314472(15) J . K^{-1} . mol^{-1}) and A the pre-exponential factor. From the slope of the curve (= -Ea/R) in Figure III.5 we calculated an activation energy for phage SPP1 release of 94.59 kJ/mol (22.6 kcal/mol).

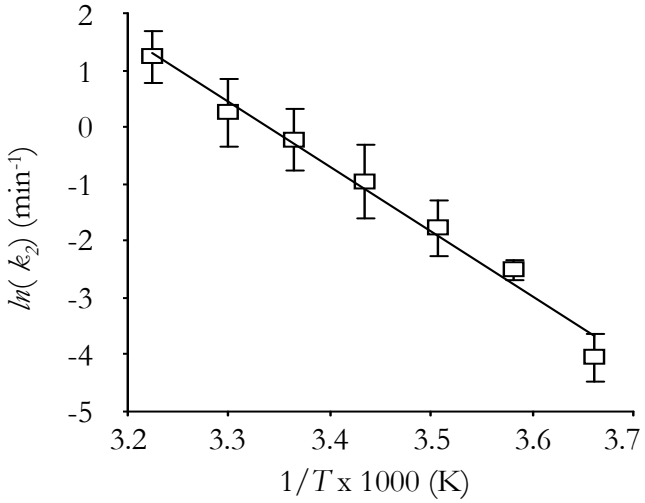


Figure III.5 | Arrhenius plot describing the variation of $ln(k_2)$ as a function of the inverse of T. Correlation coefficient is 0.9843. Error bars are represented.

III.3.3 – MUTATIONS IMPAIRING GLUCOSYLATION OF WTA AND LTA AFFECT SPP1 REVERSIBLE ADSORPTION

In the previous sections we showed that SPP1 adsorbs in a reversible way to CSJ1 strain and that the maintenance of phage/bacteria complexes is favoured at low temperatures. We took advantage of this feature to search for potential *B. subtilis* cell wall polymers (and their substituents) involved in SPP1 reversible adsorption. For that we have determined the K_{eq} value at 0ºC for CSJ1 derivative strains affected in some of the *B. subtilis* cell wall biosynthetic steps referred to above (Fig. III.1B). No significant changes in cell growth rate or cell morphology were observed between cell wall mutants and CSJ1 parental strain (not shown), provided that the required additives were supplied to the culture medium (Mg^{2+} for *gta* and *ypfP* mutants, Lazarevic *et al.* (2005), and IPTG in the case of *tagE* as the downstream *tagF* gene is essential). Mutations impairing the synthesis of UDP-Glc (*gtaC* and *gtaB*) had a drastic effect on SPP1 reversible adsorption (Fig. III.6). We were not able to detect the presence of phage/bacteria complexes with these mutants, even when cultures with A_{600} = 16 were used (we have considered as significant only averaged adsorption values ≥ 50%). In these conditions the value of K_{eq} for *gta* mutants was thus considered as zero. The K_{eq} value for the parental strain CSJ1 was 45.8 ± 7 A_{600}^{-1} (Fig. III.6).

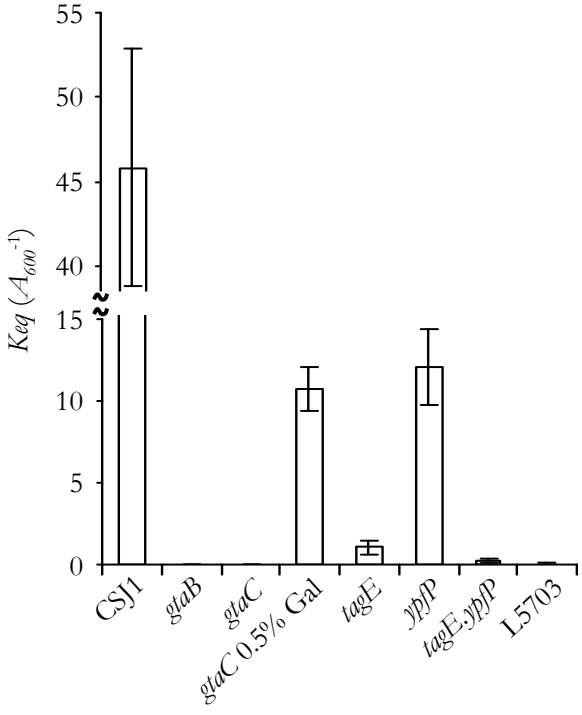


Figure III.6 | Effect of cell wall mutations in SPP1 reversible adsorption. K_{eq} was measured in CSJ1 derivative strains carrying mutations in some of the genes shown in Figure III.1B. L5703 is an interstrain hybrid producing glucosylated poly(ribitol phosphate) WTA of strain W23 and LTA of strain 168. K_{eq} values are the average of at least three measurements using cultures set to different A_{600} and pre-equilibrated at 0º C before phage addition. CSJ1 (A_{600} = 0.4, 0.6 and 0.8); *gtaB* and *gtaC* (A_{600} = 16); *gtaC* grown in the presence of 0.5% galactose (A_{600} = 0.8, 1.6 and 3.2); *tagE* (A_{600} = 3.2, 6.4 and 8); *ypfP* (A_{600} = 0.8, 1.6 and 3.2); *tagE.ypfP* (A_{600} = 8 and 16); L5803 (A_{600} = 16). Error bars are represented. The reason for the different A_{600} values is explained in section III.2.3.2).

The *gtaC* adsorption phenotype could be partially reverted by growing the mutant strain in the presence of 0.5% galactose (K_{eq} = 10.7 ± 1.4 A_{600}^{-1}), a trend also observed in studies of phage ϕ25 adsorption to *B. subtilis gtaC* mutants (Birdsell and Doyle, 1973). The molecule UDP-Glc seemed thus to be an essential precursor for the synthesis of a cell wall polymer acting as receptor for SPP1 reversible adsorption. We can draw the following observations concerning the four main cell wall polymers enumerated in Figure III.1B:

Minor WTA. Previous studies from our group have shown that *ggaAB* mutations had no effect on SPP1 adsorption (Santos, 1991);
Biofilm exopolysaccharides. The K_{eq} values measured in *eps* deletion mutants, CSJ1.EPS1 and CSJ1.EPS2, each containing a deletion of three major operon glucosyltranferases (section III.2.5.2), did not differed significantly from that calculated in the parental strain CSJ1 (not shown), indicating that the exopolysaccharide encoded by the epsA-O operon does not participate in SPP1 adsorption;
Major WTA. The K_{eq} value (1.1 ± 0.2 A_{600}^{-1}) calculated for the *tagE* mutant (carrying non-glucosylated WTA) showed a 42-fold decrease when compared to that of strain CSJ1;
LTA. The *ypfP* mutation, which changes the chemical composition of LTA lipid anchor, resulted also in a decrease of bound phages (K_{eq} = 12.1 ± 2.3 A_{600}^{-1}) although much less pronounced than in the case of *tagE*.

The fraction of adsorbed SPP1 was almost abolished in a *tagE*/*ypfP* double mutant strain (K_{eq} = 0.3 ± 0.1 A_{600}^{-1}). Interestingly, an extremely low level of adsorption was also detected with strain L5703 (K_{eq} = 0.06 ± 0.1 A_{600}^{-1}), a *B. subtilis* 168 derivative in which the *tag* operon was replaced by the *tar* operon of strain W23 (Karamata *et al.,* 1987**;** Young *et al.,* 1989). This strain produces glucosylated poly(ribitolphosphate) WTA of strain W23 and LTA of strain 168.

Moreover, a CSJ1 derivative of strain HB0048 (Cao and Helmann, 2004; Table III.1), that carries a mutation in *dltA* gene, was also tested for its ability to reversibly bind SPP1 and it behaved exactly as the parental strain CSJ1, meaning that the LTA substitutent D-alanine is not implicated in SPP1 reversible adsorption (not shown).

The *B. subtilis* Δ*tagO* mutant (Table III.1) recently described by D'Elia *et al.,* 2006, would have been a good control strain to test, since this mutant was designed to exclusively affect WTA

synthesis, instead of an absence of UDP-Glc that also affects the LTA composition. Unfortunately this mutant had several growth problems that impeded its use in any adsorption assay.

Overall, this genetic analysis strongly suggested that the major target for SPP1 reversible adsorption to *B. subtilis* cells are the glucosyl residues of poly(Gro-P) WTA. The results also indicate that poly(Gro-P) cannot be substituted by poly(ribitolphosphate) as a receptor site for SPP1.

III.3.4 – TARGETING OF GLUCOSYL RESIDUES WITH CONCANAVALIN A INHIBITS SPP1 REVERSIBLE ADSORPTION

The lectin concanavalin A (ConA) was reported to interact specifically and reversibly with α-D-glucosyl residues of the poly(Gro-P) backbone of *B. subtilis* teichoic acids (Doyle and Birdsell, 1972). It was shown that pre-incubation of *B. subtilis* cell walls with ConA inhibited adsorption of phage ϕ25 (Birdsell and Doyle, 1973), in agreement with the previous reported role of glucosyl substitution in the receptor activity of teichoic acids (Young, 1967).

We observed a similar ConA-dependent inhibition of SPP1 reversible adsorption to *B. subtilis* strain CSJ1 (Fig. III.7). ConA-treated CSJ1 cells showed a marked decrease in the ability to adsorb SPP1, being the K_{eq} value measured in these conditions about 9-fold lower than the value obtained in the control sample without ConA. The addition of α-methyl-glucoside (α-M-G), a molecule with high affinity to ConA, dissociated the ConA-cell wall complexes and restored CSJ1 adsorption capacity to levels comparable to the ConA-untreated cells. After addition of α-M-G a reproducible, slow increase of the fraction of free phages was observed both in Con A-treated and untreated cells. This phenomenon might be explained if α-M-G competes with glucosyl residues on the cell wall for SPP1 binding.

The inhibitory effect of ConA on SPP1 adsorption indicated that the phage binds to glucosyl residues of poly(Gro-P), a result that agreed with the reduced level of SPP1 reversible adsorption to the *tagE* mutant (see above).

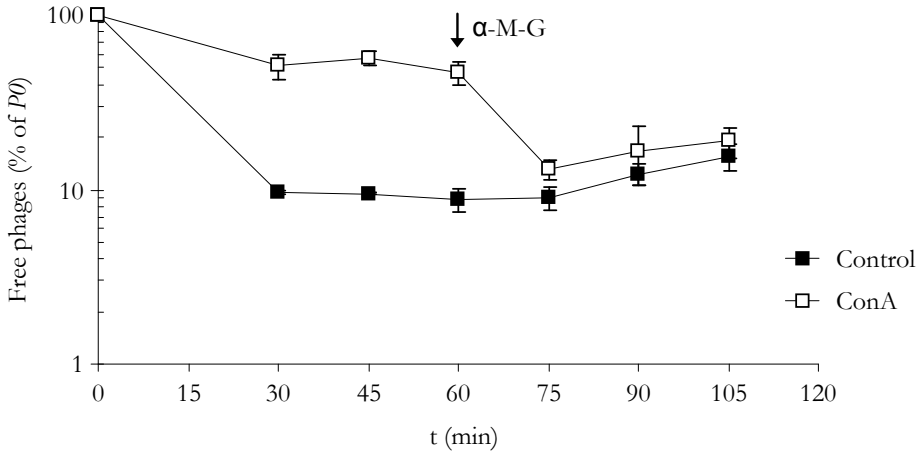


Figure III.7 | Effect of ConA on SPP1 reversible adsorption to CSJ1 cells. Strain CSJ1 was concentrated to A_{600} = 1.6 and incubated at 37ºC for 10 min in the presence of ConA (open symbols) or ConA dilution buffer (filled symbols). SPP1 was then added, the mixtures placed on a rotating wheel at room temperature and the fraction of free phages in equilibrium determined. The arrow indicates the time of addition of α-M-G. The plotted values correspond to the average of three independent experiments. Standard deviation bars are shown.

III.3.5 – IMPACT OF REVERSIBLE ADSORPTION IN SPP1 IRREVERSIBLE BINDING AND INFECTION

As noticed in the previous sections we were able to study SPP1 reversible adsorption without interference of irreversible binding by working in a *yueB* genetic background (CSJ1 strain). In order to elucidate the role of reversible interaction in SPP1 irreversible binding to YueB we have studied the effect of the abovementioned cell wall mutations in a wild type background of *yueB* expression (L16601 as recipient strain).

III.3.5.1 – SPP1 irreversible adsorption to cell wall mutants

As shown in Figure III.8, cell wall mutations essentially reduced the rate of irreversible binding proportionally to their effect on reversible adsorption, except for those strains carrying the *ypfP* mutation (see below). Mutation *gtaC*, which eliminated SPP1 reversible adsorption (see above), drastically slowed down the rate of irreversible binding when compared to the wild type strain (L16601). Similar results were obtained with a *gtaB* strain (not shown).

The initial rate of irreversible adsorption to the wild type and mutant strains (see exponential fitting curves in Fig. III.8) was used to calculate the irreversible adsorption constant (k_{ads}), which integrates the contribution of k_1, k_2 and k_3. The k_{ads} value obtained for the *gtaC* mutant strain

(0.03 ± 0.11 min^{-1} . A_{600}^{-1}) was about 30-fold lower than that obtained in the wild type strain (1.07 ± 0.10 min^{-1} . A_{600}^{-1}). The *tagE* mutation caused also a significant reduction of k_{ads}, exhibiting an 8-fold decrease compared to the wild type strain. Unexpectedly, the *ypfP* deficiency produced only a slight decrease of SPP1 irreversible adsorption ($k_{ads} = 0.91 \pm 0.16$ min^{-1} . A_{600}^{-1}), in spite of its apparent negative effect on reversible adsorption (Fig. III.6). Also surprising, the k_{ads} value obtained with the double-mutant strain *tagE/ypfP* was close to that of *tagE*, although these mutations resulted in significantly different K_{eq} values (about 0.2 and 1 A_{600}^{-1}, respectively, see above).

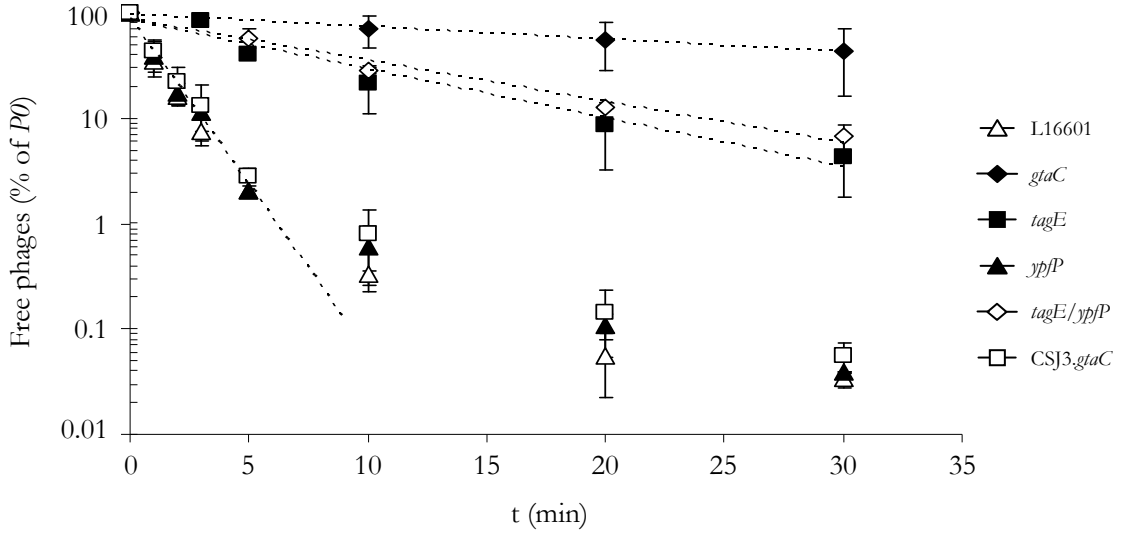


Figure III.8 | SPP1 irreversible adsorption to *B. subtilis* cell wall mutants. Cultures of the wild type strain L16601 (open triangles) and of its derivatives carrying mutations *gtaC* (filled diamonds), *tagE* (filled squares), *ypfP* (filled triangles) and *tagE.ypfP* (open diamonds) were grown to $A_{600} = 0.8$ and equilibrated at 37ºC before phage addition. A culture of strain CSJ3.*gtaC* (open squares), which is also a L16601 derivative carrying the *gtaC* mutation but overexpressing YueB was equally studied. The curves fitting the initial exponential decay of free phages (dashed lines) were used to calculate irreversible adsorption constants (k_{ads}). For clarity a single fitting curve is represented for the data of L16601, *ypfP* and CSJ3.*gtaC*. The plotted values are the average of at least three independent measurements with indication of standard deviation.

Interestingly, when the *gtaC* mutation was tested in a *yueB*-overexpressing background (CSJ3 as parental strain, see Table III.1), the SPP1 irreversible adsorption rate raised almost to wild type levels ($k_{ads} = 0.84 \pm 0.20$ min^{-1} . A_{600}^{-1}) as result of a 25-fold increase of k_{ads} relatively to that measured in strain L16601.*gtaC* (Fig. III.8). This result suggests that a deficiency in reversible adsorption can be compensated by increasing the surface concentration of YueB.

III.3.5.2 – Analysis of YueB production in cell wall mutants

A negative effect of cell wall mutations in k_{ads} could be explained if they affect YueB production. Although no evidence or previous reference pointed for the interference of cell mutations in the production of membrane proteins, we decided to rule out this hypothesis by analyzing YueB production in the tested cell wall mutants. As judged by western-blot analysis, none of the mutations affected the level of YueB production (Fig. III.9).

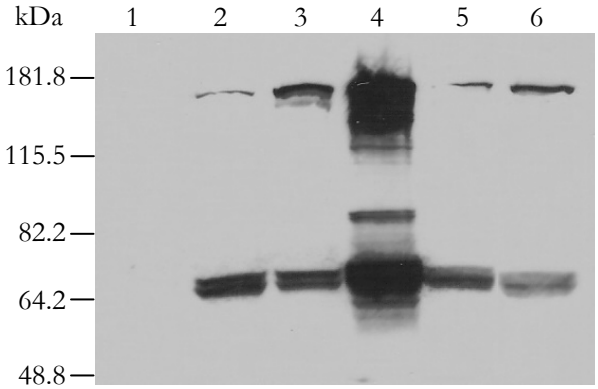


Figure III.9 | Western blot analysis of YueB polypeptides in the protein extracts obtained from the different *B. subtilis* cell wall mutant strains. Strains are designated accordingly to their mutation. Lanes: 1, CSJ1; 2, L16601; 3, L16601.*gtaC*; 4, CSJ3.*gtaC*+IPTG; 5, L16601.*tagE*; 6, L16601.*ypfP*. L16601 is the wild type strain; CSJ1 is a *yueB* non-producing strain.

III.3.5.3 – Analysis of SPP1 infection

We have also studied the effect of reversible adsorption mutations in SPP1 efficiency of infection in solid medium. All mutations allowed SPP1 platting although with distinct impact in phage efficiency of plating (E.o.P.) and plaque morphology (Fig. III.10). In *gtaC* and *gtaB* mutants the E.o.P. decreased to less than 50% along with a clear reduction of phage plaque size. Strain L5703 behaved similarly, while with a smaller impact in E.o.P. and plaque morphology. Overproduction of YueB in a *gtaC* background (strain CSJ3.*gtaC*) led to a clear increase of plaque size although the E.o.P. did not rise as would be expected. The remaining studied mutations did not affect significantly E.o.P. and plaque morphology, although some heterogeneity in plaque size could be observed in strains carrying the *tagE* mutation. According to its K_{eq} value we would expect a platting phenotype in the double-mutant *tagE*/*ypfP* similar to that observed in *gta* mutants. We observed however that it resembled the *tagE* mutant instead, a result that seems to agree with the analogous irreversible adsorption profiles obtained with *tagE* and double-mutant strains (see Fig.

III.8). The effect on E.o.P. and plaque size seemed thus also to correlate with the level of reversible adsorption inhibition imposed by the different mutations, except again for the strains carrying the *ypfP* mutation. A possible explanation for the apparent contradictory results obtained with strains harbouring the *ypfP* mutation is presented in Discussion (III.4).

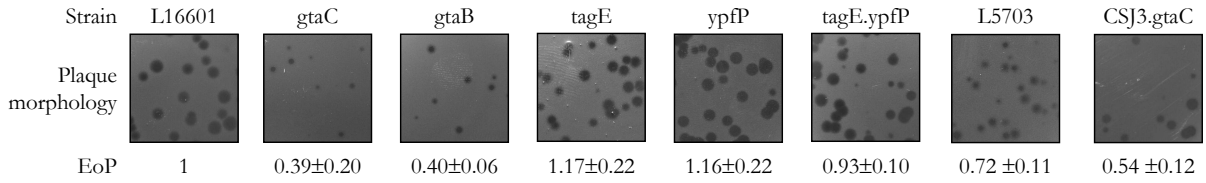


Figure III.10 | SPP1 efficiency of plating and phage plaque morphology in *B. subtilis* cell wall mutants.

In summary, the results presented above clearly show that reversible adsorption has a major contribution for the whole adsorption and infection processes. Interference with reversible adsorption has an impact on the phage E.o.P. and plaque morphology that depends on the type of cell wall mutation. The importance of reversible adsorption is particular evident in liquid media, where its inhibition can lead to a much reduced rate of SPP1 irreversible binding to YueB.

III.3.6 – FAST SPP1 DISSOCIATION FROM REVERSIBLE RECEPTORS ALLOWS RAPID YUEB RECOGNITION

We showed above that temperature has a great impact in the dissociation rate of SPP1 from reversible receptors, with k_2 increasing exponentially with *T*. Intuitively we would tend to think that high temperatures would decrease the rate of irreversible binding by shifting the equilibrium towards free phages (expression (1)). We have studied the effect of temperature on SPP1 irreversible binding to the wild type strain and to a strain virtually devoid of reversible adsorption but overproducing YueB (CSJ3.*gtaC*), by performing adsorption assays and extracting k_{ads} values as in Figure III.8. The temperatures tested were 18, 24, 30 and 37ºC because within this range k_1 is basically constant (see Fig. III.4B).

The results actually showed that irreversible adsorption to the wild type strain was clearly accelerated with the increment of *T*, whereas a much less pronounced increase of k_{ads} was observed with strain CSJ3.*gtaC* (Fig. III.11). By performing simultaneous measurements of total and irreversible adsorption to CSJ3.*gtaC* (as in Fig. III.3B) we observed that SPP1 interaction

with this strain is essentially irreversible (not shown), meaning that once bound to YueB phages are no longer recoverable.

These results strongly suggest that the major temperature-limited step in the whole process of SPP1 adsorption is the phage release from reversible receptors. High *T* promotes faster dissociation rates which are required for rapid recognition of the YueB receptor.

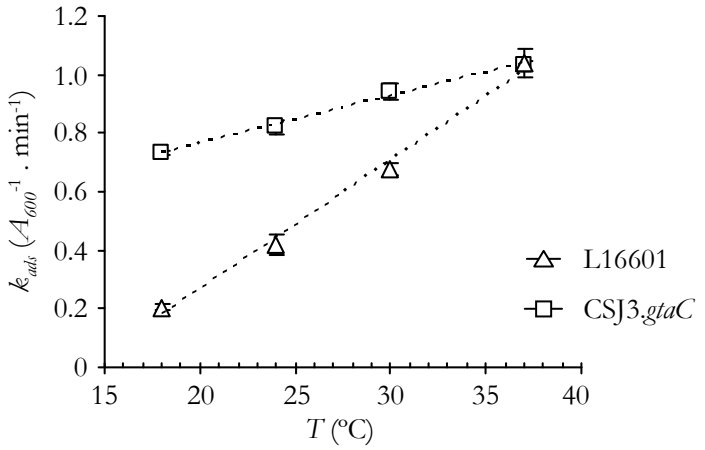


Figure III.11 | Effect of temperature on SPP1 irreversible adsorption. Cultures of strains L16601 and CSJ3.*gtaC* were grown to an A_{600} of 0.8 and equilibrated at 18, 24, 30, and 37°C before SPP1 addition. Irreversible adsorption constants (k_{ads}, from three independent assays) were extracted from exponential fitting curves as described in Figure III.8. Error bars are represented.

III.4. DISCUSSION

As far as we know the only suggestion of the involvement of glucosylated teichoic acids in SPP1 adsorption comes from the work of Yasbin *et al.* (1976). This work showed that the percentage of SPP1 total adsorption to *B. subtilis* strains carrying a *gtaC*, *gtaB* or a *gtaA* (i.e., *tagE*) mutation was reduced when compared to the wild type strain. Somewhat contradictorily however, SPP1 could still plate in these mutants with only a slight decrease of E.o.P., whereas the same mutations led to resistance to other *B. subtilis* phages like ϕ29 and SP01. The ensemble of our work provides an explanation for the results of this early study.

In this chapter we showed that in fact SPP1 binds specifically to *B. subtilis* cell wall polymers, but in a fully reversible way. We also demonstrated that SPP1 binding to the membrane protein YueB, the receptor that triggers viral DNA ejection, can still occur in the absence of the reversible adsorption step, although at considerably lower rates. This explains in the one hand the SPP1 capacity to plate in *B. subtilis* mutant strains devoid of reversible adsorption receptors and, in the other hand, the plaque phenotype observed in these mutants.

III.4.1 – KINETIC STUDY OF SPP1 REVERSIBLE ADSORPTION TO WTA

The genetic analysis and the adsorption studies conducted in this work indicate that SPP1 reversible interaction with the host cell surface primarily involves glucosyl residues of WTA and that receptor activity could not be substituted by glucosylated poly(ribitolphosphate) WTA. In contrary to lactobacilli phages (Räisänen *et al.,* 2007), interfering with the other substituent of the poly(Gro-P) backbone (D-alanine) does not affect SPP1 adsorption. This was indicated by the identical K_{eq} values measured in CSJ1 and in a derivative strain *dltA* (not shown).

We observed distinct reversible adsorption inhibition depending on the mutations impairing glucosylation of WTA. Those preventing the synthesis of UDP-Glc (*gtaC* and *gtaB*) had the most drastic effect as they completely abolished reversible adsorption. A *tagE* strain, which does not catalyse the transference of glucosyl groups from UDP-Glc to WTA, allowed some reversible adsorption although about 40-fold less efficiently than the control strain. The pleiotropic nature of *gta* mutations, which affect the composition of different cell wall polymers, can be the basis of their more severe effect on SPP1 adsorption. Other possible explanation for differences between

gta and *tagE* strains could be a residual level of glucosylation in the latter strain carried out by other cell glucosyltranferase(s) acting of the available pool of UDP-Glc. A highest negative effect in adsorption efficiency of *gtaC* and *gtaB* mutations when compared to mutation *tagE* was also previously reported for *B. subtilis* phage ϕ25 (Givan *et al.,* 1982).

The results presented here clearly showed that reversible adsorption has a major contribution for the whole SPP1 adsorption and infection processes. The cell wall mutations that decreased reversible adsorption affected proportionally SPP1 irreversible adsorption and plating efficiency, except for mutation *ypfP* (see Figs. III.8 and III.10). Based on K_{eq} determinations this mutation produced a ~4-fold decrease of SPP1 reversible adsorption. Its transfer to a *tagE* background further reduced the fraction of phage/bacteria complexes at the equilibrium state. Interestingly, the *ypfP* mutation alone had little impact on the rate of irreversible adsorption and no obvious effect on phage E.o.P. and plaque morphology. Moreover, regarding these features, the *tagE* or the double-mutant *tagE/ypfP* produced essentially the same results despite their significantly different K_{eq} values. These apparent contradictory results might be explained if we assume that mutation *ypfP* stimulates phage release but has a minor impact in phage adsorption. In this situation the *ypfP* mutation would mimic the effect of rising the temperature, which we have shown to markedly increase phage dissociation (k_2) without significant interference in phage adsorption (k_1). An increased rate of phage dissociation will result in a lower K_{eq} value (as observed for strain CSJ1 at 37ºC, Fig. III.4B) but, provided that adsorption can still occur, it will not produce a reduction of irreversible binding. In fact our results indicate that high desorption rates are required for rapid recognition of receptor YueB (see below). We have neither checked this hypothesis nor addressed the possible mechanism by which the mutation *ypfP* could be promoting phage release.

ypfP (or *ugtP*) codes for a multifunctional UDP-glucose:diacylglycerol glucosyltransferase that is required for the formation of the neutral glycolipids of the cell membrane and for the synthesis of diglucosyldiacylglycerol that acts as an anchor for LTA (Salzberg and Helmann, 2008). Recently it was also shown to couple nutritional availability to cell division in *B. subtilis* (Weart *et al.*, 2007). In *B. subtilis* and *S. aureus ypfP* mutants LTA are still anchored to the membrane but by the lipid moiety diacylglycerol instead of diglucosyldiacylglycerol (Jorasch *et al.,* 1998; Kiriukhin *et al.,* 2001). Depending on the *S. aureus* genetic background this modification of the lipid anchor can lead to an increased LTA turnover and release into the culture supernatant (Kiriukhin *et al.,* 2001) or to a reduction of both the content of LTA and the length of the poly(Gro-P) chain

(Fedtke *et al.,* 2007). The available information in *B. subtilis* is that the *ypfP* mutation does not change the incorporation of the poly(Gro-P) moiety of LTA (Lazarevic *et al.,* 2005), although it is not known how it affects the polymers turnover and release to the media. We might speculate that as a result of the *ypfP* mutation, structural, spatial and/or ionic changes of the cell wall environment promote dissociation of SPP1/Glc-WTA complexes, reducing K_{eq} value without affecting irreversible adsorption.

We also demonstrated that the temperature-dependence of the SPP1 dissociation rate from reversible receptors was consistent with an Arrhenius law, allowing the estimation of an activation energy of 22.6 kcal/mol for phage release. Interestingly, this value is comparable to the activation energy required for the YueB triggering of SPP1 DNA ejection *in vitro* (29.5 kcal/mol, Raspaud *et al.,* 2007). It can be therefore of biological significance that phage release from reversible receptors is maximal when the system's energy is close to the value allowing productive interaction with receptor YueB.

III.4.2 – THE ADSORPTION MODEL

The importance of reversible adsorption is particularly evident in liquid media, where its inhibition can lead to a much reduced rate of SPP1 irreversible binding to YueB. The most striking result of this work was the finding that the speed of SPP1 irreversible binding to YueB is mainly determined by the rate of phage release from reversible receptors, in a way that a decrease of k_2 results in a lower k_{ads} value (see Fig. III.11). Taking this fact and the knowledge that fast SPP1 initial adsorption to the *B. subtilis* surface occurs via the reversible interaction with WTA, we propose a model for reversible adsorption as a mechanism employed by SPP1 for the rapid recognition of receptor YueB (Fig. III.12).

In a wild type cell envelope context the vast majority of SPP1 particles will initially adsorb to WTA, most probably because the concentration of these polymers in the cellular surface exceeds by far that of YueB. At the temperature routinely used for *B. subtilis* growth in the laboratory (37ºC) phages will be bound for only a few seconds being immediately released for subsequent adsorption/desorption cycles. This strategy allows a dynamic association with reversible receptors permitting SPP1 to "scan" the cellular surface until it is captured by YueB. At lower temperatures WTA will retain SPP1 for longer periods increasing the time required for YueB

recognition. In the model depicted in Figure III.12 the YueB receptor can either bind directly free SPP1 virus particles (right side) or capture phages that bound to a WTA molecule located in its close vicinity (left side).

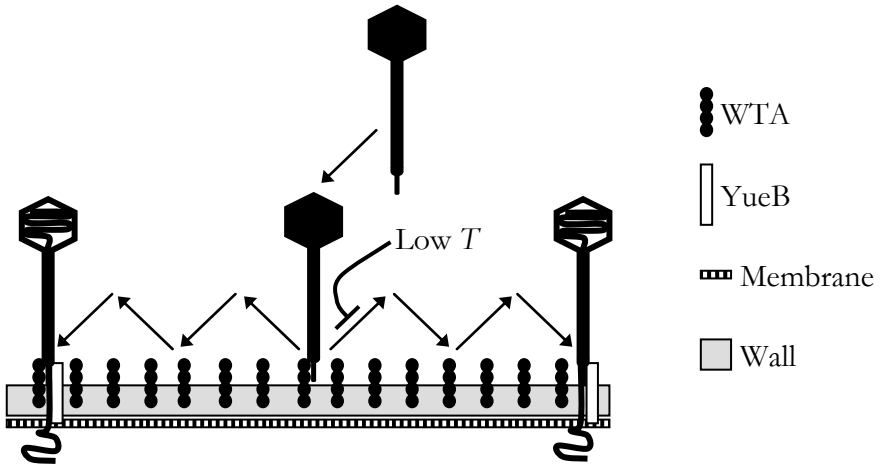


Figure III.12 | Model representation of the SPP1 strategy of adsorption to *B. subtilis* cells. Phage SPP1 initially adsorbs reversibly to the bacterial surface WTA's. At high temperature SPP1 is released from these cell wall polymers after a few seconds and immediately interacts with neighbour WTA molecules. These rounds of adsorption/desorption allow a dynamic association of SPP1 with WTA receptors, providing a mechanism to rapidly find the irreversible receptor YueB. Low temperatures delay YueB recognition by increasing the residence time of SPP1 on WTA.

SPP1 adsorption seems thus to follow the so called strategy of reduction of dimensionality (RD) as a mean to increase the rate of irreversible binding (Adam and Delbruck, 1968; Berg and Purcell, 1977; see III.1). According to RD theory phage particles initially search a bacterium by means of three-dimensional movements. Association with reversible receptors imposes then movement restrictions allowing phages to diffuse only laterally, like in a two-dimensional (2D) fluid, until they encounter the irreversible receptors. This reduction of dimensionality results in enhancement of the binding rate to irreversible receptors. As observed by Berg and Purcell (1977) 2D diffusion is advantageous to binding only if reversible association is strong enough to keep phage on the bacterial surface but, at the same time, weak enough to allow phage lateral jumping. This requirement seems to be fulfilled by the *B. subtilis*/SPP1 system.

Furthermore, the working system is best explained according to the modified sequential theory proposed by Christensen (see III.1), with some small addendums. We would say that there are not two types of reversible binding. Instead what we believe that happens is that, if the reversible binding occurs within a limited radius from the irreversible receptor, the irreversible binding has a higher probability to occur than a subsequent reversible step. This is compatible with our work

that not only shows that the adsorption/desorption cycles diminish the time lapse for irreversible adsorption, but also that that step is avoidable. Finally, the overall representation could be: PB*←P+B↔PB→PB* according to models presented in section III.1.1.

III.4.3 – COMPARISON TO OTHER PHAGE SYSTEMS

The model here presented probably applies to other phages with a two-step adsorption process involving two different cellular receptors. The lactococcal phage c2 uses wall carbohydrates for reversible adsorption and the YueB homologous protein Pip for irreversible binding (Geller *et al.,* 1993; Monteville *et al.,* 1994), whereas phage T5 targets, respectively, the outer membrane components lipopolysaccharide (LPS) and FhuA. Although the lack of the reversible interaction step with LPS has no effect in T5 E.o.P., its occurrence was shown to accelerate adsorption by a factor of 15 in liquid media (Heller and Braun, 1979; 1982). In our case, the rate of SPP1 irreversible adsorption to the *B. subtilis* wild type strain was about 30 and 8-fold higher than the rate measured in *gtaC* and *tagE* mutants, respectively. In contrast, the routinely used phage λ, which apparently targets a single receptor (LamB) for both steps of adsorption, seems to follow a completely different strategy. In this case, reversible adsorption (k_1) was shown to be weakly dependent on the temperature whereas phage dissociation (k_2) and irreversible binding (k_3) are completely unaffected by this parameter (Moldovan *et al.,* 2007). However, the kinetics of adsorption of the wild type phage λ, which displays distal tail fibers, may be substantially different as the phage may target other receptors (like LPS) in addition to LamB.

It is worth to note that adsorption of other *B. subtilis* phages like the *Podoviridae* ϕ29 and the *Myoviridae* ϕ25 seems also to involve glucosylated WTA and a yet unidentified membrane receptor (Young, 1967; Jacobson and Landman, 1977; Yasbin, 1976). These phages exhibit however clear deviations from the SPP1 adsorption mechanism. Although adsorption of phage ϕ29 to isolated cell walls is completely reversible, mutations impairing glucosylation of WTA result in phage resistance (Young, 1967; Jacobson and Landman, 1977). Mutations *gtaB* and *gtaC* also confer resistance to ϕ25 (Yasbin, 1976). In addition, this phage cannot be released as an infective virus particle after adsorption to wild type cell walls as the interaction triggers the tail contraction, although without phage DNA release (Young, 1967). It seems therefore that in these phages the reversible step is essential to allow the subsequent interaction with membrane receptors (that may be the same). On the contrary, the ability of SPP1 to bind YueB and infect host cells without

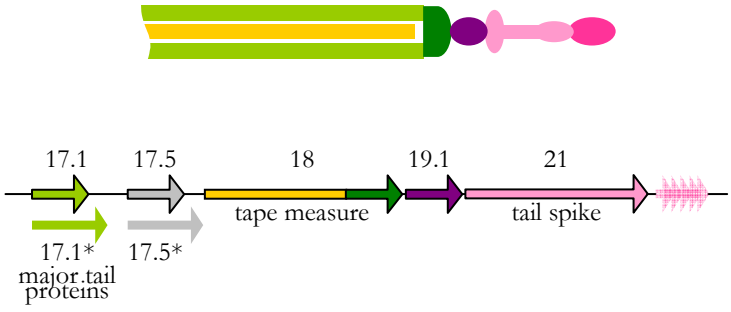
previous adsorption to glucosylated WTA (as in *gtaC* mutants) suggests that interaction with this polymer is not an essential prerequisite that "activates" the phage for the subsequent interaction with YueB. The fact that overexpression of YueB in a *gtaC* strains accelerates irreversible adsorption to wild type levels supports the idea that the major role of reversible adsorption is to favour the encounters between SPP1 and YueB.

PART IV

PHAGE SPP1 RECEPTOR BINDING PROTEINS

IV.1. INTRODUCTION

Tails from tailed phages are crucial for the specific binding to bacterial receptors and for the viral DNA channeling from the capsid to the host cell cytoplasm. The tail structures that mediate phage adsorption carry proteins with affinity to cell receptors and are generally designated as receptor binding proteins (RBP). The interaction between RBP and host receptors triggers conformational changes in the virion structure that culminate in the opening of the capsid and in the exit of the phage DNA through the lumen of the tail tube. The structure of SPP1 tail and tail tip was recently clarified by Plisson *et al.,* 2007. The authors described a 160 nm long tail tube that serves as a flexible link between the phage capsid and the 31 nm long tail tip. The two major tail proteins (MTPs) that compose the tail helical structure are gp17.1 and gp17.1*, arranged with a defined stoichiometry of ~3:1 (Auzat *et al.,* 2008). Gp17.1* is generated through a programmed translational frameshift at the 3' end of gene *17.1* (Fig. IV.1). The amino terminus shared by the two MTPs is responsible for the assembly of the tail tube whereas the extra carboxyl terminus 9.1 kDa of gp17.1* shield the structure and is exposed to the milieu (Auzat *et al.,* 2008). Gp18 fills the lumen of the tube (tape measure protein) and probably also forms the cap structure that closes the tube (Fig. IV.1). As far as the tip and the rod distal portion of the tip are concerned, gp19.1 and gp21 are predicted to be structural components together with yet unidentified proteins (Plisson *et al.,* 2007). Given its position in the tail structure and previous work from São-José *et al* (2006) and Plisson *et al.* (2007) the SPP1 tail tip as been assigned as the bacterial receptor recognition apparatus.



Adapted from Plisson *et al.*, 2007

Figure IV.1 | Schematic representation of SPP1 tail proteins. Gene organization on the SPP1 genome is shown in the lower part of the scheme. Genes coding for structural tail proteins are represented as colored arrows, keeping the same color code as in the upper scheme; gray arrow represents a non-structural gene coding for protein assisting tail assembly. The gene *18* segments encoding gp18 N- and C-terminus are shown in gold and dark green, respectively, to illustrate that these two regions form the tail tube internal tape measure and the cap. The constituents of most distal tip structure (represented as strong rose), known as the tip rod, are still yet unidentified. Proteins with a putative function in tail assembly are represented as semitransparent arrows.

The tail tip can be divided morphologically into three regions: a sphere-like region, a broad flattened domain and a terminal rod (violet, rose and pink, respectively, in Fig. IV.1). The positions of gp19.1 and gp21 in the SPP1 tip proposed by Plisson *et al.* (2007) (Fig. IV.1) were derived from the organisation of homologous proteins mapped by immunoelectron microscopy in the tip of the SPP1-related phage Tuc2009 (McGrath *et al.,* 2006). The primary sequence of gp19.1 shares homology with the protein where the base plate is anchored in the tail tube of phages Tuc2009 (Orf49; E value = 5 e^{-7}) and TP901-1 (Orf46; E value = 1 e^{-6}). Both products were assigned as the distal tail proteins (Dit) that play their part in the formation of the initiator complex at the baseplate structure of these phages (Sciara *et al.,* 2008). Plisson *et al.,* 2007 showed that the predicted secondary structure of gp19.1 is compatible with that of the head-binding domain of the *Salmonella typhimurium* P22 phage tail spike (1LKT; Steinbacher *et al.,* 1997), and that the documented structure of this protein fits well the density map of the SPP1 sphere-like region of the tip (Fig. IV.1). Blast and PSI-Blast searches of gp21 sequence showed that the N-terminal region is homologous to host specificity proteins (CDS53 from BCJA1c phage of *B. clarkii*, E value = 3 e^{-94}), putative and assigned tail fibre proteins (Orf97, YP_164732, from *Lactobacillus plantarum* bacteriophage LP65, E value = 3 e^{-7}) and putative and assigned tail associated lysozyme (Tal) proteins (Orf50, Tal, from Tuc2009, E value = 1 e^{-18}; Orf47, Tal, from TP901-1, E value = 1 e^{-17}; Orf43, ABI54246, from *Lactococcus* phage P335, E value = 2 e^{-16}). The C-terminal region of gp21 shares no relevant homologies. Plisson *et al.* (2007) proposed that there is a similarity between the expected secondary structure of gp21 and that of the main tail spike domain (β-helix domain) of phage P22, 1TSP, which is responsible for O-antigen recognition and hydrolysis (Fig. I.4; Steinbacher *et al.,* 1994; Baxa *et al.,* 1999).

Although the available evidence suggests that gp21 is the tail tip component responsible for YueB recognition (São-José *et al.,* 2006; Plisson *et al.,* 2007) the direct interaction between these two proteins has never been experimentally demonstrated. The goal of the work presented in this chapter was to identify SPP1 proteins or tail structures that specifically target the *B. subtilis* receptors (YueB and Glc-WTA), and by this way confirm the role of the tail tip in adsorption.

IV.2. MATERIALS AND METHODS

IV.2.1 – BACTERIAL STRAINS, BACTERIOPHAGES AND PLASMIDS

All biological material used in this work is listed in Table IV.1.

Table IV.1 | *E. coli* and *B. subtilis* strains, bacteriophages, vectors and plasmids used in this work.

Strains, phages and plasmids	Genotype or relevant features	Reference
***E. coli* strains**		
TG1	*SupE thi-1* Δ(*lac-proAB*) Δ(*mcrB-hsdSM*)*5* (r_K^- m_K^-) [*F' traD36 proAB lacI*^{q} *Z*Δ*M15*]; Cloning strain	Stratagene
BL21	*E. coli* B F$^-$ *dcm ompT hsdS*(r_B^- m_B^-) *gal*; Protein expression strain	Stratagene
ECB-MBP21	BL21 carrying pCB21-MBP; Amp^R	This work
ECB-GST21	BL21 carrying pCB21-GST; Amp^R	This work
ECB-MBP21-Chap	BL21 carrying pCB21-MBP and pKJE3; Amp^R; Cm^R	This work
ECB-GST21-Chap	BL21 carrying pCB21-GST and pKJE3; Amp^R; Cm^R	This work
***B. subtilis* strains**		
L16601	*B. subtilis* 168; SPP1 indicator strain	Margot and Karamata, 1996
HA101B	Suppressor strain	Okubo and Yanagida, 1968
CSJ3	L16601 derivative overexpressing *yueB* upon IPTG induction; Ery^R	São-José *et al.,* 2004
SP1154	*B. subtilis* wild type YB886 derivative with a replicative plasmid, pIA2, carrying SPP1 gene *21* with a N-terminal fusion with a histidine tag; Cm^R	I. Auzat, unpublished results
CSJ1.*gtaC*	CSJ1 derivative; *gtaC*ΩpBS640; Cm^R	This work
***B. subtilis* phages**		
SPP1	Lytic phage	Riva *et al.,* 1968
SPP1*sus31*	SPP1 derivative carries mutation CAG to TAG (amber) replacing codon Gln31 by a stop codon in gene *13*.	Becker *et al.,* 1997
SPP1*sus45*	SPP1 derivative that carries mutation CAA to TAA (ochre) replacing codon Gln133 by a stop codon in gene *17.1*.	Auzat *et al.,* 2008
SPP1mut3	SPP1 derivative that shows no reversible adsorption; carries a single nucleotide substitution in gene *21*	This work
SPP1mut40	SPP1 derivative that shows no reversible adsorption; carries a single nucleotide substitution in gene *21*	This work

SPP1rec3.43 SPP1rec3.45 SPP1rec3.47	SPP1mut3 derivatives with wild type levels of reversible adsorption; carry wild type gene *21*	This work
SPP1rec40.16 SPP1rec40.17 SPP1rec40.92	SPP1mut40 derivatives with wild type levels of reversible adsorption; carry wild type gene *21*	This work
Vectors and plasmids		
pMAL-c2	Expression vector; allows production of N-terminal MBP-tagged fusion proteins upon IPTG induction; Amp^R	New England Biolabs
pGEX-4T-1	Expression vector; allows production of N-terminal GST-tagged fusion proteins upon IPTG induction; Amp^R	Amersham Biosciences
pG-KJE3	Chaperone expression plasmid; allows expression of GroEL, GroES, DnaK, DnaJ and GrpE upon arabinose induction; Cm^R	Nishihara *et al.,* 1998
pCB21-MBP	pMAL-c2 derivative expressing MBP-gp21; Amp^R;	This work
pCB21-GST	pGEX-4T-1 derivative expressing GST-gp21; Amp^R	This work

IV.2.2 – OLIGONUCLEOTIDES

The oligonucleotides used in this work are presented in Table IV.2. Oligonucleotides numbered from 28758 to 31001, were used in gene *21* sequencing reactions. Primer pair gp21EcoFwd/gp21SalRev was used for gene *21* cloning in vectors pMAL-c2 and pGEX-4T-1.

Table IV.2 | Oligonucleotides used in this work.

Oligonucleotide	Sequence (5'→3')[a]
gp21EcoFwd	AGA**gaattc**TCAAAAAACATATGGATAATGAGCAAC
gp21SalRev	AAT**gtcgac**TTAGTTCACTTCATAGAAAATTAAAAATC
28758	CCCATCCAGCTACATCGCAGG
30125	CCGAGATTAAAGATTATGG
30625	GCGTGTGAATACTATCAAC
30686	GGTTCACAAGTGGCTATG
31005	GGCAACTAACAGCGGATGGG
31011	CATAGCCACTTGTGAACC

[a] Boldface lowercase letters indicate endonuclease restriction sites (*Eco*RI-**gaattc**; *Sal*I-**gtcgac**).

IV.2.3 – GROWTH CONDITIONS AND MICROBIOLOGICAL METHODS

IV.2.3.1 – *E. coli*, *B. subtilis* and bacteriophage growth conditions

E. coli and *B. subtilis* growth conditions, media supplements and antibiotic concentrations were as described in section II.2.3.1 and II.2.3.2. Strains producing YueB780 were grown at 30ºC before thermal induction of protein production (section IV.2.6.1). Strains producing gp21 fusion proteins were grown overnight at 28ºC (section IV.2.6.6). Phage propagation and purification was also as described previously (II.2.3.2 and II.2.3.3).

IV.2.3.2 – Measurement of SPP1 adsorption to *B. subtilis* cells

For SPP1 and derivatives, total and irreversible adsorption was measured as described in II.2.3.4 and III.2.3.2.

IV.2.3.3 – Preparation of SPP1*sus* lysates

In day one, a suspension of the mutant phages SPP1*sus31* and SPP1*sus45* was titrated in the permissive *B. subtilis* strain HA101B. Strain HA101B was always cultured in the minimal medium MIII (1 mM $MgSO_4.7H_2O$, 10 mM $(NH_4)_2SO_4$, 10 mM Na-citrate.$2H_2O$, 8 mM K_2HPO_4, 4 mM KH_2PO_4, 1.5 mg/mL asparagine) supplemented with leucine and tryptophan (50 μg/mL), 0.15 mM $MnSO_4$, 15 mM $MgSO_4.7H_2O$, 0.7% glucose, 0.05% casein hydrolysate and 150 mM $FeCl_3$. On the 2nd day, one isolated plaque of each phage was collected and resuspended in 0.5 mL of TBT. The suspensions were titrated both in strain HA101B and in the non-permissive host L16601. The later served to control the percentage of revertants in each step of amplification. On the 3rd day, phages from confluent plates in strain HA101B were recovered by resuspension in 3 mL of LB medium. The phage suspension was collected after 2 hours incubation at room temperature, with gentle shaking. Cells were removed by centrifugation and the lysate titrated as described above. On the 4th day a medium scale liquid lysate (100 mL) was prepared. For that, strain HA101B was grown to A_{600}=0.85, supplemented with 100 mM $CaCl_2$, infected (i.m.=0.8) and incubated at 37ºC with shaking. After 2 hours incubation, 1% of chloroform and a pinch of lysozyme were added to promote complete lysis. After an incubation period of 10-15 minutes at 37ºC the infection mixture was centrifuged to remove cells and the supernatant filtrated with a

0.45 μm disposable filter. The final lysates were again titrated in the permissive and in the non-permissive host and contained less than 1% of revertants.

IV.2.4 – ISOLATION OF SPP1 REVERSIBLE ADSORPTION MUTANTS

For the isolation of SPP1 mutants impaired in reversible adsorption we took advantage on our knowledge on the kinetics of SPP1 reversible adsorption (see III.3.1 and III.3.2 for details). Essentially we have selected and amplified phages that remained free after 1 hour of contact with highly dense cell suspensions of CSJ1 equilibrated at 0ºC. We have shown that these conditions highly favoured the maintenance of bacteria/phage complexes. The procedure used is summarized in Figure IV.2.

Fifty isolated phage plaques from the final lysate obtained after 4 rounds of enrichment for non-adsorbing phages (Fig. IV.2) were recovered and resuspended in 500 μL of TBT + 0.01% gelatine + 10% chloroform; 20 μL were used to make a small-scale infection by incubating with 20 μL of a overnight culture of *B. subtilis* strain L16601 and 2 mL of LB supplemented with 15 mM $CaCl_2$. Infection was let to occur at 37ºC with agitation and cell lysis was monitored in each tested isolated plaque comparing to a similar infection prepared with WT SPP1. From the 20 infection tubes that showed a delayed lysis compared to the wild type infection, 10 of them were shown to be initiated with a phage input similar to WT SPP1 (~10^7 pfu/mL), meaning that delayed lysis was not due to a low phage input. Phages SPP1mut1, 3, 5, 10, 13, 15, 20, 30, 35, and 40 were selected for confirmation of low reversible adsorption and gene *21* sequencing.

IV.2.5 – ISOLATION OF GENE *21* RECOMBINANT PHAGES FROM PHAGE MUTANT POPULATION

For the selection of recombinant-gp21 phages, SPP1mut3 and SPP1mut40 (reversible adsorption mutants) were used to infect *B. subtilis* strain SP1154, which carries a wild type copy of gene *21* in a replicative plasmid. Infection of the wild type *B. subtilis* strain L16601 was carried out in parallel to account for phage reversion. One hundred isolated phage plaques of each infection event were recovered, resuspended in TBT and screened for wild type levels of reversible adsorption by the ratio *P*/*P0*, which was ≤0.02 for a wild type adsorption and ≥0.3 for the mutants (*P0* is the phage

input and *P* the fraction of free phages after 60 minutes of incubation with CSJ1, adjusted to A_{600}=0.8, at 0ºC). Phages that presented a *P/P0* ratio similar to WT adsorption were selected for further adsorption studies and the sequence of gene *21* confirmed for selected phages.

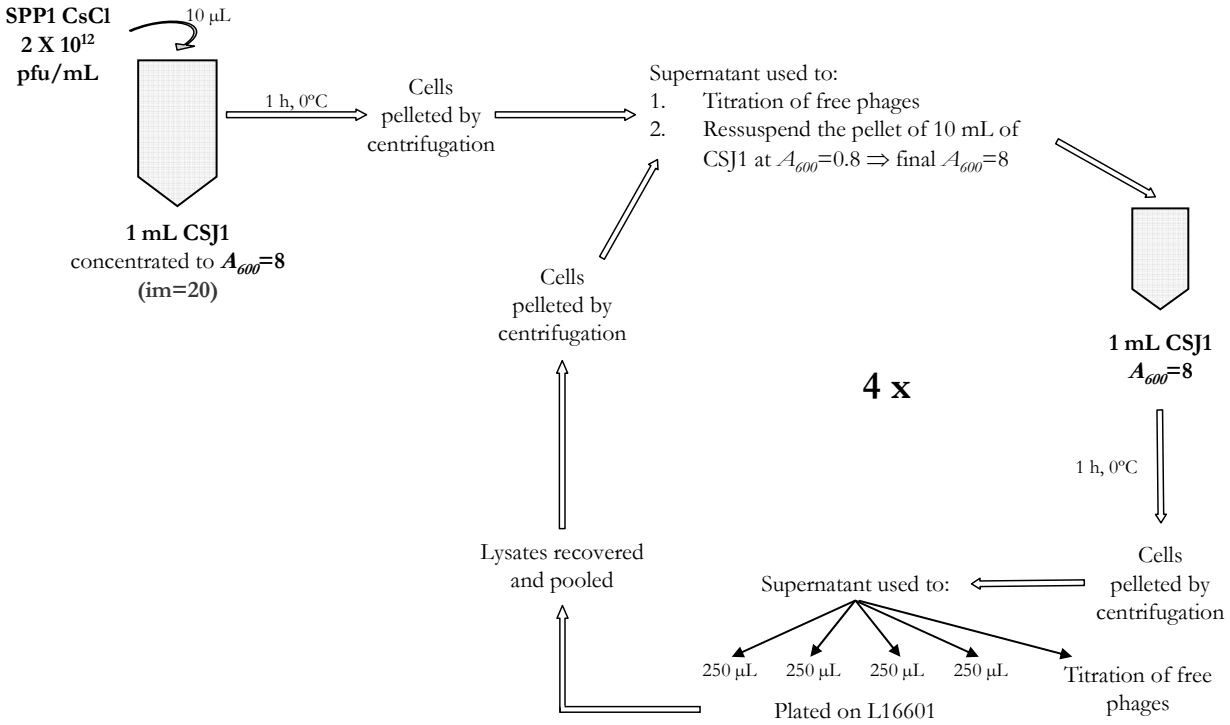


Figure IV.2 | Schematic representation of the protocol used to isolate SPP1 reversible adsorption mutants. Note that SPP1 adsorption is not required for efficient platting (see III.3).

IV.2.6 – MOLECULAR AND BIOCHEMICAL METHODS

DNA extraction and separation by gel electrophoresis was as described in II.2.4.1. PCR amplifications, restriction endonuclease digestions, DNA ligations, and DNA purification and extraction from gels were as described in II.2.4.4. All the oligonucleotides used in this work were purchased from Invitrogen and are listed in Table IV.2. Development of *E. coli* and *B. subtilis* competence and transformation was as described by Chung *et al.,* 1989 and Yasbin *et al.,* 1973, respectively.

Phage DNA extraction was performed using CsCl-purified SPP1 phages adjusted to around 10^{11} pfu/mL in TBT (keeping a final volume higher than 300 µL). Phages were incubated for 1 hour at 55ºC in the presence of 50 mM EDTA to promote DNA release. Three rounds of phenol extractions were then carried out by slowly inverting the tube for 5 minutes and centrifuging for

the same time at room temperature. The upper aqueous phase was transferred to a new tube with an equal volume of chloroform and, after centrifugation in the same conditions the purified DNA was dialyzed overnight against TE.

To monitor protein production by *E. coli* strains total extracts were produced as described in II.2.4.6. For phage protein extraction, purified SPP1 virions were incubated with 50 mM EDTA at 55ºC (Tavares *et al.,* 1996). After 1 hour, 100 mM $MgCl_2$ and 10 U of benzonase (Roche Applied Science) were added, and the samples were incubated for 2 hours at 37ºC for DNA digestion. For protein analysis, an equal volume of 2X SDS-PAGE loading buffer was added to the sample. Analysis by SDS-PAGE was as described previously (II.2.4.7). Proteins were visualized either by Coomassie Blue staining (Sambrook and Russel, 2001) or following the Rapid Silver Staining Method (Colligan *et al.*, Current Protocols in Protein Science).

Western blot analysis of YueB780 polypeptides was as described in II.2.4.7. Immuno-detection of SPP1 virion proteins (anti-SPP1 serum) and of the SPP1 tail protein gp21 (anti-gp21 N-ter serum, a gift of Dr. Isabelle Petitpas) was performed with rabbit polyclonal sera diluted 1:10,000. Maltose binding Protein (MBP) and Glutathione S-transferase (GST) fusion proteins were detected using, respectively, the Anti-MBP antiserum (NEB) and the Anti-GST antibody (Amersham Biosciences), following the suppliers' instructions. Antigen/antibody complexes were detected with the Chemioluminescence Western Blotting Kit (Roche Applied Science), or with the ECL (ECL Western Blotting Detection Reagents, GE Healthcare Life Science) and secondary antibody (Anti-Rabbit IgG and F(ab)2 fragment-peroxidase developed in Goat; Sigma), according to the manufacturer's instructions.

IV.2.6.1 – YueB780 purification

Production of the His_6-tagged receptor version YueB780 was basically as described in São-José *et al.,* 2006. Cellular pellets from 500 mL of induced cultures were resuspended in 10 mL of YueB Buffer A (300 mM NaCl, 50 mM HEPES, 50 mM imidazole, pH 8) and sonicated on ice with 4-6 cycles of 1 minute (pulse 3; 60%; 36-40 W, 1 minute pause between cycles) in a Vibra Cell sonicator (Bioblock Scientific). A 200 μL sample was kept for analysis of crude extract. Cell debris and insoluble material were removed by centrifugation and the supernatant filtered through a pore size of 0.22 μm.

YueB780 was purified from the cleared extracts by affinity chromatography using a HisTrap™ HP column (Amersham Biosciences) coupled to an Äkta FPLC system (Amersham Biosciences). YueB780 was eluted in a single step with 185 mM imidazole (30% of YueB Buffer B: 300 mM NaCl, 50 mM HEPES, 500 mM imidazole, pH 8). The fraction corresponding to the peak from the affinity step was further purified by gel filtration in a Superose 6 column (Amersham Bioscience) and using YueB Buffer (without imidazole) as solvent.

IV.2.6.2 – Production of SPP1*sus* total protein extracts from infected cells

A 10 mM $CaCl_2$ supplemented culture of *B. subtilis* wild type strain, A_{600}=0.8, was infected either with SPP1*sus31* or SPP1*sus45* lysate in order to have an i.m.=5, in a 200 or 500 mL final volume. After an incubation period of 10 minutes at 37ºC with shaking, the infection mixtures were centrifuged at room-temperature, using a pre-warmed rotor. The cellular pellets were resuspended in 1/10 volume of pre-warmed LB medium, and incubated in the same conditions until 1 hour post-infection. To complete cell lysis, 100 μg/mL lysozyme, 300 mM NaCl, 50 mM imidazole, 0.01 mg/mL DNase and EDTA-free Anti-Protease Cocktail (Roche) were added to the lysates, which were then incubated on ice for 10 minutes. After 3x1 minute sonication (pulse 3, amplitude 80%, 33-37 W), the crude extracts were cleared by centrifugation.

IV.2.6.3 – Binding assays of SPP1*sus* proteins to a YueB780 affinity resin

The SPP1*sus* crude extracts (see previous section) were used in YueB binding assays. To deplete the extracts from proteins that non-specifically bound to Ni-NTA agarose beads, 5% of glycerol and 500 μL of Ni-NTA Agarose beads (Qiagen) washed with LB, were added to the extracts. The extracts were left at 4ºC on a rotating wheel for 1 hour to saturate the Ni-beads. After, the beads were removed by centrifugation and the cleared SPP1*sus* extracts were assayed for YueB binding capacity by two different approaches.

In a first assay, one milliliter of Ni-NTA Agarose beads (Qiagen) was washed with ejection buffer (TBT with 300 mM NaCl; São-José *et al.,* 2006) and resuspended in 900 μL ejection buffer, 50 mM imidazole and 0.117 mg of YueB780 (670 pmol of dimers; purified as described in section IV.2.6.1), and incubated on a rotating wheel at 4ºC for at least 1 hour. The rationale for the YueB780 amount was the following. In the conditions of infection we had 10^8 cfu/mL in a total volume of 200 mL. In a SPP1 wild type infection those cells would release an average of 200 pfu

per infected cell. This means that a total of approximately 4 x 10^{12} pfu would be obtained. These correspond to 6.64 x 10^{-12} moles of SPP1 or 6.6 pmol. São-José *et al.,* 2006 showed that 99% of the SPP1 pfu were inactivated by YueB780 at a receptor/phage molar ratio of ~100:1. To guarantee this ratio in our binding assays we bound 670 pmol of YueB780 dimers to Ni-NTA beads (a slight excess of the theoretical 660 pmol).

The Ni-YueB780 matrix, prepared as described above, was poured onto a Econo-Pac Mini column (Bio-Rad) coupled to a peristaltic pump, and washed with 2 column volumes of ejection buffer supplemented with 50 mM imidazole. The SPP1*sus* extracts were gently transferred to the top of the matrix and the first flowthrough was collected and re-passed (an aliquot was kept for analysis). Approximately 1 column volume of ejection buffer+50 mM imidazole was used to wash the matrix (an aliquot of the first and last milliliters was collected). The complexes of YueB780/SPP1*sus* proteins were eluted by gravity flow with 1.2 mL of ejection buffer supplemented with 500 mM imidazole. The eluate was fractionated in 200 μL samples. For protein analysis, 20 μL of each fraction was used.

In the second assay we only used a cleared 500 mL SPP1*sus31* crude extract prepared as described above. The YueB780 was bound to a HisTrap™ FF (GE Healthcare, Uppsala-Sweden) coupled to an Äkta FPLC system (Amersham Biosciences). The matrix was washed with two column volumes of ejection buffer supplemented with 50 mM imidazole (column buffer). The pre-filtered SPP1*sus31* lysate was passed through the column twice at a flow of 1 mL per minute. Proteins bound to the YueB780-matrix were eluted in a stepwise manner by passing increasing percentages of elution buffer (ejection buffer supplemented with 500 mM imidazole). The 1.5 mL relevant samples were concentrated 10X with Vivaspin2 concentrators (VivaScience, Sartorius Stedim Biotech) and 15 μL used for analysis.

IV.2.6.4 – Analysis of SPP1*sus* structures by linear sucrose gradient

A linear 5-20% sucrose gradient was used to separate protein and protein complexes of the selected fraction from the first binding assay (see above). The gradient was made with a Radian™ gradient maker (BioComp Instruments). The sucrose solutions were prepared in 0.5X ejection buffer. After centrifugation in a Beckman SW 55 rotor (35000 rpm, 3 hours, 4ºC) fractions of 750 μL were collected starting from the bottom of the tube were taken with the help of a peristaltic pump. For protein analysis, 20 μL of each fraction were used.

IV.2.6.5 – MALDI TOF Mass Spectrometry analysis

The proteins from selected fractions of the SPP1*sus31*-YueB780 binding assay (section IV.2.6.3) were separated in a 10% SDS-PAGE, bands were recovered and submitted to complete digestion with trypsin followed by mass spectrometry analysis by Matrix-assisted laser desorption/ionisation-time of flight (MALDI-TOF). MALDI-TOF MS analysis was executed by Dr Jean-Pierre LeCaer in the ICSN, Gif-sur-Yvette, France. Peptide mapping analysis was performed using the ProFound tool by Dr. Isabelle Auzat from CNRS, Gif-sur-Yvette, France (http://bioinformatics.genomicsolutions.com/service/prowl/profound.html).

IV.2.6.6 – Cloning and heterologous expression of gp21 fusion proteins

For cloning and overexpression of gp21, the coding sequence of gene *21* was PCR amplified with Pfu polymerase (Stratagene) and the primer pair gp21EcoFwd/gp21SalRev (Table IV.2). The PCR product was cloned in vectors pMAL-c2 and pGEX-4T-1 so that its 5' end was fused in frame to the 3' end of sequences coding MBP (Maltose Binding Protein; NEB) and GST (Glutathione S-Transferase; Amersham Biosciences). The resulting plasmids pCB21-MBP and pCB21-GST were isolated and amplified in *E. coli* strain TG1. Protein overproduction was performed in the *E. coli* protein expression strain BL21. To improve the solubility of the fusion proteins, strain BL21 carrying plasmid pG-JKE3 that expresses a set of chaperone proteins (GroEL, GroES, DnaK, DnaJ, GrpE) upon arabinose induction (Table IV.1) was also used. *E. coli* BL21 strains carrying plasmids coding for MBP-gp21 or GST-gp21 fusion proteins were grown overnight at 28ºC with the appropriate antibiotics and, in the case of MBP-gp21, with 0.2% glucose for transcription repression. A fresh isolated colony was re-inoculated 1:100 on the next day and incubated at 37ºC at 150 rpm. Having reached an A_{600}=0.4, cultures were transferred to 25ºC and agitated in the same conditions. For *E. coli* strains containing pKJE3 3 mg/mL arabinose was added 10 minutes before induction of gp21 production to induce the production of chaperones. Production of recombinant gp21 was induced when cultures reached an A_{600} of 0.5-0.6 by adding 0.3 mM IPTG. After 3 to 4 hours of protein production the cells were collected by centrifugation and kept at -20ºC until use.

To produce total protein extracts, the frozen cells were gently resuspended in 10 mL of lysis buffer (50 mM Tris-HCl pH7.5, 150 mM NaCl, 10 mM $MgSO_4$, 0.1% Triton). Cell disruption

was performed as described in section IV.2.6.1. The cleared extracts were used for GST- and MBP-gp21 purification by FPLC (section IV.2.6.7).

IV.2.6.7 – Purification of GST- and MBP-gp21 fusion proteins

The GST-gp21 fusion was purified using a GSTrap FF 5 mL affinity column (Amersham Biosciences), following the manufacturer's instructions. The binding/wash buffer was PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and the elution buffer 50 mM Tris-HCl, 10 mM reduced L-glutathione (Sigma), pH 8.0. The proteins bound to the column were eluted in one step with 100% of elution buffer. The GST-gp21 enriched fractions and other relevant fractions were concentrated with Vivaspin 0.5 mL concentrators (VivaScience, Sartorius Stedim Biotech) and stored at 4ºC until SDS-PAGE/Western blot analysis.

MBP-gp21 fusion proteins were purified with an amylose resin (NEB), following the instructions of the manufacturer. The binding/wash buffer was 20 mM Tris-HCl (pH 7.4), 0.2 M NaCl and 1 mM EDTA. Elution was carried out with the same buffer supplemented with 10 mM maltose (Sigma) in a single step elution. An MBP-gp21 enriched fraction was desalted with the Hi-trap Desalting Column from Amersham Biosciences to remove EDTA and exchange the sample buffer to YueB Buffer A (section IV.2.6.1). A MBP-gp21 enriched sample and other relevant fractions were stored at 4ºC until SDS-PAGE/Western blot analysis.

IV.2.6.8 – Binding of gp21 fusion proteins to a YueB780 affinity resin

The GST-gp21 and the MBP-gp21 enriched fractions were then passed through a YueB780-bound Ni column (section IV.2.6.1) to see if YueB780 had the ability to capture gp21. Elution of potential gp21-YueB780 complexes was made as for YueB780. Fractions were collected and kept at 4ºC until SDS-PAGE/Western blot analysis.

IV.2.7 – BIOINFORMATICS ANALYSIS

Protein homology searches were carried out by BLAST and PSI-BLAST analysis (http://www.ncbi.nlm.nih.gov). Multiple sequence alignments of DNA sequences were performed with ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html). PsiPred was

used for prediction of protein secondary structures (http://bioinf.cs.ucl.ac.uk/psipred/). Prediction of β-helix folds from protein sequence data was performed using BetaWrap (http://groups.csail.mit.edu/cb/betawrap/).

IV.3. RESULTS

It is clear from the previous results presented in this thesis, that SPP1 has two steps of adsorption. The reversible step, which can be studied independently in a CSJ1 background, involves the glucosylated teichoic acids and the irreversible step that involves the membrane protein YueB. Since two types of receptors are sequentially recognized by SPP1 the question arises: is the same phage component involved in both processes, or does SPP1 use different anti-receptors for each interaction? We decided to address this question by searching directly which phage component was implicated in each receptor recognition.

IV.3.1 – IDENTIFICATION OF SPP1 TAIL STRUCTURES/PROTEINS BINDING TO YUEB

São-José *et al.,* 2006 showed that a soluble ectodomain of the receptor YueB fused to a hexa-histidine tag, named YueB780, could bind to the tail extremity distal from the SPP1 head and trigger phage DNA ejection. As it was not clear to which tail tip component YueB780 was binding to, we decided to study the receptor binding capacity of phage tail structures/proteins. For that we produced lysates of two SPP1 mutants blocked in two major steps of the SPP1 morphogenic pathway which accumulate different substructures: SPP1*sus31* that is defective for the production of the major head protein, gp13 (Becker *et al.,* 1997), and SPP1*sus45* which is defective in the major tail protein, gp17.1 (Auzat *et al.,* 2008). SPP1*sus31* infection of a non-permissive *B. subtilis* strain is expected to produce tails and tail tips, whereas SPP1*sus45* will produce heads and tail tips.

IV.3.1.1 – Binding of SPP1*sus* structural proteins to YueB780

Taking advantage of the His-tag in YueB780, we have immobilized it on a solid support (Ni-NTA Agarose beads, Qiagen) to obtain an affinity matrix for tail substructures and proteins that bind to the receptor. This affinity matrix was packed in a column and washed as described in section IV.2.6.3. The prepared lysates and a control extract produced from a non-infected culture were first incubated with free Ni-NTA beads to eliminate proteins that non-specifically bound to the support. The depleted extracts were then run through the YueB780-bound matrix and the receptor-SPP1 components complexes eluted with 500 mM imidazole and collected by gravity

flow (section IV.2.6.3). The eluted fractions were separated in a 10% SDS-PAGE, followed by silver staining and anti-SPP1 western blot analysis (Fig. IV.3).

The results for both phage lysates showed that several proteins co-eluted with YueB780 with estimated molecular weights, based on the electrophoretic mobility, of 135, 83, 32 and 22 kDa (the latter only visible for SPP1*sus31*; Fig. IV.3A). Analysis of fraction 3 of each elution by silver staining (Fig. IV.3B) revealed similar protein amounts and a few differences in the overall profile of detected polypeptides. Interestingly, a ~135 kDa protein is common to the mutant phages fractions and absent in the control. The estimated molecular weight and the relative signal intensity indicate that the 32 and the 22 kDa bands correspond to the major tail proteins gp17.1* and gp17.1, respectively (30 ± 2 and 23 ± 1 kDa, respectively; Auzat *et al.,* 2008). Although the ratio of the proteins gp17.1 and gp17.1* in the virion is 3:1, the larger species is known to produce a stronger reaction with the anti-SPP1 antibodies (Auzat *et al.,* 2008). The detection of the gp17.1* in SPP1*sus45* extracts can be explained by a small percentage of revertants and/or the infective phages present in the lysate used for infection (section IV.2.3.3).

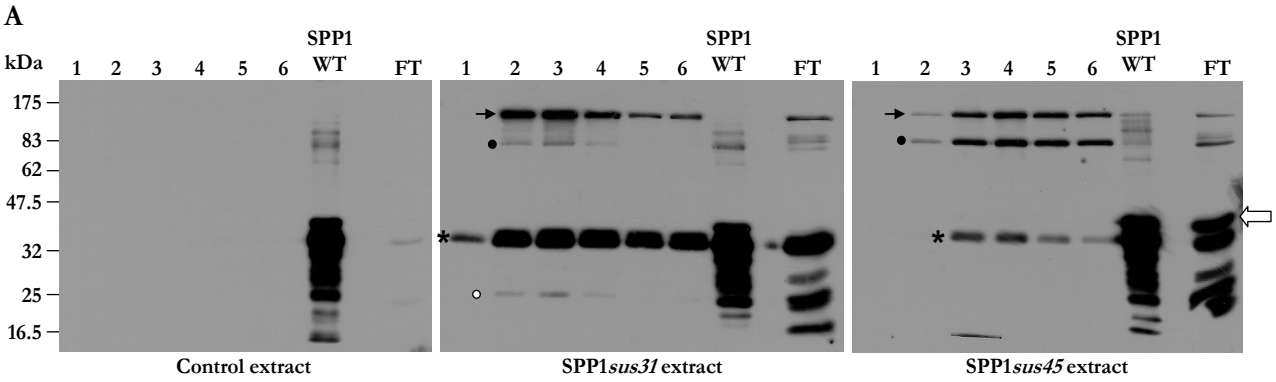


Figure IV.3 | (**A**) Anti-SPP1 western blot analysis of the SPP1 polypeptides in the different eluted fractions (1 to 6) obtained after running SPP1*sus31* and SPP1*sus45* extracts through the YueB780 affinity matrix. The black arrow and closed circle indicate the high molecular weight protein and its putative proteolytic product that co-eluted with YueB780, respectively. The asterisk indicates the major tail protein gp17.1* and the open circle the gp17.1. The block white arrow highlights the major head protein gp13 (note its absence in the SPP1*sus31* FT lane). The position of the bands of the molecular weight marker is shown on the left. FT, flowthrough. SPP1 WT, total SPP1 proteins prepared as described in section IV.2.6 (Although we used the same protein extract in all experiments we obtained some irreproducibility in the detection profile of the high molecular weight species). (**B**) Silver staining analysis of the third fraction of the eluates. Black arrow points for the high molecular weight protein as in **A**.

The estimated mass of the high molecular weight protein detected in both lysates (~135 kDa) is compatible with that of the virion tail fiber gp21 or the tape measure protein gp18, which have theoretical masses of ~124 kDa and ~110 kDa, respectively. As it will be shown below we collected evidences indicating that this protein is gp21. The detection profile observed in the SPP1*sus31* extract suggests that the ~83 kDa polypeptide corresponds to a proteolytic product of the 135 kDa protein. It is interesting to observe though, that the relative amount of these two proteins differs in the two extracts. In the case of SPP1*sus45* the smaller product is clearly more abundant than in the case of the SPP1*sus31* extract. The fact that in a SPP1*sus31* infection tail tips are expected to be assembled to tail tubes, as opposed to SPP1*sus45* that should produce isolated tail tips, may explain a stronger resistance in the first case to proteolysis in the infected cells and/or during extract preparation.

We observed in both experiments the co-elution of phage structural proteins, namely the major tail proteins, a large protein probably corresponding to gp21 and head proteins (the later seen in overexposed SPP1*sus45* blot; not shown). If we compare the relative amount of these proteins in the virion (see the SPP1 WT control lane in Fig. IV.3A), we see that high molecular weight proteins like gp21 are clearly underrepresented when compared to the major tail and head proteins. In the eluted factions we observed a clear increase of the ratio between the 135 kDa and the major tail protein gp17.1*. Moreover, given the results obtained with SPP1*sus45* we see that the presence of gp17.1* is not mandatory for the 135 kDa protein to adsorb to YueB780. Overall, the results suggested that YueB binds to a tip structure formed by the 135 kDa protein (or a complex containing it) and that the two proteins are specifically eluted from the affinity matrix.

IV.3.1.2 – Separation SPP1*sus31*-YueB780 complexes by sedimentation

In order to evidence the existence of a complex formed by the 135 kDa protein and YueB780, samples of the SPP1*sus31* eluted fractions (2 to 4) were pooled and separated by a linear 5-20% sucrose gradient (section IV.2.6.4). Seven fractions were collected and analyzed by western blot (fractions 1 to 7, from the bottom to the top of the gradient). Anti-SPP1 and anti-YueB780 immuno-detection reactions were performed sequentially (Fig. IV.4). It should be noted that the 87.3 kDa subunit that forms the YueB780 homodimer has an abnormal electrophoretic mobility behaving like a globular ~120 kDa protein (São-José *et al.,* 2006). The eluted fractions of the control extract referred to above were similarly analyzed. This control allowed us to determine

the position of YueB780 in the sucrose gradient that, as seen in Fig. IV.4, was in fractions 5, 6 and 7 (top of the gradient).

Analysis of the sucrose gradient fractions of the SPP1*sus31* eluate indicated that the tail protein gp17.1* was not associated with YueB780, since most of the protein was detected in fractions 3 and 4. The 135 kDa polypeptide was detected in factions 4, 5 and 6. In addition, the YueB780 sedimentation profile seemed to change in the presence of SPP1*sus31* proteins since fraction 5 appears to become enriched in the full length YueB780 band. These results further support a specific interaction between a structure containing the 135 kDa protein and YueB780.

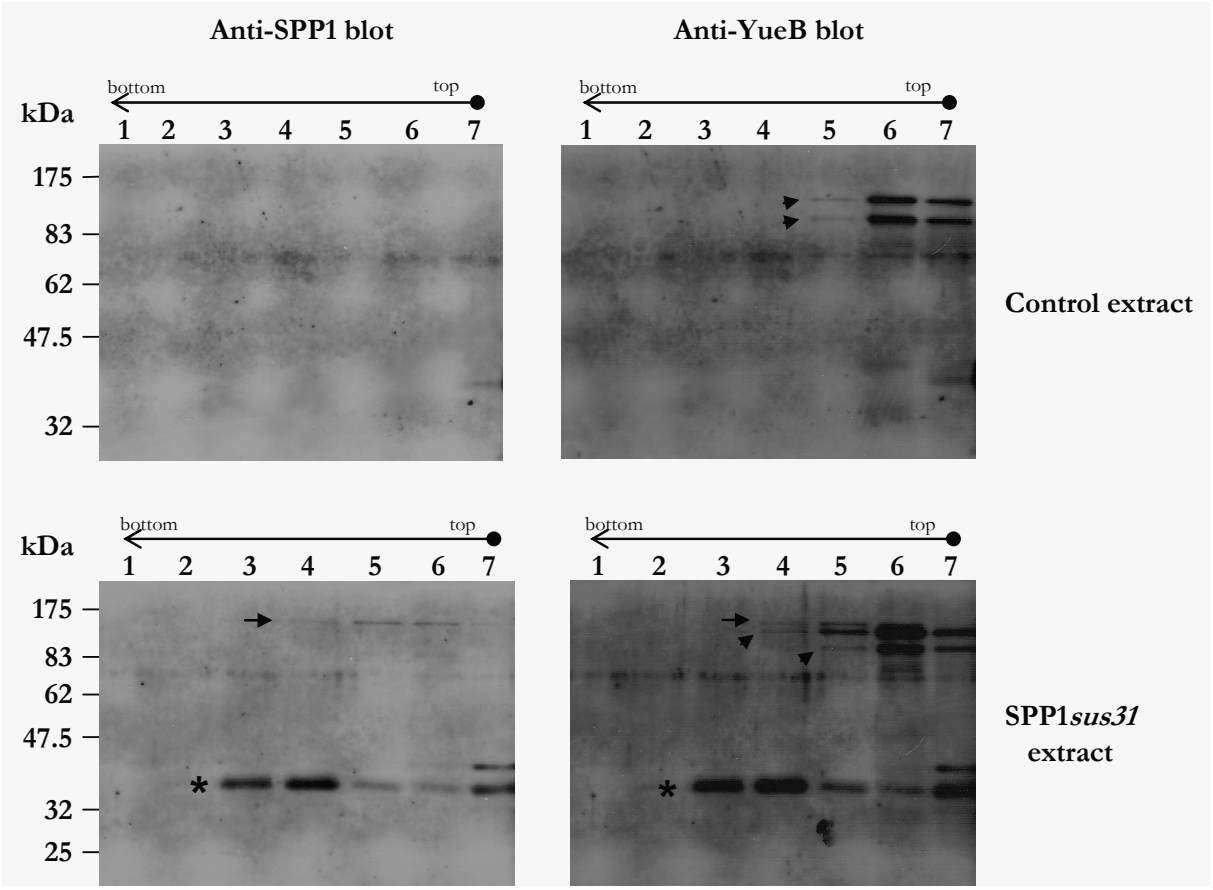


Figure IV.4 | Western blot analysis of the fractions resulting from sucrose gradient separation of the SPP1*sus31* proteins that co-eluted with YueB780. Immunodetection of YueB was performed after detection with anti-SPP1 in the same membrane. The arrows, arrowheads and asterisks indicate gp21, YueB780 and gp17.1*, respectively. The molecular weight marker is shown on the left. The sample numbers are on top of the gel.

IV.3.1.3 – Analysis of SPP1*sus31* binding to YueB780 by FPLC

Given the limited amount of proteins obtained in the experiments described above (IV.3.1.1 and IV.3.1.2) we tried to obtain a higher amount of the 135 kDa protein-YueB780 complex by running a larger amount of SPP1*sus31* lysate in a YueB-bound affinity resin coupled to an Äkta-FPLC system (Part IV.2.6.3). Elution with imidazole was done in a stepwise manner and, as seen in Fig. IV.5A, the complex starts to elute with 72.5 mM imidazole (5% of Buffer B), corresponding to fraction 9 in the Figure IV.5B (anti-SPP1 and anti-YueB western blots).

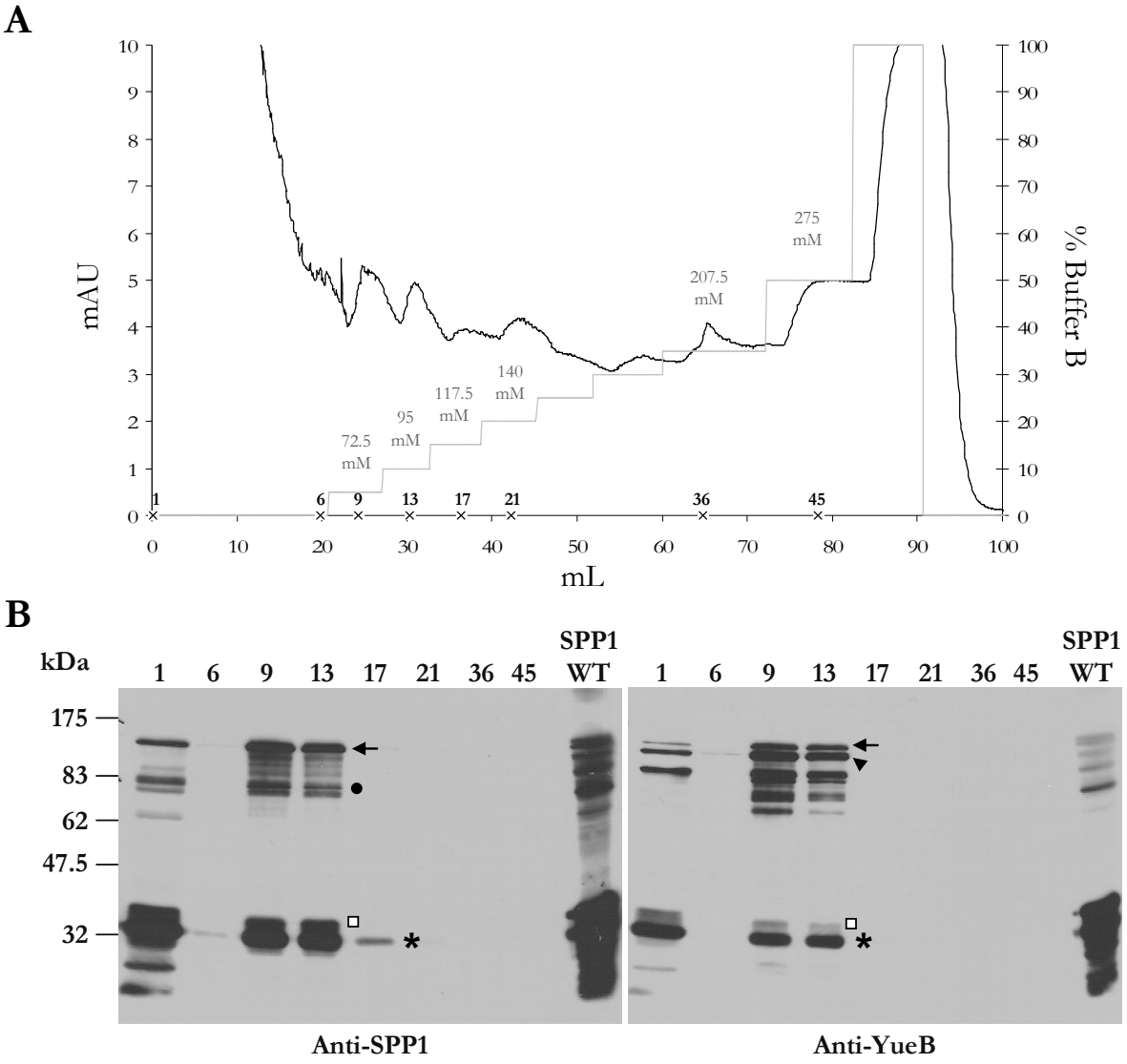


Figure IV.5 | (**A**) Affinity chromatography of YueB780 and SPP1*sus31* lysate. The black curve in the graphic corresponds to protein absorption at 280 nm. The gray curve shows the imidazole concentration increments used to elute the SPP1*sus31* tail proteins-YueB780 complexes (100% Buffer B has 500mM imidazole). The imidazole concentration used in each relevant step is indicated above the curve. The indicated fractions were 10-fold concentrated and the same amount of each was used for analysis by western blot. (**B**) Western blot analysis of the selected fractions. Immunodetection of YueB was

performed after detection with anti-SPP1 in the same membrane. The molecular weight marker is shown on the left. The sample numbers are on top of the gel. SPP1 WT, as in Figure IV.3. Arrows, closed circles and asterisks as in Fig. IV.3; arrowheads as in Fig. IV.4. The open square denotes an unidentified protein (see text).

A part of the 135 kDa protein and YueB780 eluted during the wash step (fraction 1). In the anti-SPP1 blot we could observe that fractions 9 and 13 are enriched in the high molecular weight proteins (comparing with the gp17.1* signal). The anti-YueB blot demonstrated that that SPP1 protein is indeed co-eluting with YueB780. Unfortunately with this technique we failed to separate the SPP1 135 kDa-YueB780 complexes from the other tail proteins. The gp21 and YueB780 degradation patterns observed in fractions 9 and 13 (the first two elution peaks) are compatible with the ones observed in Figs. IV.3 and IV.4. In this experiment it became clear that a ~34 kDa polypeptide (denoted with an open square in Fig. IV.5) was also being co-eluted in SPP1*sus31* extracts (also apparent in Fig. IV.3). The predicted molecular weight and position suggests that most likely this is residual capsid protein.

IV.3.1.4 – MALDI-TOF MS analysis of the YueB780 co-eluted protein

The molecular weight of the protein that co-eluted with YueB780 was compatible with that of gp21. To define unambiguously the identity of the protein, fractions 9 to 14 were pooled, concentrated (50µL) and applied to a 10% SDS-PAGE. The band visualized by silver staining was digested with trypsin and the tryptic fragments submitted to MALDI-TOF mass spectrometry (analysis performed by Jean-Pierre LeCaer; Part IV.2.6.5). Peptide mapping using ProFound was executed by Dr. Isabelle Auzat and revealed as first hit the protein gp21 of bacteriophage SPP1 (22% similarity, E value = 0.062), when compared to the viruses database. However, when the database was extended to All Taxa, the first relevant hit was the *B. subtilis* YueB protein (16% similarity, E value = 0.35) followed by SPP1 gp21 (19% similarity, E value = 0.6). This result can be explained by a contamination with YueB780 of the 135 kDa band extracted from gels due to the close electrophoretic mobility of these proteins.

IV.3.2 – INTERACTION OF GP21 FUSION PROTEINS WITH *B. SUBTILIS* YUEB780

IV.3.2.1 – Production of MBP-gp21 and GST-gp21

To further document the interaction of gp21 with YueB, we aimed to obtain a purified form of gp21 to be used in *in vitro* studies. In the genome sequence available in public databases (NC_004166) the coding sequence of gene *21* has 3336 base pairs and its deduced product has 1111 amino acid residues (expected molecular weight of 124 kDa). However, results from N-terminal sequencing of gp21 from the virion indicated that translation of gene *21* initiates at a TTG codon localized downstream of the GTG triplet assigned as start codon, and that the starting methionine is removed from the protein (Paulo Tavares and Paulette Decottignies, personal communication). The N-terminal sequence of the unprocessed gp21 is MSKNIWI and not MRSLSKNIWI. After this correction, the coding sequence of gp21 has 3327 nucleotides and its product has 1108 amino acid residues with an expected molecular weight of 123.5 kDa.

As we found that most probably the C-terminal end of gp21 is interacting with the reversible receptor (section IV.3.3) and since the C-terminal end of gp21 is predicted to form the structure facing the distal rod of the tail tip (Plisson *et al.,* 2007), we decided to fuse affinity tags to the amino terminus of gp21. Vectors pMAL-c2 and pGEX-4T-1 were used to fuse MBP (Maltose Binding Protein; NEB) and GST (Glutathione S-Transferase; Amersham Biosciences), respectively, to the N-terminus of gp21. In addition to provide an affinity tag for purification, these domains have been described to confer increased solubility to the proteins to which they are fused (Kim and Lee, 2008; Waugh, 2005). Fusion proteins MBP-gp21 and GST-gp21 are predicted to have ~166 kDa and ~152 kDa, respectively. Increased amounts of soluble gp21 fusion proteins were achieved when co-expressed with chaperone proteins encoded by plasmid pG-KJE3 (GroEL, GroES, DnaK, DnaJ and GrpE; Nishihara *et al.,* 1998, Table IV.1; Fig. IV.6). As shown below, in addition to the expected full-length fusion proteins western-blot analysis also revealed a significant amount of truncated versions.

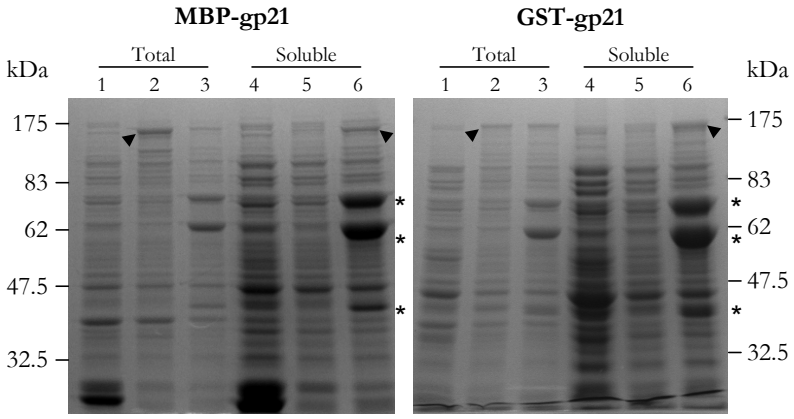


Figure IV.6 |SDS-PAGE/Coomassie blue staining of the total (1, 2, 3) and soluble (4, 5, 6) fractions of MBP-gp21 and GST-gp21 protein extracts produced from induced cultures of BL21 derivatives. Control extracts with the vectors used in gene *21* cloning are in lanes 1 and 4; extracts of ECB-MBP21 and ECB-GST21 strains are in lanes 2 and 5 of the corresponding gel; extracts of strains co-expressing the chaperones, ECB-MBP21-Chap and ECB-GST21-Chap, are in lane 3 and 6 of the corresponding gel. Arrowheads point to the expected gp21 fusions. Asterisks indicate the visible overexpressed chaperone proteins: DnaK (~70 kDa), GroEL (~60 kDa) and DnaJ (~40 kDa). Molecular weight marker is shown on the sides of the blots.

IV.3.2.2 – Study of GST-gp21 and MBP-gp21 binding to YueB780

The gp21 fusions were assayed for their ability to bind to YueB780. Essentially, the experiments consisted in running through a YueB780-bound Ni column purified fractions of the gp21 fusions and analyze their binding to YueB780 (see details in section IV.2.6.8).

GST-gp21 was purified with a glutathione affinity column and the elution resulted in a single peak (not shown) enriched in partially purified GST-gp21 (lane GST-gp21 in Fig. IV.7B and C). After an efficient binding of YueB780 to the Ni matrix (not shown), the GST-gp21 enriched fraction was loaded. Fractions 15 and 16 (Fig. IV.7A) correspond to the flowthrough peak of GST-gp21 (pooled and analyzed in Fig. IV.7B and C). Fractions 24 and 25 match the elution peak obtained with 30% of YueB Buffer B (185 mM imidazole) and were also pooled and analyzed. After this we injected 100% of YueB Buffer B and collected fraction 33 that corresponds to a weak, broad peak, also analyzed. From the western analysis we can conclude that most GST-gp21 did not bind to the affinity matrix (Fr. 15+16, Fig. IV.7B and C). A small fraction though could bind to the matrix and was subsequently co-eluted with YueB780 (Fr. 24+25). The fraction just before the co-elution peak (Fr. 21 in Fig. IV.7A) was also analyzed (not shown) and no traces of GST-gp21 or YueB780 were detected. This indicates that the two proteins co-eluted when YueB780 was desorbed from the Ni matrix with imidazole.

Purified MBP-gp21 was also tested for YueB780 interaction and the results were very similar to those obtained with GST-gp21. After MBP-gp21 purification on an amylose resin, the partially purified MBP-gp21 fraction was desalted before loading the YueB780-bound Ni resin (MBP-gp21 in Fig. IV.8B and C). The Coomassie blue staining in Figure IV.8B shows that the amount of partially purified MBP-gp21 loaded on the YueB780-bound Ni column was scarce. Even though, as can be observed in the western blots presented in Figure IV.8C, MBP-gp21 is present in the co-elution fraction (Fr. 13) although the vast majority of the fusion protein is lost in the flowthrough (Fr. 4). After the elution with 185 mM imidazole (30% Buffer B; Fr. 13) the column was eluted with 500mM of imidazole (100% Buffer B) and the resulting peak was analyzed (Fr. 28 and 29 pooled).

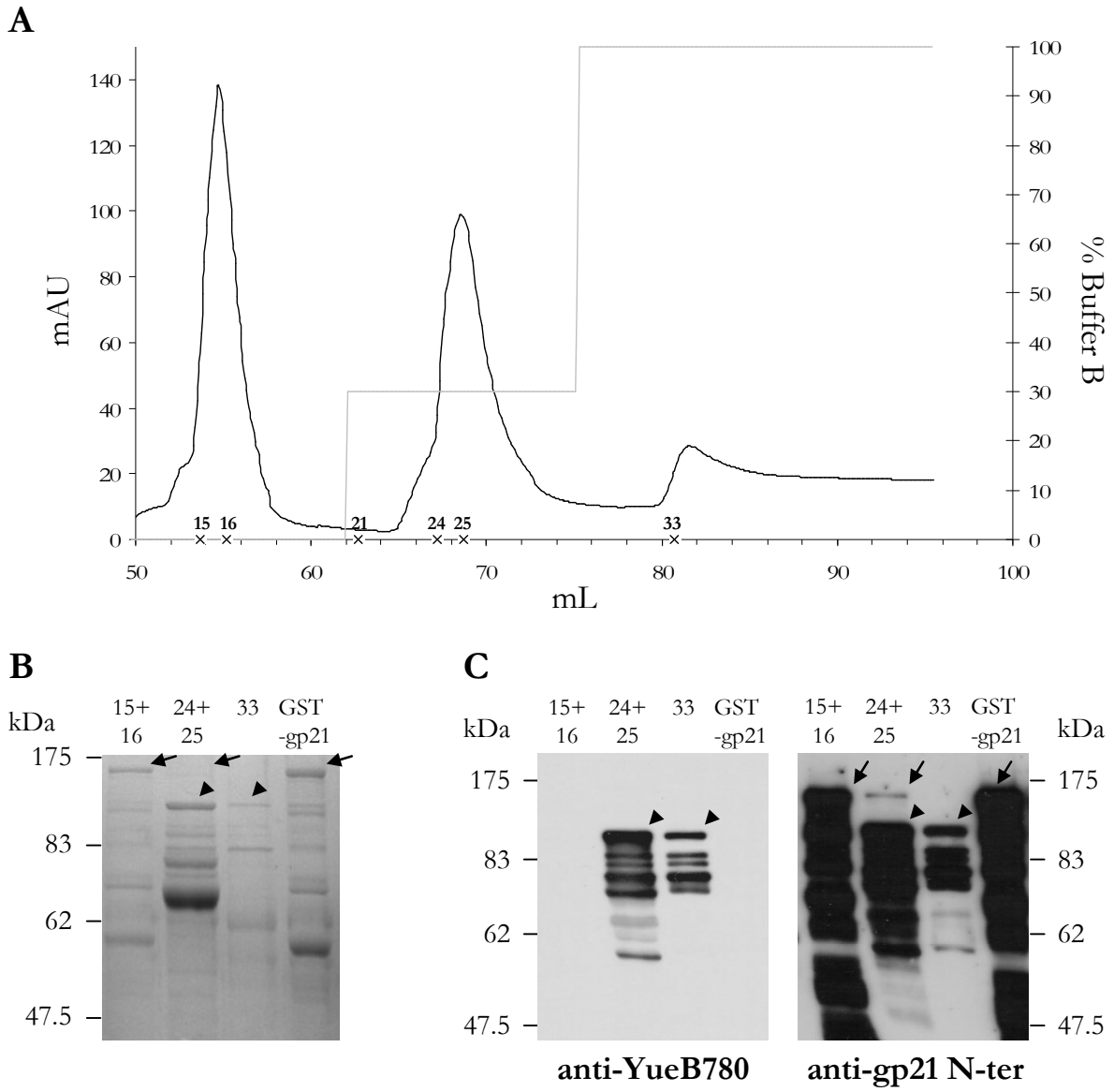


Figure IV.7 | (**A**) GST-gp21 binding to a YueB affinity matrix. The black curve in the graphic corresponds to protein absorption (280 nm). The gray curve shows the imidazole concentration increments used for elution of GST-gp21-YueB780 complexes. The fractions analyzed are indicated. (**B**)

Coomassie blue staining of the selected fractions. Lanes are designated according to fractions in **A**: 15+16, GST-gp21 flowthrougth of the nickel column; 24+25, GST-gp21 and YueB780 co-elution peak with 185 mM of imidazole; 33, elution with 500mM of imidazole; GST-gp21, elution peak of the GST-gp21 purification in a glutathione affinity column. Arrows and arrowheads indicate the GST-gp21 and YueB780 full length polypeptides, respectively. (**C**) Western blot analysis of the selected fractions. Lanes and symbols as in **B**. Immunodetection of GST-gp21 was performed after detection with anti-YueB780 in the same membrane. The position of the molecular weight marker is shown on the side of each blot. Anti-gp21 N-terminal antibodies were kindly provided by Dr. Isabelle Petitpas.

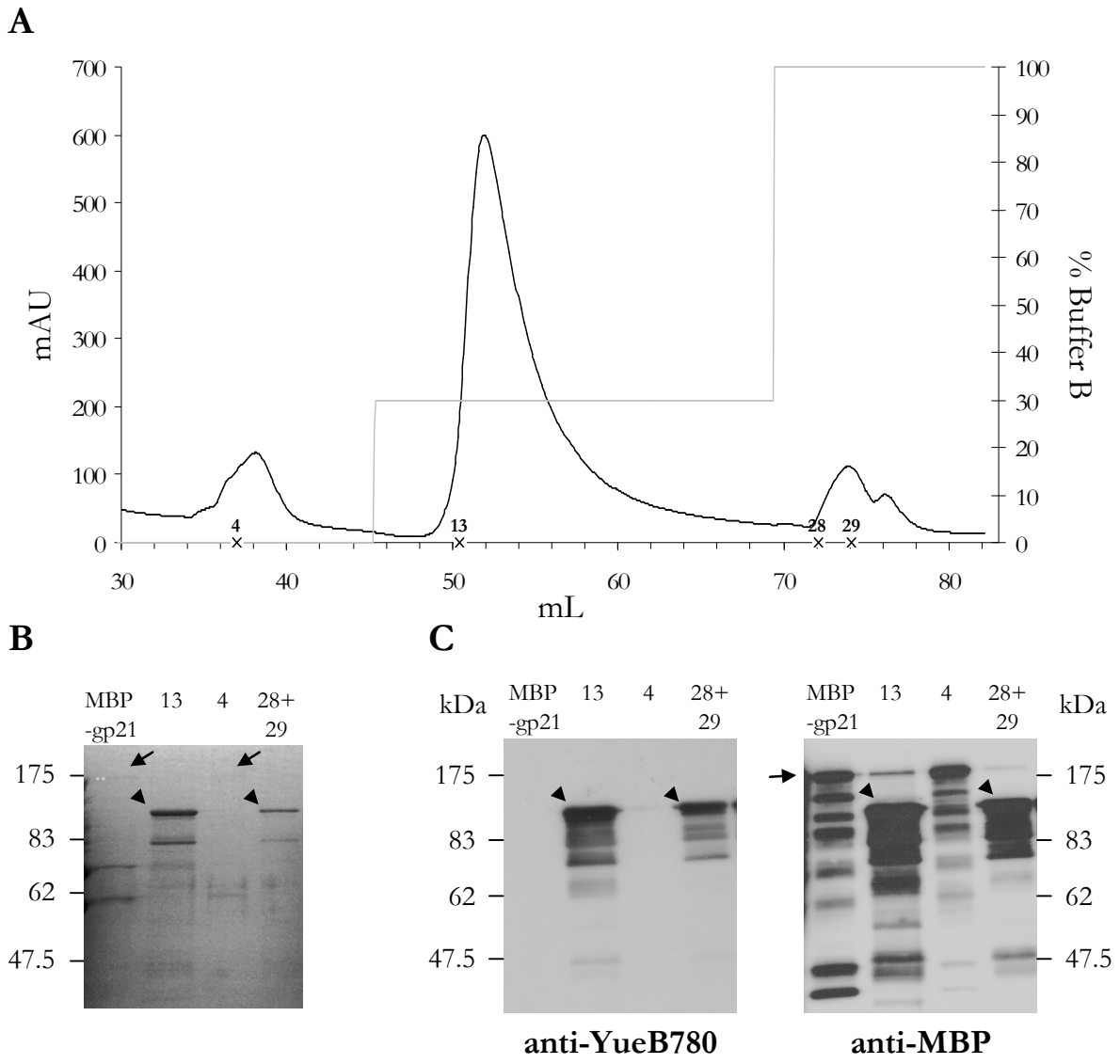


Figure IV.8 | (**A**) MBP-gp21 binding to a YueB affinity matrix. Black and gray curves as in Fig. IV.7A. The fractions analyzed by western blot are indicated. (**B**) Coomassie blue staining of the selected fractions. Lanes are designated according to fractions in **A**: MBP-gp21, elution peak of MBP-gp21 purification (desalted); 13, co-elution peak of YueB780 and MBP-gp21 with 185 mM of imidazole; 4, MBP-gp21 flowthrougth of the nickel column; 28+29, elution peak with 100% of YueB Buffer B. Arrows and arrowheads indicate the MBP-gp21 and YueB780 full length polypeptides, respectively. (**C**) Western blot analysis of the selected fractions. Lanes and symbols as in **B**. Immunodetection of MBP fused proteins

was performed after detection with anti-YueB780 in the same membrane. The position of the molecular weight marker is shown.

The results presented above suggested that a small fraction of the produced gp21 fusion proteins might bind to the receptor YueB780 immobilized in a nickel column. However, it is difficult to conclude about the specificity of this interaction since we have not controlled a possible unspecific binding of the tagged gp21 to the Ni column. Moreover, we have not dissociated the chaperones from purified gp21 fusions before running through the YueB780-bound column (see Discussion, IV.4).

IV.3.3 – THE C-TERMINAL REGION OF GP21 IS INVOLVED IN SPP1 REVERSIBLE ADSORPTION

IV.3.3.1 – Phages with abnormal reversible adsorption have mutations in gp21 gene sequence

To identify SPP1 proteins that target glucosylated WTA we have developed a strategy for the isolation of phage mutants specifically impaired in reversible adsorption. For that we took advantage on our previous study of SPP1 adsorption (see Part III), which showed that SPP1 can infect *B. subtilis* in solid medium in the absence of the reversible adsorption step. Essentially, we performed several cycles of selection of SPP1 mutants unable to adsorb to strain CSJ1 (*yueB*$^-$; wild type WTA) in liquid media, with the enrichment of the selected mutants in solid medium between each cycle (see section IV.2.4 for protocol details).

Between ten selected phages, each propagated from a single plaque after the selection (enrichment protocol), that showed an altered adsorption phenotype (SPP1mut1, 3, 5, 10, 13, 15, 20, 30, 35, and 40) we will only present the adsorption results of SPP1mut3 and SPP1mut40, as they are representative of the other phages (not shown).

The study of phages SPP1mut3 and SPP1mut40 adsorption to strain CSJ1 at 0ºC (a condition where we have a very low rate of desorption) showed that these phages are severely affected in their ability to reversibly adsorb to *B. subtilis* cells when compared to the wild type SPP1 (Fig. IV.9A). As would be expected for phages impaired in the reversible adsorption step, the rate of

irreversible adsorption to YueB measured in the mutant phages is reduced relatively to the wild type phage (4-fold reduction; Fig. IV.9B). It is interesting to note that, as described for the *gta* mutants in Part III, as long as the membrane receptor YueB is present, these phages can still adsorb irreversibly at very low rates and infect *B. subtilis*.

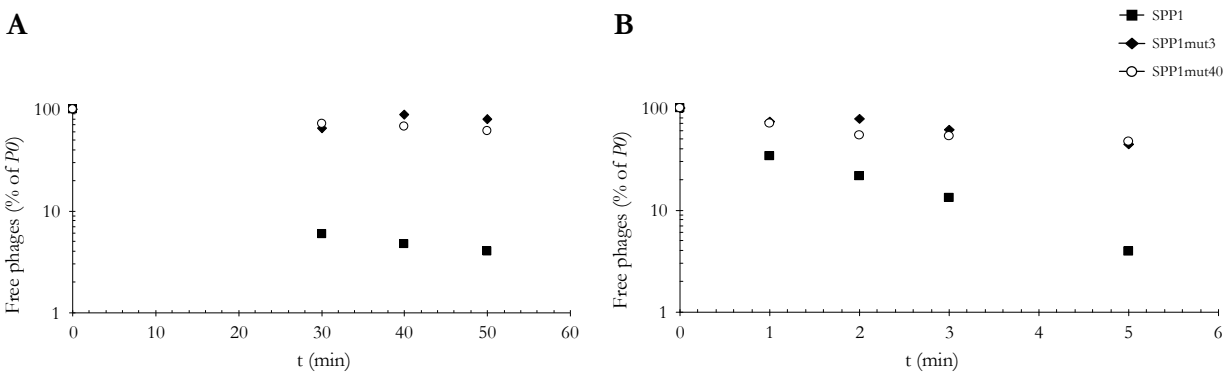


Figure IV.9 | Adsorption profiles of SPP1 wild type (filled squared) and mutant phages SPP1mut3 (filled diamonds) and SPP1mut40 (open circles). Free phages are represented as a normalized concentration (*P*x100)/*P0*, where *P* is the fraction of free phages (plaque forming units per ml (pfu/ml)) at time *t* and *P0* the initial input. (**A**) Reversible adsorption to *B. subtilis* strain CSJ1 set to A_{600} = 0.4 and equilibrated at 0ºC before phage addition. Reversible adsorption was measured at the equilibrium state. (**B**) Irreversible adsorption to wild type strain L16601 set to A_{600} = 0.8, at 37ºC.

For CsCl purification the mutant phages were multiplied in solid medium using 200mL Petri dishes and recovering the phages resulting from confluent lysis plaques with 70mL of TBT Buffer. After purification we decided to calculate the efficiency of platting of each mutant regarding the wild type phage. To guarantee that we used the same number of virus particles in the titration assays, we quantified the total protein content of the tested lysates. For the same amount of virions, based on protein quantity, we obtained approximately the same number of phage plaques, meaning that there was no significant variation of the E.o.P between the wild type and the two SPP1 mutant phages.

Since the results presented in the previous sections have indicated that gp21 is involved in SPP1 binding to YueB780, we next decided to sequence the gp21 gene of the 10 mutant phages to look for mutations that could explain the adsorption phenotype. Among the 10 phages, two distinct point mutations in the 3' end of gene *21* were identified (Fig. IV.10). In phages SPP1mut35 and SPP1mut40, a single nucleotide substitution of an adenine by a guanine (nucleotide 3045 in gene *21* coding sequence) leads to a glutamine substitution by an arginine in the gp21 primary

sequence. In the other eight selected mutants a transition of a thymine to a cytosine (nt 3180), leads to a change from a phenylalanine to a serine.

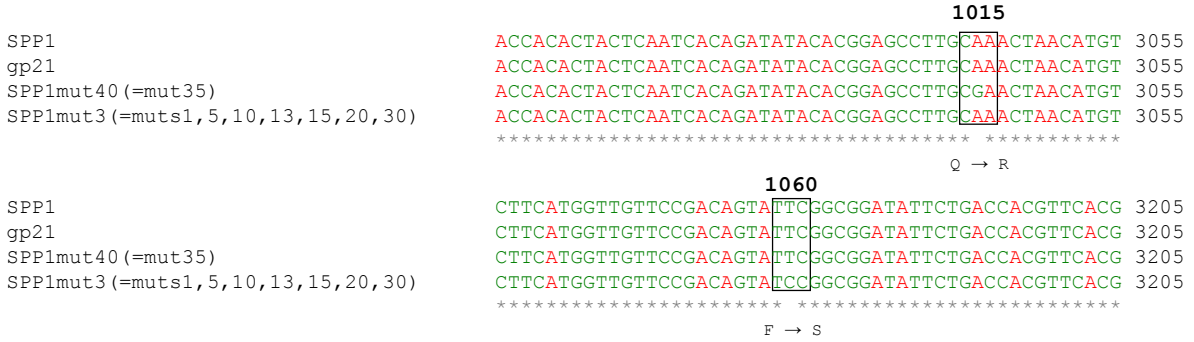


Figure IV.10 | CLUSTAL W2 multiple sequence alignment highlighting the sequenced point mutations found in gene *21* sequence of the mutant phages. SPP1 refers to the sequenced wild type gene *21*. gp21 refers to the deposited sequence in the database. Under the alignment is represented the resulting amino acid substitution. Q, glutamine; R, arginine; F, phenylalanine; S, serine. The boxes indicate the in frame triplet. Residue numbers regarding the 1108 aminoacid gp21 protein are presented above each modified triplet (CLUSTAL W version 2.0.10).

IV.3.3.2 – Recombination with a wild type gene *21* sequence restores reversible adsorption to mutant phages

In order to sustain the evidence that the mutations found in gene *21* caused the altered adsorption phenotype, we aimed to correct the sequence of the gene in the mutants to see if we could restore normal reversible adsorption. Mutant phages SPP1mut3 and SPP1mut40 were titrated in the *B. subtilis* strain SP1154, which carries a wild type copy of gene *21* in a replicative plasmid (pIA2, see Table IV.1). The rationale was that a recombination event between mutant and wild type copies would restore the correct gene *21* in the genome of the mutant phages. We collected 100 isolated phage plaques from each titration and the eluted phages were screened for a wild type adsorption phenotype to strain CSJ1 at 0ºC (for details see section IV.2.5). In parallel the wild type *B. subtilis* strain was also used to titrate both mutants and the same number of isolated plaques collected to account for the spontaneous revertants. Eluted phage plaques of wild type SPP1 were also used as controls in adsorption assays. Essentially, the phages eluted from each plaque were scored for the parameter P/P0, where P is the fraction of free phages in the equilibrium state and P0 the initial input. Typically, in these conditions the value P/P0 was ≤0.02 for the wild type SPP1 and ≥0.3 for the mutant phages. Table IV.3 shows the results for each set of 100 eluted phage plaques concerning the P/P0 parameter.

Table IV.3 | Values of P/P0 measured for each set of 100 isolated phage plaques of each titration event. Highlighted in red are the values that were approximate to the wild type value. Total row refers to the sum of tested phages with wild type parameter.

	SPP1mut3 in SP1154	SPP1mut40 in SP1154	SPP1mut3 in L16601	SPP1mut40 in L16601		SPP1mut3 in SP1154	SPP1mut40 in SP1154	SPP1mut3 in L16601	SPP1mut40 in L16601
Isolate #	**P/P0**	**P/P0**	**P/P0**	**P/P0**	**Isolate #**	**P/P0**	**P/P0**	**P/P0**	**P/P0**
1	0.52	>1	0.81	>1	**51**	>1	>1	>1	0.79
2	>1	>1	>1	>1	**52**	>1	>1	>1	>1
3	>1	0.13	>1	>1	**53**	>1	>1	>1	>1
4	0.01	>1	>1	>1	**54**	>1	>1	>1	>1
5	>1	>1	0.81	>1	**55**	>1	>1	0.6	>1
6	0.07	>1	>1	0.77	**56**	>1	>1	>1	>1
7	>1	>1	>1	>1	**57**	0.03	0.04	0.38	0.73
8	>1	>1	>1	0.73	**58**	>1	>1	>1	>1
9	0.81	>1	>1	>1	**59**	>1	0.05	ND	0.68
10	>1	>1	>1	>1	**60**	0.07	>1	ND	0.68
11	>1	>1	>1	0.42	**61**	>1	0.05	0.64	>1
12	>1	>1	0.63	>1	**62**	0.07	>1	>1	0.86
13	>1	>1	0.45	>1	**63**	>1	>1	>1	>1
14	>1	>1	0.59	>1	**64**	>1	>1	>1	0.87
15	>1	>1	0.05	>1	**65**	>1	>1	>1	>1
16	>1	0.07	0.39	>1	**66**	>1	0.05	0.53	>1
17	0.05	0.08	>1	>1	**67**	>1	0.04	>1	0.58
18	0.12	>1	ND	>1	**68**	>1	>1	0.17	>1
19	0.12	>1	>1	>1	**69**	>1	>1	0.45	>1
20	>1	>1	>1	>1	**70**	0.04	>1	>1	>1
21	>1	>1	>1	>1	**71**	0.12	>1	0.32	0.78
22	>1	0.04	0.9	>1	**72**	0.04	>1	0.38	0.6
23	0.04	0.01	>1	>1	**73**	0.05	>1	0.41	>1
24	>1	>1	0.59	>1	**74**	>1	0.13	>1	ND
25	0.97	>1	0.49	>1	**75**	0.1	>1	>1	0.52
26	>1	>1	>1	>1	**76**	>1	>1	>1	>1
27	>1	>1	>1	0.92	**77**	>1	>1	>1	0.82
28	>1	>1	0.63	>1	**78**	>1	>1	0.4	>1
29	>1	0.05	>1	0.92	**79**	>1	>1	>1	0.69
30	>1	>1	>1	>1	**80**	>1	>1	>1	>1
31	>1	0.37	0.4	>1	**81**	>1	>1	>1	0.64
32	>1	0.21	0.44	>1	**82**	0.03	>1	0.64	>1
33	>1	>1	0.54	>1	**83**	>1	>1	>1	>1
34	>1	>1	0.59	0.97	**84**	>1	0.6	>1	0.94
35	>1	>1	>1	>1	**85**	>1	>1	>1	0.6
36	0.1	>1	0.72	>1	**86**	>1	>1	>1	>1
37	>1	0.09	>1	0.81	**87**	>1	>1	>1	>1
38	0.16	>1	>1	>1	**88**	0.95	>1	>1	>1
39	>1	>1	>1	>1	**89**	>1	>1	>1	>1
40	>1	>1	>1	>1	**90**	>1	0.2	>1	>1
41	>1	>1	>1	>1	**91**	>1	>1	0.63	>1
42	>1	>1	>1	>1	**92**	>1	0.03	>1	>1
43	0.02	0.79	0.74	>1	**93**	0.05	0.45	>1	>1
44	>1	>1	0.3	>1	**94**	>1	>1	>1	>1
45	0.06	0.04	>1	0.53	**95**	0.07	0.3	0.47	>1
46	0.11	>1	>1	0.5	**96**	>1	0.05	>1	>1
47	0.04	>1	>1	0.45	**97**	0.82	>1	>1	>1
48	>1	0.21	>1	0.8	**98**	0.2	>1	0.97	>1
49	0.06	>1	>1	0.93	**99**	>1	>1	>1	>1
50	>1	0.06	>1	0.99	**100**	>1	>1	>1	>1
					Total	**17**	**15**	**1**	**0**

ND - The phage input was too low to measure P/P0

When SPP1mut3 and SPP1mut40 were titrated in strain SP1154, 17 and 15 potential recombinants, respectively, were selected. These phages had a P/P0 value <0.1 (highlighted in

red). In the control titration of both mutants in the wild type strain all phage isolates gave a P/P0 value ≥0.3 (except for phage #15 of the SPP1mut3 titration), which is the result expected for mutants in reversible adsorption. This result indicated that spontaneous reversion did not significantly contributed for the isolation of phages with increased reversible adsorption.

The potential recombinants mentioned above were propagated and checked again for their reversible adsorption to CSJ1. Three potential recombinants that derived from each mutant phage and that showed a reversible adsorption profile similar to the wild type (Fig. IV.11) were selected for sequencing gene *21*. The results confirmed that the single point mutations of SPP1mut3 and mut40 had been corrected for the wild type sequence, either through recombination or gene conversion event. Overall, the evidences presented here indicate that the C-terminal region of gp21 is involved in Glc-WTA recognition during SPP1 reversible adsorption.

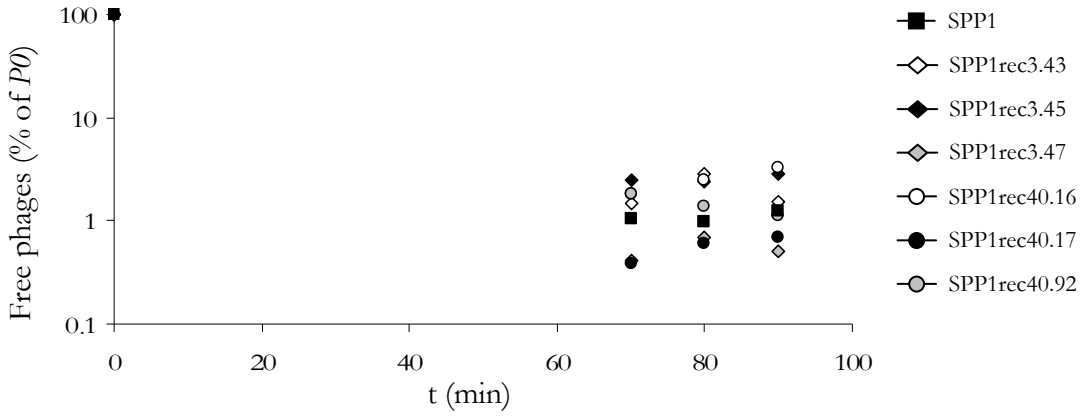


Figure IV.11 | Adsorption profiles of SPP1 wild type (filled squared) and recombinant phages from SPP1mut3 (diamonds) and SPP1mut40 (circles). Free phages are represented in Figure IV.9. Reversible adsorption at the equilibrium state to *B. subtilis* strain CSJ1 set to A_{600} = 0.8 and equilibrated at 0ºC before phage addition.

IV.4. DISCUSSION

After having identified both receptors that phage SPP1 uses to efficiently adsorb to *B. subtilis* cellular surface, we initiated a study to identify the SPP1 RBP(s). We were particularly interested to disclose if SPP1 uses the same or distinct RBPs for each type of adsorption. We have tried to answer that question by two main approaches that will be discussed below.

Regarding the virion components concerned in irreversible adsorption, we tried to capture them based on their specific binding to a YueB780-affinity matrix. Total protein extracts produced from *B. subtilis* cells infected with SPP1*sus* mutants, which were defective for the production of phage capsids or tails, were used as the source of phage structures. When these extracts were run through the affinity matrix a ~135 kDa protein was captured. The specificity of the interaction between this protein and YueB780 was further documented by sucrose gradient analysis. Mass spectrometry identified this protein as being the tail tip forming protein, gp21. These results indicate that gp21 or a gp21-containing substructure is responsible for binding to YueB.

Unfortunately, when we studied the binding capacity of partially purified gp21 fusions overproduced in *E. coli* to the same affinity matrix the results obtained were not fully convincing about the RBP nature of gp21. Though, some possible explanations for the poor binding to YueB can be put forward. It was hypothesized that gp21 forms a trimer in the SPP1 tail fiber (Plisson *et al.*, 2007) and it is possible that this oligomeric state is in fact the active form of gp21. It may happen that the conditions used in the assays might not favor oligomerization or the correct folding of gp21. In addition, formation of functionally active gp21 may depend on other phage proteins that should be present in SPP1*sus* infections but which are absent during gp21 production in *E. coli*. Therefore, the excess ratio of gp21 fusion proteins to YueB780 that was required to detect their interaction can be explained if the fraction of active gp21 loaded in the YueB780 affinity column was too low. Moreover, it has been shown that the purified YueB780 is not 100% active since we need about 130 YueB780 dimers to inactivate one SPP1, a ratio that should be far from the biological situation. Also, as we did not dissociate the chaperone proteins from gp21 fusions before loading on the YueB780-bound column, we can not differentiate between a specific although weak interaction of gp21 to YueB780 and an unspecific binding of the chaperones to YueB780.

A different approach was used for identification of the protein involved in the reversible binding. In this case we designed a protocol, based on our previously developed studies (Part III), for isolation of mutant phages that showed the incapacity to reversibly adsorb to a Δ*yueB* strain. In the conditions used in this protocol we forced minimum phage release from the reversible receptors and collected those that were unable to adsorb. The mutant phages could however be propagated in the wild type strain, in solid medium, in accordance to our results showing the dispensability of reversible adsorption in this condition (section III.3.1). Rescue of adsorption capacity in the selected phages was accomplished after recombination (or gene conversion event) with a wild type sequence of gene *21*, present in a plasmid in the infected cells. Sequencing of gene *21* in the mutant phages revealed two distinct point mutations in the 3' end of the gene (Fig. IV.12) and the restore of the wild type sequence in the rescued phages. As this result might uncover a second receptor binding function for the tail fiber protein we have performed a preliminary study where we checked if the tagged gp21 proteins could bind to *B. subtilis* cells. For this we selected the wild type strain L16601 and the mutants CSJ1 and CSJ3.*gtaC*. As shown in Part III the latter two strains only display reversible (glucosylated WTA) and irreversible receptors (YueB), respectively, with CSJ3.*gtaC* overproducing YueB in the presence of IPTG. After an incubation period at 0ºC or 37ºC samples were centrifuged and MBP-g21, GST-gp21 content both in the pellet and supernatant was analyzed by SDS-PAGE. Although it needs confirmation, the results indicated that cells CSJ3.*gtaC* bound a higher fraction of gp21 when compared to the other strains (not shown). This again suggests a specific interaction between the gp21 fusion proteins and the cellular YueB but adds nothing to its ability to target the cell wall receptors. The apparent correlation between the observed mutations in gene *21* and the adsorption phenotype remains to be confirmed, although it is remarking that correction of those mutations results in phenotype rescue. No other genes coding for tail proteins were tested or sequenced in this analysis.

Nevertheless, some analysis can be presented that support the RBP function of gp21. Blast and PSI-Blast searches with gp21 primary sequence showed that its N-terminal region is homologous to several putative host specificity proteins and tail fibre proteins (see IV.1). As expected for a protein that is involved in specific receptor recognition, no significant homology was observed for the C-terminal half of gp21 other than a small region of about 150 aa that shared similarity with two putative tail fiber proteins (Orfs 53 and 97 of *Bacillus clarkii* phage BCJA1c and *Lactobacillus plantarum* phage LP65, respectively; Fig. IV.12). Sciara *et al.* (2008) assumed a tail associated lysozyme function (Tal) for gp21 given its homology with Tal of phage Tuc2009 (Fig.

IV.12). However, the homology between these two proteins is restricted to both N-terminal regions, whereas the lysozyme activity is specified by the C-terminal portion of $Tal_{Tuc2009}$. Nevertheless, the authors went forward with evolutionary considerations, assessing that Tal proteins might fulfil very different functions like adhesion, DNA injection initiation and cell lysis. On top of that, bioinformatics analyses of the secondary structure of gp21 revealed a right-handed β-helix fold (residues 678 to 844), a feature commonly present in RBPs (Part I). The coordinates of this fold do not correspond to those described in Plisson *et al.* (2007) because in this report the authors used for this analysis an uncorrected gp21 sequence (Acc. N0 O48463), which lacks 262 N-terminal residues. In the same report the authors suggested a structural similarity between gp21 and the main domain (1TSP; Steinbacher *et al.,* 1994) of the tail spike from phage P22. P22-like TSPs possess elongated grooves in which they specifically recognize polysaccharide shapes (Barbirz *et al.*, 2009 and references therein).

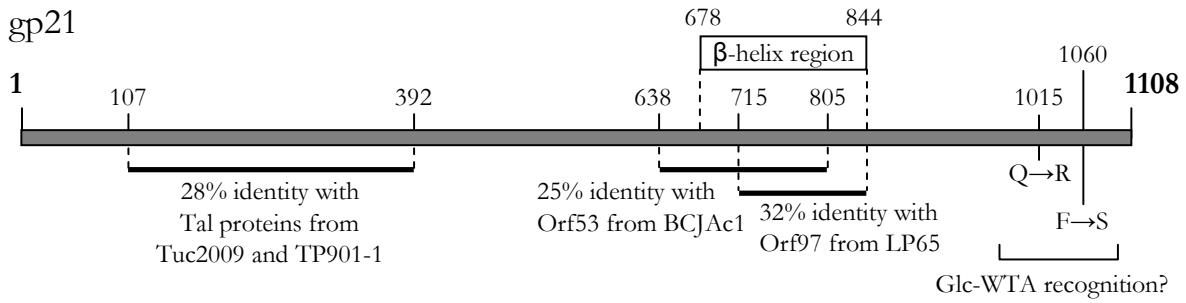


Figure IV.12 | Schematic representation of gp21 (in gray) displaying its major features and their localization regarding the primary sequence Residue numbers are on the top part of the scheme. Homologous sequences were retrieved using PSI-Blast search; the percentage of identity is indicated. Prediction of the right-handed β-helix region was performed with BetaWrap (Bradley *et al.*, 2001). The amino acid substitutions resulting from the mutations found in SPP1 mutants impaired in reversible adsorption are also indicated. These are probably included in a region that is responsible for glucosylated-WTA recognition.

Puzzling data evolves from two studies on the interaction of SPP1 with purified YueB. Plisson *et al.* (2007) observed that after ejection, the cap structure opens, the tail helical structure rearranges and the inner tube that fills the inside of the tail is expelled. This analysis also showed that the tail rearrangements involved the loss of the tip and the authors advanced that after this lost it is probably the cap structure that adheres to the host surface. The electron microscope analysis performed by São-José *et al.* (2006) showed YueB780 bound to the tail extremity of full SPP1 virus particles but also to that of empty phages. In none of the situations could the tail fiber be easily discerned. We can advance two interpretations for these observations: i) YueB is bound by

(an)other phage structural component; ii) the association to YueB modifies the fiber structure making it "invisible" while loosely bound to the tail.

Also to be taken in consideration regarding this matter is the programmed translational frameshift in the major tail protein gene *17.1*. This frameshift adds an immunoglobulin-like domain to gp17.1 generating gp17.1*, being both products present in the SPP1 tail in a 3:1 ratio (Auzat *et al.,* 2008). The gp17.1* appendages protrude from the helical tail given it a "hairy" look. These Ig-like extensions were predicted for several phages and it was speculated that they might serve to increase infectivity by binding to or degrading polysaccharides (Fraser *et al.,* 2007). Auzat *et al.* (2008) tested this possibility and observed that, independently of the SPP1 tail being composed of just one of the MTP forms, the phage adsorption rate was as the wild type. Even so, they proposed that this domain could mediate the initial contact and attachment of SPP1 to the bacterial surface but that this interaction would be difficult to detect. Taking in account our work on the kinetic of adsorption presented in Part III we would expect that, if the phage no longer possesses the structures that enable it to reversibly adsorb to the Glc-WTA, that would mimic the absence of the reversible receptor, and the effect on adsorption would be noticeable (SPP1 adsorbed at very low rates to strain L16601.*gtaC* and exhibited a lower E.o.P and a small plaque phenotype in this strain; section III.3.3). So, if the Ig-like domains present in gp17.1* are involved in some type of interaction, we believe that this should not correspond to the reversible adsorption described here. Anyway, it may happen that these domains stabilize phage adhesion to the cell surface after adsorption and/or during DNA ejection.

The two step adsorption is a widely used strategy among the tailed phage world. Phages conduct this by using their tail components and appendages. Usually phages that do possess the later structures use them for the first step of adsorption in order to keep associated to the cell surface and facilitate subsequent interactions. As enumerated before in this thesis, the L-shaped fibers of phage T5 and the long tail fibers of T4 are examples of this. Phages that do not possess these complementary structures pose a difficulty (to the scientist, not to the phage): how to scan the bacterial surface until the irreversible receptor is found or the phage can be properly positioned to eject the DNA? In phage SPP1, where the reversible step is known to exist, we presented evidences that the same tail component, gp21, may be involved in both steps of adsorption. This can also be the case of other Siphoviruses that, like SPP1, do not have additional structures other than the tail tip RBPs. These include the routinely used phage λ (a mutant devoid of long fibers) and the lactococcal phage c2.

PART V

CONCLUDING REMARKS

The major goals that we projected were accomplished in this work. We identified the *B. subtilis* receptor involved in SPP1 reversible adsorption and for the first time we were able to present a model for phage adsorption that integrates both molecular and kinetics data. It describes how the reversible and irreversible binding steps, each to a distinct cellular receptor, contribute to the whole adsorption process and, consequently, to the efficiency of infection. This study has provided us tools to genetically dissect virion and cellular components involved in SPP1 adsorption. These were used in this work to isolate SPP1 mutants impaired in reversible adsorption. The same approach can be used to isolate mutants affected in YueB recognition. We presented results suggesting that SPP1 uses a single RBP, gp21, to target the two cellular receptors. We expect continuing our partnership with the group of Dr. Paulo Tavares to confirm the RBP nature of gp21 along with its precise contribution to the tail tip structure. This collaborative work also aims to define precisely the YueB motifs involved in SPP1 binding and to deepen the study of its localization in the cell envelope. Concerning the yukE operon it will be interesting to study its predicted function as a type VII secretion system. It was suggested that originally this system had a general cellular function that in some bacteria has evolved to serve virulence or other specialized role. The yukE operon of *B. subtilis* may represent an archetypal type VII secretion system and a good model to answer many questions about its working mechanism.

PART VI

REFERENCES

Abdallah, A M., N. C. Gey van Pittius, P. A. Champion, J. Cox, J. Luirink, C. M. Vandenbroucke-Grauls, B. J. Appelmelk and W. Bitter. (2007) Type VII secretion – mycobacteria show the way. Nat. Rev. Microbiol. **5**: 883-91

Abedon, S. T., and R. R. Culler. (2007) Optimizing bacteriophage plaque fecundity. J. Theor. Biol. **249**: 582-592

Ackermann, H.-W. and M. S. DuBow. (1987) *In* Viruses of Prokaryotes Vol. 1 & 2. CRC Press, Boca Raton.

Ackermann, H. -W. (1998) Tailed bacteriophages: the order caudovirales. Adv. Virus Res. **51**: 135-201

Ackermann, H. -W. (2003) Bacteriophage observations and evolution. Res. Microbiol. **154**: 245-251

Ackermann, H. -W. (2009) Phage classification and characterization. Methods Mol. Biol. **501**: 127-140

Adam, G., and M. Delbruck. (1968) Reduction of dimensionality in biological diffusion processes. *In* Structural chemistry and molecular biology. A. Rich and N. Davidson (Eds.), pp. 198-215. W. H. Freeman & Company, San Francisco, CA.

Adams, M. H. (1959) Bacteriophages. Wiley Interscience, New York, N.Y.

Aksyuk, A. A., P. G. Leiman, L. P. Kurochkina, M. M. Shneider, V. A. Kostyuchenko, V. V. Mesyanzhinov, and M. G. Rossmann. (2009) The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. EMBO J. Feb 19

Alonso, J. C., G. Lüder, A. C. Stiege, S. Chai, F. Weise, and T. A. Trautner. (1997) The complete nucleotide sequence and functional organization of *Bacillus subtilis* bacteriophage SPP1. Gene. **204**: 201-212

Alonso, J. C., P. Tavares, R. Lurz, and T. A. Trautner. (2006) Bacteriophage SPP1. *In* The Bacteriophages. R. Calendar (Ed.), pp. 331-349. Oxford University Press, New York, USA.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389-3402

Altschul, S. F., J. C. Wootton, E. Michael Gertz, R. Agarwala, A. Morgulis, A. A. Schäffer, and Y. Yu. (2005) Protein database searches using compositionally adjusted substitution matrices. FEBS J. **272**: 5101-5109

Archibald, A. R., and H. E. Coapes. (1976) Bacteriophage SP50 as a marker for cell wall growth in *Bacillus subtilis*. J. Bacteriol. **125**: 1195-1206

Archibald, A. R., I. C. Hancock, and C. R. Harwood. (1993) Cell wall structure, synthesis, and turnover. *Bacillus subtilis* and other Gram-positive bacteria. A. L. Sonenshein, J. A. Hoch and R. Losick (Eds.). pp. 381-410. ASM Press, Washington, DC.

Ashlee, M. E., R. Losick, and R. Kolter. (2007) *Bacillus subtilis* genome diversity. J. Bacteriol. **189**: 1163-1170

Auzat, I., A. Dröge, F. Weise, R. Lurz, and P. Tavares. (2008) Origin and function of the two major tail proteins of bacteriophage SPP1. Mol. Microbiol. **70**: 557-569

Axelrod, D., and M. D. Wang. (1994) Reduction-of-dimensionality kinetics at reaction-limited cell surface receptors. Biophys. J. **66**: 588-600

Barbirz, S., M. Becker, A. Freiberg, and R. Seckler. (2009) Phage tailspike proteins with beta-solenoid fold as thermostable carbohydrate binding materials. Macromol. Biosci. **9**: 169-173

Baxa, U., S. Steinbacher, A. Weintraub, R. Huber, and R. Seckler. (1999) Mutations improving the folding of phage P22 tailspike protein affect its receptor binding activity. J. Mol. Biol. **293**: 693-701

Becker, B., N. de la Fuente, M. Gassel, D. Günther, P. Tavares, R. Lurz, T. A. Trautner and J. Alonso. (1997) Head morphogenesis genes of the *Bacillus subtilis* bacteriophage SPP1. J. Mol. Biol. **268**: 822-839

Benson, A. K., and W. G. Haldenwang. (1993) Regulation of sigma B levels and activity in *Bacillus subtilis*. J. Bacteriol. **175**: 2347-2356

Berg, H. C., and E. M. Purcell. (1977) Physics of chemoreception. Biophys. J. **20**: 193-219

Berg, 0. (1978) On diffusion-controlled dissociation. Chem. Phys. **31**: 47-57

Behrens, B., G. Lüder, M. Behncke, T. A. Trautner, and A. T. Ganesan. (1979) The genome of *B. subtilis* phage SPP1: physical arrangement in phage genes. Mol. Gen. Genet. **175**: 351-357

Bhavsar, A. P., L. K. Erdman, J. W. Schertzer, and E. D. Brown. (2004) Teichoic acid is an essential polymer in *Bacillus subtilis* that is functionally distinct from teichuronic acid. J. Bacteriol. **186**: 7865-7873

Bhavsar, A. P., and E. D. Brown. (2006) Cell wall assembly in *Bacillus subtilis*: how spirals and spaces challenge paradigms. Mol. Microbiol. **60**: 1077-1090

Birdsell, D. C., and R. J. Doyle. (1973) Modification of bacteriophage ϕ25 adsorption to *Bacillus subtilis* by concanavalin A. J. Bacteriol. **113**: 198-202

Bos, M. P., V. Robert, and J. Tommassen. (2007) Biogenesis of the gram-negative bacterial outer membrane. Annu. Rev. Microbiol. **61**: 191-214

Boulanger, P., P. Jacquot, L. Plançon, M. Chami, A. Engel, C. Parquet, C. Herbeuval, and L. Letellier. (2008) Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. J. Biol. Chem. **283**: 13556-13564

Bradley, D. E. (1967) Ultrastructure of bacteriophage and bacteriocins. Bacteriol. Rev. **31**: 230-314

Bradley, P., L. Cowen, M. Menke, J. King, and B. Berger. (2001) BETAWRAP: successful prediction of parallel beta-helices from primary sequence reveals an association with many microbial pathogens. Proc. Natl. Acad. Sci. USA. **98**: 14819-14824

Branda, S. S., J. E. González-Pastor, E. Dervyn, S. D. Ehrlich, R. Losick, and R. Kolter. (2004) Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. J. Bacteriol. **186**: 3970-3979

Brodin, P., I. Rosenkrands, P. Andersen, S. T. Cole and R. Brosch. (2004) ESAT-6 proteins: protective antigens and virulence factors? Trends Microbiol. **12**: 500-508

Bryson, K., L. J. McGuffin, R. L. Marsden, J. J. Ward, J. S.Sodhi, and D. T. Jones. (2005) Protein structure prediction servers at University College London. Nucl. Acids Res. **33**(Web Server issue):W36-38

Burger, K. J. (1980) Membrane binding of bacteriophage SPP1 DNA. Mol. Gen. Genet. **179**: 373-376

Burts, M. L., W. A. Williams, K. DeBord and D. M. Missiakas. (2005) EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. Proc. Natl. Acad. Sci. USA **102**: 1169-1174

Burts, M. L., A. C. DeDent, and D. M. Missiakas. (2008) EsaC substrate for the ESAT-6 secretion pathway and its role in persistent infections of *Staphylococcus aureus*. Mol. Microbiol. **69**: 736-746

Cabrita, L. D., and S. P. Bottomley. (2004) Protein expression and refolding - a practical guide to getting the most out of inclusion bodies. Biotechnol. Annu. Rev. **10**: 31-50

Calendar, R. (Ed). (2007) The Bacteriophages. Oxford University Press, New York, USA.

Cao M. and J. D. Helmann. (2004) The *Bacillus subtilis* extracytoplasmic-function σ^X factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides. J. Bacteriol. **186**: 1136-1146

Casadaban, M. J., and S. N. Cohen. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138:**179-207

Casjens, S. R. (2005) Comparative genomics and evolution of the tailed-bacteriophages. Curr. Opin. Microbiol. **8**: 451-458

Cerritelli, M. E., and F. W. Studier. (1996) Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. J. Mol. Biol. **258**: 286-298

Cerritelli, M. E., J. S. Wall, M. N. Simon, J. F. Conway, and A. C. Steven. (1996) Stoichiometry and domainal organization of the long tail-fiber of bacteriophage T4: a hinged viral adhesin. J. Mol. Biol. **260**: 767-780

Chappell, J. D., A. E. Prota, T. S. Dermody, and T. Stehle. (2002) Crystal structure of reovirus attachment protein sigma1 reveals evolutionary relationship to adenovirus fiber. EMBO J. **21**: 1-11

Cheng Y, N. Promadej, J. W. Kim, and S. Kathariou. (2008) Teichoic acid glycosylation mediated by *gtcA* is required for phage adsorption and susceptibility of *Listeria monocytogenes* serotype 4b. Appl. Environ. Microbiol. **74**: 1653-1655

Choi, K. H., J. McPartland, I. Kaganman, V. D. Bowman, L. B. Rothman-Denes, and M G. Rossmann. (2008) Insight into DNA and protein transport in double-stranded DNA viruses: the structure of bacteriophage N4. J. Mol. Biol. **378**: 726-736

Christensen, J. R. (1965) The kinetics of reversible and irreversible attachment of bacteriophage Ti. Virology. **26**:727-737

Chua, J. E., P. A. Manning, and R. Morona. (1999) The *Shigella flexneri* bacteriophage Sf6 tailspike protein (TSP)/endorhamnosidase is related to the bacteriophage P22 TSP and has a motif common to exo- and endoglycanases, and C-5 epimerases. Microbiology. **145**: 1649-1659

Chung, C. T., S. L. Niemela, and R. H. Miller. (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc. Natl. Acad. Sci. USA **86**: 2172-2175

Coligan J. E., B. M. Dunn, H. L. Ploegh, D. W. Speicher, and P. T. Wingfield. Current Protocols in Protein Science. Vol. 1. Series Editor Virginia Benson Chanda. John Wiley & Sons, Inc.

Collins, F. C., and G. E. Kimball. (1949) Diffusion-controlled reaction rates. J. Colloid Sci. **4**: 425-437

Condon, C. (2003) RNA processing and degradation in *Bacillus subtilis*. Microbiol. Mol. Biol. Rev. **67**: 157-174

Condon, C. (2007) Maturation and degradation of RNA in bacteria. Curr. Opin. Microbiol. **10**: 271-278

Coros, A., B. Callahan, E. Battaglioli, and K. M. Derbyshire. (2008) The specialized secretory apparatus ESX-1 is essential for DNA transfer in *Mycobacterium smegmatis*. Mol. Microbiol. **69**: 794-808

Coyette, J., and J.-M. Ghuysen. (1968) Structure of the cell wall of *Staphylococcus aureus*, strain Copenhagen. IX. Teichoic acid and phage adsorption. Biochemistry **7**: 2385-2389

Daffé, M., and P. Draper. (1998) The envelope layers of mycobacteria with reference to their pathogenicity. Adv. Microb. Physiol. **39**: 131-203

Daniel, A., and D. W. Michelle. (1994) Reduction-of-dimensionality kinetics at reaction-limited cell surface receptors. Biophys. J. **66**: 588-600

Davison, S., E. Couture-Tosi, T. Candela, M. Mock, and A. Fouet. (2005) Identification of the *Bacillus anthracis* (gamma) phage receptor. J. Bacteriol. **187**: 6742-6749

de Frutos, M., S. Brasilès, P. Tavares, and E. Raspaud. (2005a) Effect of spermine and DNase on DNA release from bacteriophage T5. Eur. Phys. J. E. Soft Matter. **17**: 429-434

de Frutos, M., L. Letellier, E. Raspaud. (2005b) DNA ejection from bacteriophage T5: analysis of the kinetics and energetics. Biophys. J. **88**: 1364-1370

de Hoon, M. J., Y. Makita, K. Nakai, and S. Miyano. (2005) Prediction of transcriptional terminators in *Bacillus subtilis* and related species. PLoS Comput. Biol. **1**: 212-221

Dekhtyar, M., A. Morin and V. Sakanyan. (2008) Triad pattern algorithm for predicting strong promoter candidates in bacterial genomes. BMC Bioinformatics **9**: 233. doi:10.1186/1471-2105-9-233

Delcour, A. H. (2008) Outer membrane permeability and antibiotic resistance. Biochim Biophys Acta. 2008 Nov 27

D'Elia, M. A., K. E. Millar, T. J. Beveridge and E. D. Brown. (2006) Wall teichoic acid polymers are dispensable for cell viability in *Bacillus subtilis*. J. Bacteriol. **188**: 8313-8316

Demchick, P., and A. L. Koch. (1996) The permeability of the wall fabric of *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. **178**: 768-773

Deveau, H., M. R. Van Calsteren, and S. Moineau. (2002) Effect of exopolysaccharides on phage-host interactions in *Lactococcus lactis*. Appl. Environ. Microbiol. **68**: 4364-4369. *Erratum* in: Appl. Environ. Microbiol. (2003) **69**: 723

DiMauro, A. J., D. Lin, S. Guo, D. B. Karr, J. J. Tanner, and P. Guo. (2007) Crystallization of Phi29 spindle-shaped nano-bar anti-receptor with glycosidase domain. J. Nanosci. Nanotechnol. **7**: 2616-2622

Douglas, L. J., and M. J. Wolin. (1971) Cell wall polymers and phage lysis of *Lactobacillus plantarum*. Biochemistry. **10**: 1551-1555

Doyle, R. J., and D. C. Birdsell. (1972) Interaction of concanavalin A with the cell wall of *Bacillus subtilis*. J. Bacteriol. **109**: 652-658

Dröge, A., and P. Tavares. (2000) In vitro packaging of DNA of the *Bacillus subtilis* bacteriophage SPP1. J. Mol. Biol. **296**: 103-115

Duda, R. L., K. Martincic, and R. W. Hendrix. (1995) Genetic basis of bacteriophage HK97 prohead assembly. J. Mol. Biol. **247**: 636-647

Duong, F., J. Eichler, A. Price, M. R. Leonard, and W. Wickner. (1997) Biogenesis of the gram-negative bacterial envelope. Cell **91**: 567-573

Duplessis, M., and S. Moineau. (2001) Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. Mol. Microbiol. **41**: 325-336

Duplessis, M., C. M. Lévesque, and S. Moineau. (2006) Characterization of *Streptococcus thermophilus* host range phage mutants. Appl. Environ. Microbiol. **72**: 3036-3041

Dupont, K., J. Thomas, K. V. Finn, J. Jytte and S.-T. Birgitte. (2004) Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. Appl. Environ. Microbiol. **70**: 5825-5832

Earl, A. M., R. Losick, and R. Kolter. (2007) *Bacillus subtilis* genome diversity. J. Bacteriol. **189**: 1163-1170

Earnshaw, W. C., and S. R. Casjens (1980) DNA packaging by the doublestranded DNA bacteriophages. Cell. **21**: 319-331

Edgar, R., A. Rokney, M. Feeney, S. Semsey, M. Kessel, M. B. Goldberg, S. Adhya, A. B. Oppenheim. (2008) Bacteriophage infection is targeted to cellular poles. Mol. Microbiol. **68**: 1107-1116

Ellwood, D. C., and D. W. Tempest. (1969) Control of teichoic acid and teichuronic acid biosynthesis in chemostat cultures of *Bacillus subtilis* var. niger. Biochem. J. **111**: 1-5

Errington, J., J. Bath, and L. J. Wu. (2001) DNA transport in bacteria. Nat. Rev. Mol. Cell Biol. **2**: 538-545

Esche, H., M. Schweiger, and T. A .Trautner. (1975) Gene expression of bacteriophage SPPI. I. Phage directed protein synthesis. Mol. Gen. Genet. **142**: 45-55

Esche, H. (1975) Gene expression of bacteriophage SPP1. II. Regulatory aspects. Mol. Gen. Genet. **142**: 57-66

Esquinas-Rychen, M., and B. Erni. (2001) Facilitation of bacteriophage lambda DNA injection by inner membrane proteins of the bacterial phosphoenol-pyruvate: carbohydrate phosphotransferase system (PTS). J. Mol. Microbiol. Biotechnol. **3**: 361-370

Estrela, A. I., H. M. Pooley, H. de Lencastre, and D. Karamata. (1991) Genetic and biochemical characterization of *Bacillus subtilis* 168 mutants specifically blocked in the synthesis of the teichoic acid poly(3-O-beta-D-glucopyranosyl-N-acetylgalactosamine 1-phosphate): *gneA*, a new locus, is associated with UDP-N-acetylglucosamine 4-epimerase activity. J. Gen. Microbiol. **137**: 943-950

Evilevitch, A., L. Lavelle, C. M. Knobler, E. Raspaud, and W. M. Gelbart. (2003) Osmotic pressure inhibition of DNA ejection from phage. Proc. Natl Acad. Sci. USA. **100**: 9292-9295

Fauquet, C. M., M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball. (Eds.). (2004) Virus Taxonomy: The Classification and Nomenclature of Viruses. The Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier, Academic Press

Fedtke, I., D. Mader, T. Kohler, H. Moll, G. Nicholson, R. Biswas, K. Henseler, F. Götz, U. Zähringer, and A. Peschel. (2007) A *Staphylococcus aureus ypfP* mutant with strongly reduced lipoteichoic

acid (LTA) content: LTA governs bacterial surface properties and autolysin activity. Mol. Microbiol. **65**: 1078-1091

Ferguson, A. D., E. Hofmann, J. W. Coulton, K. Diederichs, and W. Welte. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science. **282**: 2215-2220

Formstone, A., R. Carballido-López, P. Noirot, J. Errington, and D. J. Scheffers. (2008) Localization and interactions of teichoic acid synthetic enzymes in *Bacillus subtilis*. J. Bacteriol. **190**: 1812-1821

Fraser, J. S., K. L. Maxwell, and A. R. Davidson. (2007) Immunoglobulin-like domains on bacteriophage: weapons of modest damage? Curr. Opin. Microbiol. **10**: 382-387

Fredrick, K., T. Caramori, Y. F. Chen, A. Galizzi and J. D. Helmann. (1995) Promoter architecture in the flagellar regulon of *Bacillus subtilis*: highlevel expression of flagellin by the sigma D RNA polymerase requires an upstream promoter element. Proc. Natl. Acad. Sci. USA. **92**: 2582-2586

Freymond, P. P., V. Lazarevic, B. Soldo, and D. Karamata. (2006) Poly(glucosyl-N-acetylgalactosamine 1-phosphate), a wall teichoic acid of *Bacillus subtilis* 168: its biosynthetic pathway and mode of attachment to peptidoglycan. Microbiology. **152**: 1709-1718

Fuller, D. N., D. M. Raymer, J. P. Rickgauer, R. M. Robertson, C. E. Catalano, D. L. Anderson, S. Grimes, and D. E. Smith. (2007) Measurements of single DNA molecule packaging dynamics in bacteriophage λ reveal high forces, high motor processivity, and capsid transformations. J. Mol. Biol. **373**: 1113-1122

Gamow, R. I. (1969) Thermodynamic treatment of bacteriophage T4B adsorption kinetics. J. Virology. **4**: 113-115

Garufi, G., E. Butler, and D. Missiakas. (2008) ESAT-6-like protein secretion in *Bacillus anthracis*. J. Bacteriol. **190**: 7004-7011

Geller, B. L., R. G. Ivey, J. E. Trempy, and B. Hettinger-Smith. (1993) Cloning of a chromosomal gene required for phage infection of *Lactococcus lactis* subsp. lactis C2. J. Bacteriol. **175**: 5510-5519

German, G. J., and R. Misra. (2001) The TolC protein of *Escherichia coli* serves as a cell-surface receptor for the newly characterized TLS bacteriophage. J. Mol. Biol. **308**: 579-585

Gey Van Pittius, N. C., J. Gamieldien, W. Hide, G. D. Brown, R. J. Siezen, and A. D. Beyers. (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. Genome Biol. 2, research0044.1–0044.18.

Gibbs, K. A., D. D. Isaac, J. Xu, R. W. Hendrix, T. J. Silhavy, J. A. Theriot. (2004) Complex spatial distribution and dynamics of an abundant *Escherichia coli* outer membrane protein, LamB. Mol. Microbiol. **53**: 1771-1783

Givan, A. L., K. Glassey, R. S. Green, W. K. Lang, A. J. Anderson, and A. R. Archibald. (1982) Relation between wall teichoic acid content of *Bacillus subtilis* and efficiency of adsorption of bacteriophages SP 50 and phi 25. Arch. Microbiol. **133**: 318-322

González-Huici, V., M. Salas, and J. M. Hermoso. (2004) The push-pull mechanism of bacteriophage phi29 DNA injection. Mol. Microbiol. **52**: 529-540

González-Huici, V., M. Salas, and J. M. Hermoso. (2006) Requirements for *Bacillus subtilis* bacteriophage phi29 DNA ejection. Gene. **374**: 19-25

Goude, R., and T. Parish. (2008) The genetics of cell wall biosynthesis in *Mycobacterium tuberculosis*. Future Microbiol. **3**: 299-313

Grayson, P., A. Evilevitch, M. M. Inamdar, P. K. Purohit, W. M. Gelbart, C. M. Knobler, and R. Phillips. (2006) The effect of genome length on ejection forces in bacteriophage lambda. Virology. **348**: 430-436

Grayson, P., and I. J. Molineux. (2007) Is phage DNA 'injected' into cells-biologists and physicists can agree. Curr. Opin. Microbiol. **10**: 401-409

Gründling, A., and O. Schneewind. (2007) Synthesis of glycerol phosphate lipoteichoic acid in *Staphylococcus aureus*. Proc. Natl. Acad. Sci. USA. **104**: 8478-8483

Guo, S., D. Shu, M. N. Simon, and P. Guo. (2003) Gene cloning, purification, and stoichiometry quantification of phi29 anti-receptor gp12 with potential use as special ligand for gene delivery. Gene. **315**: 145-152

Hatfull, G. F. (2008) Bacteriophage genomics. Curr. Opin. Microbiol. **11**: 447-453

Hayashi, S., and H. C. Wu. (1990) Lipoproteins in bacteria. J. Bioenerg. Biomembr. **22**: 451-471

Hayes, W. (1984) Max Delbrück and the birth of molecular biology. Soc. Res. (New York). **51**: 641-673

Hayhurst, E. J., L. Kailas, J. K. Hobbs, and S. J. Foster. (2008) Cell wall peptidoglycan architecture in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA. **105**: 14603-14608

Heller, K., and V. Braun. (1979) Accelerated adsorption of bacteriophage T5 to *Escherichia coli* F, resulting from reversible tail fiber-lipopolysaccharide binding. J. Bacteriol. **139**: 32-38

Heller, K. J., and V. Braun. (1982) Polymannose O antigens of *Escherichia coli*, the binding sites for the reversible adsorption of bacteriophage T5 + via the L-shaped tail fibers. J. Viro. **141**: 222-227

Heller, K. J., and H. Schwarz. (1985) Irreversible binding to the receptor of bacteriophages T5 and BF23 does not occur with the tip of the tail. J. Bacteriol. **162**: 621-625

Heller, K. J. (1992) Molecular interaction between bacteriophage and the gram-negative cell envelope. Arch. Microbiol. **158**: 235-248

Hendrix, R.W. (2002) Bacteriophages: evolution of the majority. Theor. Popul. Biol. **61**: 471-480

Hendrix, R. W., G. F. Hatfull, and M. C. Smith. (2003) Bacteriophages with tails: chasing their origins and evolution. Res. Microbiol. **154**: 253-257

Hershey, A. D., and M. Chase. (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. J. Gen. Physiol. **36**: 39-56

Ho, T. D., and J. M. Slauch. (2001) OmpC is the receptor for Gifsy-1 and Gifsy-2 bacteriophages of *Salmonella*. J. Bacteriol. **183**: 1495-1498

Hsu, T., S. M. Hingley-Wilson, B. Chen, M. Chen, A. Z. Dai, P. M. Morin, C. B. Marks, J. Padiyar, C. Goulding, M. Gingery, D. Eisenberg, R. G. Russell, S. C. Derrick, F. M. Collins, S. L. Morris, C. H. King, and W. R. Jacobs Jr. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. Proc. Natl. Acad. Sci. USA. **100**: 12420-12425

Inamdar, M. M., W. M. Gelbart, and R. Phillips. (2006) Dynamics of DNA ejection from bacteriophage. Biophys. J. **91**: 411-420

Iwashita, S., and S. Kanegasaki. (1976) Enzymic and molecular properties of base-plate parts of bacteriophage P22. Eur. J. Biochem. **65**: 87-94

Iyer, L. M., A. M. Burroughs and L Aravind. (2006) The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like β-grasp domains. Genome Biology. **7:** R60 (doi:10.1186/gb-2006-7-7-r60)

Jacobson, E. D., and O. E. Landman. (1975) Interaction of protoplasts, L forms, and bacilli of *Bacillus subtilis* with 12 strains of bacteriophage. J. Bacteriol. **124**: 445-448

Jacobson, E. D., and O. E. Landman. (1977) Adsorption of bacteriophages phi29 and 22a to protoplasts of *Bacillus subtilis* 168. J. Virol. **21**: 1223-1227

Jeembaeva, M., M. Castelnovo, F. Larsson, and A. Evilevitch. (2008) Osmotic pressure: resisting or promoting DNA ejection from phage? J. Mol. Biol. **381**: 310-323

Johnson, J. E, and W. Chiu. (2007) DNA packaging and delivery machines in tailed bacteriophages. Curr. Opin. Struct. Biol. **17**: 237-243

Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292**: 195-202

Jorasch, P., F. P. Wolter, U. Zähringer, and E. Heinz. (1998) A UDP glucosyltransferase from *Bacillus subtilis* successively transfers up to four glucose residues to 1,2-diacylglycerol: expression of *ypfP* in *Escherichia coli* and structural analysis of its reaction products. Mol. Microbiol. **29**: 419-430

Jordan, S., M. I. Hutchings, and T. Mascher. (2008) Cell envelope stress response in Gram-positive bacteria. FEMS Microbiol. Rev. **32**: 107-146

Kanamaru, S., P. G. Leiman, V. A. Kostyuchenko, P. R. Chipman, V. V. Mesyanzhinov, F. Arisaka, and M. G. Rossman. (2002) Structure of the cell puncturing device of bacteriophage T4. Nature. **415:** 553-557

Karamata, D., H. M. Pooley, and M. Monod. (1987) Expression of heterologous genes for wall teichoic acid in *Bacillus subtilis* 168. Mol. Gen. Genet. **207**: 73-81

Kearns, D. B., F. Chu, S. S. Branda, R. Kolter, and R. Losick. (2005) A master regulator for biofilm formation by *Bacillus subtilis*. Mol. Microbiol. **55**: 739-749

Kemp, P., L. R. Garcia, and I. J. Molineux. (2005) Changes in bacteriophage T7 virion structure at the initiation of infection. Virology. **340**: 307-317

Kenny, J. G., S. McGrath, G. F. Fitzgerald, and D. Sinderen. (2004) Bacteriophage Tuc2009 encodes a tail-associated cell wall-degrading activity. J. Bacteriol **186**: 3480-3491

Kiino, D. R., R. Licudine, K. Wilt, D. H. Yang, and L. B. Rothman-Denes. (1993) A cytoplasmic protein, NfrC, is required for bacteriophage N4 adsorption. J. Bacteriol. **175**: 7074-7080

Kim, S., and S. B. Lee. (2008) Soluble expression of archaeal proteins in *Escherichia coli* by using fusion-partners. Protein Expr. Purif. **62**: 116-119

Kiriukhin, M. Y., D. V. Debabov, D. L. Shinabarger, and F. C. Neuhaus. (2001) Biosynthesis of the glycolipid anchor in lipoteichoic acid of *Staphylococcus aureus* RN4220: role of YpfP, the diglucosyldiacylglycerol synthase. J. Bacteriol. **183**: 3506-3514

Klumpp, J., J. Dorscht, R. Lurz, R. Bielmann, M. Wieland, M. Zimmer, R. Calendar, and M. J. Loessner. (2008) The terminally redundant, nonpermuted genome of Listeria bacteriophage A511: a model for the SPO1-like myoviruses of gram-positive bacteria. J. Bacteriol. **190**: 5753-5765

Kozak, M. (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene. **361**: 13-37

Kostyuchenko, V. A., P. R. Chipman, P. G. Leiman, F. Arisaka, V. V. Mesyanzhinov, and M. G. Rossmann. (2005) The tail structure of bacteriophage T4 and its mechanism of contraction. Nat. Struct. Mol. Biol. **12**: 810-813

Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessières, A. Bolotin, S. Borchert, *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature. **390**: 249-56

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. **227:** 680-685

Lanni, Y. (1969) Functions of two genes in the first-step-transfer DNA of bacteriophage T5. J. Mol. Biol. **44**: 173-183

Lazarevic, V., B. Soldo, N. Médico, H. Pooley, S. Bron, and D. Karamata. (2005) *Bacillus subtilis* α-phosphoglucomutase is required for normal cell morphology and biofilm formation. Appl. Environ. Microbiol. **71**: 39-45

Leiman, P. G., S. Kanamaru, V. V. Mesyanzhinov, F. Arisaka, and M. G. Rossmann. (2003) Structure and morphogenesis of bacteriophage T4. Cell. Mol. Life Sci. **60**: 2356-2370

Leiman, P. G., P. R. Chipman, V. A. Kostyuchenko, V. V. Mesyanzhinov, and M. G. Rossman. (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. Cell. **118**: 419-429

Leiman, P. G., A. J. Battisti, V. D. Bowman, K. Stummeyer, M. Mühlenhoff, R. Gerardy-Schahn, D. Scholl and I. J. Molineux. (2007) The structures of bacteriophages K1E and K1-5 explain processive degradation of polysaccharide capsules and evolution of new host specificities. J. Mol. Biol. **371**: 836-849

Leiman, P. G., and I. J. Molineux. (2008) Evolution of a new enzyme activity from the same motif fold Mol. Microbiol. **69**: 287-290

Letellier, L., P. Boulanger, L. Plançon, P. Jacquot, and M. Santamaria. (2004) Main features on tailed phage, host recognition and DNA uptake. Front. Biosci. **9**: 1228-1339

Li, X., and H. L. T. Mobley. (2001) Bacterial pili and fimbriae. Encyclopedia of Life Sciences. 1-7

Lindeberg, A. A. (1973) Bacteriophage receptors. Annu. Rev. Microbiol. **27**: 205-241

Lubbers, M. W., N. R. Waterfield, T. P. Beresford, R. W. Le Page, and A. W. Jarvis. (1995) Sequencing and analysis of the prolate-headed lactococcal bacteriophage c2 genome and identification of the structural genes. Appl. Environ. Microbiol. **61**: 4348-4356

Ludwig, H., G. Homuth, M. Schmalisch, F. M. Dyka, M. Hecker, and J. Stulke. (2001) Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the *gapA* operon. Mol. Microbiol. **41:** 409-422

Lurz, R., E. V. Orlova, D. Günther, P. Dube, A. Dröge, F. Weise, M. van Heel, and P. Tavares. (2001) Structural organisation of the head-to-tail interface of a bacterial virus. J. Mol. Biol. **310**: 1027-3107

McGrath, S., H. Neve, J. F. Seegers, R. Eijlander, C. S. Vegge, L. Brondsted, K. J. Heller, G. F. Fitzgerald, F. K. Vogensen, and D. van Sinderen. (2006) Anatomy of a lactococcal phage tail. J. Bacteriol. **188**: 3972-3982

Madigan, M. T., J. M. Martinko, and J. Parker. (1997) Brock; Biology of Microorganisms, 8th ed., Prentice-hall, Inc. New Jersey, USA.

Maniloff, J., and H.-W. Ackermann. (1998) Taxonomy of bacterial viruses: establishment of tailed virus genera and the other *Caudovirales*. Arch. Virol. **143**: 2051-2063

Marchler-Bauer, A., and S. H. Bryant. (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res. **32**: 327-331

Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. DeWeese-Scott, L. Y. Geer, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, D. Zhang, and S. H. Bryant. (2005) CDD: a Conserved Domain Database for protein classification. Nucleic Acids Res. **33**: 192-196

Marchler-Bauer, A., J. B. Anderson, M. K. Derbyshire, C. DeWeese-Scott, N. R. Gonzales, M. Gwadz, L. Hao, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, D. Krylov, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, S. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, N. Thanki, R. A. Yamashita, J. J. Yin, D. Zhang, and S. H. Bryant. (2007) CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res. **35**: 237-240

Margot, P., and D. Karamata. (1996) The *wprA* gene of *Bacillus subtilis* 168, expressed during exponential growth, encodes a cell-wall-associated protease. Microbiology. **142**: 3437-3444

Maruyama, R., M. Nishizawa, Y. Itoi, S. Ito and M. Inoue. (2002) The enzymes with benzil reductase activity conserved from bacteria to mammals. J. Biotechnol. **94**: 157-169

Matias, V. R., and T. J. Beveridge. (2005) Cryo-electron microscopy reveals native polymeric cell wall structure in *Bacillus subtilis* 168 and the existence of a periplasmic space. Mol. Microbiol. **56**: 240-251

Matias, V. R., and T. J. Beveridge. (2008) Lipoteichoic acid is a major component of the *Bacillus subtilis* periplasm. J. Bacteriol. **190**: 7414-7418

Mauël, C., M. Young, and D. Karamata. (1991) Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid in *Bacillus subtilis* strain 168, are organized in two divergent transcription units. J. Gen. Microbiol. **137**: 929-941

McPartland, J., and L. B. Rothman-Denes. (2009) The tail sheath of bacteriophage N4 interacts with the *Escherichia coli* receptor. J. Bacteriol. **191**: 525-532

Milanesi, G., and G. Cassani. (1972) Transcription after bacteriophage SPP1 infection in *Bacillus subtilis*. J. Virology. **10**: 187-192

Miller, J. H. (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Mitraki, A., S. Miller, and M. J. van Raaij. (2002) Review: conformation and folding of novel beta-structural elements in viral fiber proteins: the triple beta-spiral and triple beta-helix. J. Struct. Biol. **137**: 236-247

Mitraki, A., K. Papanikolopoulou, and M. J. van Raaij. (2006) Natural triple beta-stranded fibrous folds. Adv. Protein Chem. **73**: 97-124

Moak, M., and I. J. Molineux. (2000) Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection. Mol. Microbiol. **37**: 345-355

Moldovan, R., E. Chapman-McQuiston, and X. L. Wu. (2007) On kinetics of phage adsorption. Biophys. J. **93**: 303-315

Molineux, I. J. (2001) No syringes please, ejection of T7 DNA from the virion is enzyme-driven. Mol. Microbiol. **40**: 1-8

Molineux, I. J. (2006) Fifty-three years since Hershey and Chase; much ado about pressure but which pressure is it? Virology. **344**: 221-229

Monteville, M. R., B. Ardestani, and B. L. Geller. (1994) Lactococcal bacteriophages require a host cell wall carbohydrate and a plasma membrane protein for adsorption and ejection of DNA. Appl. Environ. Microbiol. **60**: 3204-3211

Morona, R., C. Krämer, and U. Henning. (1985) Bacteriophage receptor area of outer membrane protein OmpA of *Escherichia coli* K-12. J. Bacteriol. **164**: 539-543

Müller, J. J., S. Barbirz, K. Heinle, A. Freiberg, R. Seckler, and U. Heinemann. (2008) An intersubunit active site between supercoiled parallel beta helices in the trimeric tailspike endorhamnosidase of *Shigella flexneri* Phage Sf6. Structure. **16**: 766-775

Neuhaus, F. C., and J. Baddiley. (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. Microbiol. Mol. Biol. Rev. **67**: 686-723

Nishihara, K., M. Kanemori, M. Kitagawa, H. Yanagi, and T. Yura. (1998) Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. Appl. Environ. Microbiol. **64**: 1694-1699

Norrby, E. (2008) Nobel Prizes and the emerging virus concept. Arch. Virol. **153**: 1109-1123

Okubo, S., B. Strauss, and M. Stodolsky. (1964) The possible role of recombination in the infection of competent *Bacillus subtilis* by bacteriophage deoxyribonucleic acid. Virology. **24**: 552-562

Okubo, S., and T. Yanagida. (1968) Isolation of a suppressor mutant in *Bacillus subtilis*. J. Bacteriol. **95**: 1187-1188

Oliveira, L., A. O. Henriques, and P. Tavares. (2006) Modulation of the viral ATPase activity by the portal protein correlates with DNA packaging efficiency. J. Biol. Chem. **281**: 21914-21923

Orlova, E. V., B. Gowen, A. Dröge, A. Stiege, F. Weise, R. Lurz, M. van Heel, and P. Tavares. (2003) Structure of a viral DNA gatekeeper at 10 A resolution by cryo-electron microscopy. The EMBO J. **22**: 1255-1262

Oudega, B., M, Vandenbol, and G. Koningstein. (1997) A 12 kb nucleotide sequence containing the alanine dehydrogenase gene at 279 degrees on the *Bacillus subtilis* chromosome. Microbiology. **143**: 1489-1491

Pallen, M. J. (2002) The ESAT-6/WXG100 superfamily - and a new Gram-positive secretion system? Trends Microbiol. **10**: 209-212

Pedulla, M. L., M. E. Ford, J. M. Houtz, T. Karthikeyan, C. Wadsworth, J. A. Lewis, D. Jacobs-Sera, J. Falbo, J. Gross, N. R. Pannunzio, W. Brucker, V. Kumar, J. Kandasamy, L. Keenan, S. Bardarov, J. Kriakov, J. G. Lawrence, W. R. Jr. Jacobs, R. W. Hendrix, and G. F. Hatfull. (2003) Origins of highly mosaic mycobacteriophage genomes. Cell. **113**: 171-182

Pennazio, S. (2006) The origin of phage virology. Riv. Biol. **99**: 103-129

Perego, M., P. Glaser, A. Minutello, M. A. Strauch, K. Leopold, and W. Fischer. (1995) Incorporation of D-alanine into lipoteichoic acid and wall teichoic acid in *Bacillus subtilis*. Identification of genes and regulation. J. Biol. Chem. **30**: 15598-15606

Pitcher, D. G., N. A. Saunders, and R. J. Owen. (1989) Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. Lett. Appl. Microbiol. **8**: 151-156

Piuri, M., and G. F. Hatfull. (2006) A peptidoglycan hydrolase motif within the mycobacteriophage TM4 tape measure protein promotes efficient infection of stationary phase cells. Mol. Microbiol. **62**: 1569-1585

Plisson, C., H. E. White, I. Auzat, A. Zafarani, C. São-José, S. Lhuillier, P. Tavares, and E. V. Orlova. (2007) Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. EMBO J. **26**: 3720-3728

Pogliano, K., J. Pogliano, and E. Becker. (2003) Chromosome segregation in Eubacteria. Curr. Opin. Microbiol. **6**: 586-593

Price, K., S. Roels, and R. Losick. (1997) A *Bacillus subtilis* gene encoding a protein similar to nucleotide sugar transferases influences cell shape and viability. J. Bacteriol. **179**: 4959-4961

Puck, T. T., A. Garen and J. Cline. (1951) The mechanism of virus attachment to host cells. J. Exp. Med. **93**: 65-88

Purohit, P. K., M. M. Inamdar, P. D. Grayson, T. M. Squires, J. Kondev, and R. Phillips. (2005) Forces during bacteriophage DNA packaging and ejection. Biophys. J. **88**: 851-866

Pym, A. S., P. Brodin, R. Brosch, M. Huerre, S. T. Cole. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. Mol. Microbiol. **46**: 709-717

Quiberoni A, J. I. Stiefel, and J. A. Reinheimer. (2000) Characterization of phage receptors in *Streptococcus thermophilus* using purified cell walls obtained by a simple protocol. J. Appl. Microbiol. **89**: 1059-1065

Quisel, J. D., D. C. Lin, and A. D. Grossman. (1999) Control of development by altered localization of a transcription factor in *Bacillus subtilis*. Molecular Cell. **4**: 665-672

Räisänen, L, K. Schubert, T. Jaakonsaari, and T. Alatossava. (2004) Characterization of lipoteichoic acids as *Lactobacillus delbrueckii* phage receptor components. J. Bacteriol. **186**: 5529-5532

Räisänen, L., C. Draing, M. Pfitzenmaier, K. Schubert, T. Jaakonsaari, S. von Aulock, T. Hartung, and T. Alatossava. (2007) Molecular interaction between lipoteichoic acids and *Lactobacillus delbrueckii* phages depends on D-alanyl and alpha-glucose substitution of poly(glycerophosphate) backbones. J. Bacteriol. **189**: 4135-4140

Rakonjac, J., P. W. O'Toole, and M. Lubbers. (2005) Isolation of lactococcal prolate phage-phage recombinants by an enrichment strategy reveals two novel host range determinants. J. Bacteriol. **187**: 3110-3121

Raspaud, E., T. Forth, C. São-José, P. Tavares, and M. de Frutos. (2007) A kinetic analysis of DNA ejection from tailed phages revealing the prerequisite activation energy. Biophys. J. **93**: 3999-4005

Real, G., S. Autret, E. J. Harry, J. Errington, and A. O. Henriques. (2005) Cell division protein DivIB influences the Spo0J/Soj system of chromosome segregation in *Bacillus subtilis*. Mol. Microbiol. **55**: 349-367

Real, G., and A. O. Henriques. (2006) Localization of the *Bacillus subtilis murB* gene within the *dcw* cluster is important for growth and sporulation. J. Bacteriol. **188**: 1721-1732

Renshaw, P. S., K. L. Lightbody, V. Veverka, F. W. Muskett, G. Kelly, T. A. Frenkiel, S. V. Gordon, R. G. Hewinson, B. Burke, J. Norman, R. A. Williamson, and M. D. Carr. (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. EMBO J. **24**: 2491-2498

Ricagno, S., V. Campanacci, S. Blangy, S. Spinelli, D. Tremblay, S. Moineau, M. Tegoni, and C. Cambillau. (2006) Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bIL170. J. Virol. **80**: 9331-9335

Rieger-Hug, D., and S. Stirm. (1981) Comparative study of host capsule depolymerases associated with Klebsiella bacteriophages.Virology. **113**: 363-378

Riva, S., M. Polsinelli, and A. Falaschi. (1968) A new phage of *Bacillus subtilis* with infectious DNA having separable strands. J. Mol. Biol. **35**: 347-356

Rodríguez, C., R. Van der Meulen, F. Vaningelgem, G. Font de Valdez, R. Raya, L. De Vuyst, and F. Mozzi. (2008) Sensitivity of capsular-producing *Streptococcus thermophilus* strains to bacteriophage adsorption. Lett. Appl. Microbiol. **46**: 462-468

Roessner, C. A. and G. M. Ihler. (1984) Proteinase sensitivity of bacteriophage lambda tail proteins gpJ and pH* in complexes with the lambda receptor. J. Bacteriol. **157**: 165-170

Rossmann, M. G., V. V. Mesyanzhinov, F. Arisaka, and P. G. Leiman. (2004). The bacteriophage T4 DNA injection machine. Curr. Opin. Struct. Biol. **14**: 171-180

Salzberg, L. I., J. D. Helmann. (2008) Phenotypic and transcriptomic characterization of *Bacillus subtilis* mutants with grossly altered membrane composition. J. Bacteriol. **190**: 7797-7807

Sambrook, J., and D.W. Russell. (2001) Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Santos, M. A., H. de Lencastre, and L. J. Archer. (1983) *Bacillus subtilis* mutation blocking irreversible binding of bacteriophage SPP1. J. Gen. Microbiol. **129**: 3499-3504

Santos, M. A. (1991) Bacteriófagos de *Bacillus subtilis* do grupo SPP1: características gerais, especificidade de adsorção e organização genómica. Ph.D. thesis. University of Lisbon, Lisbon, Portugal.

São-José, C., C. Baptista, and M. A. Santos. (2004) *Bacillus subtilis* operon encoding a membrane receptor for bacteriophage SPP1. J. Bacteriol. **186**: 8337-8346

São-José, C., S. Lhuillier, R. Lurz, R. Melki, J. Lepault, M. A. Santos, and P. Tavares. (2006) The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of the bacteriophage SPP1 DNA. J. Biol. Chem. **281**: 11464-11470

São-José, C., M. de Frutos, E. Raspaud, M. A. Santos, and P. Tavares. (2007) Pressure built by DNA packing inside virions: enough to drive DNA ejection in vitro, largely insufficient for delivery into the bacterial cytoplasm. J. Mol. Biol. **374**: 346-355

Scholl, D., J. Kieleczawa, P. Kemp, J. Rush, C. C. Richardson, C. Merril, S. Adhya, and I. J. Molineux. (2004) Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. J. Mol. Biol. **335**: 1151-1171

Scholl, D., S. Adhya, and C. Merril. (2005) *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. Appl. Environ. Microbiol. **71**: 4872-4874

Schwartz, M. (1976) The adsorption of coliphage lambda to its host: effect of variations in the surface density of receptor and in phage-receptor affinity. J. Mol. Biol. **103**: 521-536

Sciara, G., S. Blangy, M. Siponen, S. McGrath, D. van Sinderen, M. Tegoni, C. Cambillau, and V. Campanacci. (2008) A topological model of the baseplate of lactococcal phage Tuc2009. J. Biol. Chem. **283**: 2716-2723

Sierro, N., Y. Makita, M. J. L. de Hoon and K. Nakai. (2008) DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. Nucleic Acids Res. **36** (Database issue):D93-D96; doi:10.1093/nar/gkm910

Simons, R. W., F. Houman and N. Kleckner. (1987) Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene. **53**: 85-96

Siranosian, K. J., K. Ireton, and A. D. Grossman. (1993) Alanine dehydrogenase (*ald*) is required for normal sporulation in *Bacillus subtilis*. J. Bacteriol. **175**: 6789-6796

Soldo, B., V. Lazarevic, P. Margot, and D. Karamata. (1993) Sequencing and analysis of the divergon comprising *gtaB*, the structural gene of UDP-glucose pyrophosphorylase of *Bacillus subtilis* 168. J. Gen. Microbiol. **139**: 3185-3195

Spinelli, S., A. Desmyter, C. T. Verrips, H. J. de Haard, S. Moineau, and C. Cambillau. (2006a) Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses. Nat. Struct. Mol. Biol. **13**: 85-89

Spinelli, S., V. Campanacci, S. Blangy, S. Moineau, M. Tegoni, and C. Cambillau. (2006b) Modular structure of the receptor binding proteins of *Lactococcus lactis* phages. The RBP structure of the temperate phage TP901-1. J. Biol. Chem. **281**: 14256-14262

Stanley, S. A., S. Raghavan, W. W. Hwang, and J. S. Cox. (2003) Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. Proc. Natl. Acad. Sci. USA. **100**: 13001-13006

Steinbacher, S., R. Seckler, S. Miller, B. Steipe, R. Huber, and P. Reinemer. (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. Science. **265**: 383-386

Steinbacher, S., U. Baxa, S. Miller, A. Weintraub, R. Seckler, and R. Huber. (1996) Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. Proc. Natl. Acad. Sci. USA. **93**: 10584-10588

Steinbacher, S., S. Miller, U. Baxa, N. Budisa, A. Weintraub, R. Seckler, and R. Huber. (1997) Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 A, fully refined structure of the endorhamnosidase at 1.56 A resolution, and the molecular basis of O-antigen recognition and cleavage. J. Mol. Biol. **267**: 865-880

Stent, G. S., and E. L. Wollman. (1952) On the two-step nature of bacteriophage adsorption. Biochim. Biophys. Acta. **8**: 260-269

Stuer-Lauridsen, B., T. Janzen, J. Schnabl, and E. Johansen. (2003) Identification of the host determinant of two prolate-headed phages infecting *Lactococcus lactis*. Virology. **309**: 10-17

Tavares, P., R. Lurz, A. Stiege, B. Ruckert, and T. A. Trautner. (1996) Sequence heedful packaging and fate of the cleaved DNA ends in bacteriophage SPP1. J. Mol. Biol. **264**: 954-967

Traurig, M., and R. Misra. (1999) Identification of bacteriophage K20 binding regions of OmpF and lipopolysaccharide in *Escherichia coli* K-12. FEMS Microbiol. Lett. **181**: 101-108

Tremblay, D. M., M. Tegoni, S. Spinelli, V. Campanacci, S. Blangy, C. Huyghe, A. Desmyter, S. Labrie, S. Moineau, and C. Cambillau. (2006) Receptor-binding protein of *Lactococcus lactis* phages: identification and characterization of the saccharide receptor-binding site. J. Bacteriol. **188**: 2400-2410

Tzlil, S., J. T. Kindt, W. M. Gelbart, and A. Ben-Shaul. (2003). Forces and pressures in DNA packaging and release from viral capsids. Biophys. J. **84**: 1616-1627

Uchiyama, J., M. Rashel, I. Takemura, H. Wakiguchi, S. Matsuzaki. (2008) In silico and in vivo evaluation of bacteriophage phiEF24C, a candidate for treatment of *Enterococcus faecalis* infections. Appl. Environ. Microbiol. **74**: 4149-4163

Vagner, V., E. Dervyn, and S. D. Ehrlich. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. Microbiology. **144**: 3097-3104

van den Ent, F., and J. Löwe. (2005) Crystal structure of the ubiquitin-like protein YukD from *Bacillus subtilis*. FEBS Letters. **579**: 3837-3841

van Raaij, M. J., A. Mitraki, G. Lavigne, and S. Cusack. (1999) A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature. **401**: 935-938

van Raaij, M. J., G. Schoehn, M. R. Burda, and S. Miller. (2001a). Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. J. Mol. Biol. **314**: 1137-1146

van Raaij, M. J., G. Schoehn, M. Jaquinod, K. Ashman, M. R. Burda, and S. Miller. (2001b). Identification and crystallisation of a heat- and protease-stable fragment of the bacteriophage T4 short tail fibre. Biol. Chem. **382**: 1049-1055

Vegge, C. S., L. Brøndsted, H. Neve, S. McGrath, D. van Sinderen, and F. K. Vogensen. (2005) Structural characterization and assembly of the distal tail structure of the temperate lactococcal bacteriophage TP901-1. J. Bacteriol. **187**: 4187-4197

Villanueva, N., and M. Salas. (1981) Adsorption of bacteriophage phi 29 to *Bacillus subtilis* through the neck appendages of the viral particle. J. Virol. **38**: 15-19

Vinga, I., C. São-José, P. Tavares, M. A. Santos. (2006a) Bacteriophage entry in the host cell. *In* Modern Bacteriophage Biology and Biotechnology, Wegrzyn G. (Ed.), pp. 165-203. Kerala, India: Research Signpost.

Vinga, I., A. Dröge, A. C. Stiege, R. Lurz, M. A. Santos, R. Daugelavicius, and P. Tavares. (2006b) The minor capsid protein gp7 of bacteriophage SPP1 is required for efficient infection of *Bacillus subtilis*. Mol. Microbiol. **61**: 1609-1621

Walter, M., C. Fiedler, R. Grassl, M. Biebl, R. Rachel, X. L. Hermo-Parrado, A. L. Llamas-Saiz, R. Seckler, S. Miller, and M. J. van Raaij. (2008) Structure of the receptor-binding protein of bacteriophage Det7: a podoviral tail spike in a myovirus. J. Virol. **82**: 2265-2273

Wandersman, C., and P. Delepelaire. (2004) Bacterial iron sources: from siderophores to hemophores. Annu. Rev. Microbiol. **58**: 611-647

Wang, J., M. Hofnung, and A. Charbit. (2000) The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of *Escherichia coli* K-12. J. Bacteriol. **182**: 508-512

Ward, J. B. (1981) Teichoic and teichuronic acids: biosynthesis, assembly, and location. Microbiol. Rev. **45**: 211-243

Watanabe, K., and S. Takesue. (1975) Use of L-rhamnose to study irreversible adsorption of bacteriophage PL-1 to a strain of *Lactobacillus casei*. J. Gen. Virol. **28**: 29-35

Waugh, D. S. (2005) Making the most of affinity tags. Trends Biotechnol. **23**: 316-320

Way, S. S., and C. B. Wilson. (2005) The *Mycobacterium tuberculosis* ESAT-6 homologue in *Listeria monocytogenes* is dispensable for growth in vitro and in vivo. Infect. Immun. **73**: 6151-6153

Weart, R. B., A. H. Lee, A. C. Chien, D. P. Haeusser, N. S. Hill, P. A. Levin. (2007) A metabolic sensor governing cell size in bacteria. Cell. **130**: 335-347

Weigele, P. R., E. Scanlon, and J. King. (2003) Homotrimeric, beta-stranded viral adhesins and tail proteins. J. Bacteriol. **185**: 4022-4030

Weigele, P. R., C. Haase-Pettingell, P. G. Campbell, D. C. Gossard, and J. King. (2005) Stalled folding mutants in the triple beta-helix domain of the phage P22 tailspike adhesin. J. Mol. Biol. **354**: 1103-1117

Wendlinger, G., M. J. Loessner, and S. Scherer. (1996) Bacteriophage receptors on *Listeria monocytogenes* cells are the N-acetylglucosamine and rhamnose substituents of teichoic acids or the peptidoglycan itself. Microbiology. **142:** 985-992

Williams, N., D. K. Fox, C. Shea, and S. Roseman. (1986) Pel, the protein that permits lambda DNA penetration of *Escherichia coli*, is encoded by a gene in ptsM and is required for mannose utilization by the phosphotransferase system. Proc. Natl. Acad. Sci. USA. **83**: 8934-8938

Wilkinson, M., J. Doskow, and S. Lindsey. (1991) RNA blots: staining procedures and optimization of conditions. Nucleic Acids Res. **19**: 679

Wu, L. J., and J. Errington. (1997) Septal localization of the SpoIIIE chromosome partitioning protein in *Bacillus subtilis*. EMBO J. **16**: 2161-2169

Yasbin, R. E., G. A. Wilson, and F. E. Young. (1973) Transformation and transfection in lysogenic strains of *Bacillus subtilis* 168. J. Bacteriol. **113**: 540-548

Yasbin, R. E., V. C. Maino, and F. E. Young. (1976) Bacteriophage resistance in *Bacillus subtilis* 168, W23, and interstrain transformants. J. Bacteriol. **125**: 1120-1126

Yokoyama, K., H. Mizuguchi, Y. Araki, S. Kaya, and E. Ito. (1989) Biosynthesis of linkage units for teichoic acids in gram-positive bacteria: distribution of related enzymes and their specificities for UDP-sugars and lipid-linked intermediates. J. Bacteriol. **171**: 940-946

Young, F. E. (1967) Requirement of glucosylated teichoic acid for adsorption of phage in *Bacillus subtilis* 168. Proc. Natl. Acad. Sci. USA. **58**: 2377-2384

Young, M., C. Maul, P. Margot, and D. Karamata. (1989) Pseudo-allelic relationship between non-homologous genes concerned with biosynthesis of polyglycerol phosphate and polyribitol phosphate teichoic acids in *Bacillus subtilis* strains 168 and W23. Mol. Microbiol. **3**: 1805-1812

Young, R. (1992) Bacteriophage lysis: mechanism and regulation. Microbiol. Rev. **56**: 430-481

Yu, F., and S. Mizushima. (1982) Roles of lipopolysaccharide and outer membrane protein OmpC of *Escherichia coli* K-12 in the receptor function for bacteriophage T4. J. Bacteriol. **151**: 718-722

Yu, X. C., A. H. Tran, Q. Sun, and W. Margolin. (1998) Localization of cell division protein FtsK to the *Escherichia coli* septum and identification of a potential N-terminal targeting domain. J. Bacteriol. **180**: 1296-1304