Université de Montréal

Molecular characterization of XerS/*difSL* site-specific recombination system in *Streptococcus suis*

Par

Fabio Andres Castillo Martinez

Département de microbiologie et immunologie

Faculté de Médecine

Mémoire présenté en vue de l'obtention du grade de

Maîtrise en microbiologie et immunologie

Janvier, 2020

Université de Montréal
Faculté des études supérieures et postdoctorales

Ce mémoire intitulé

Molecular characterization of XerS/*difSL* site-specific recombination system in *Streptococcus suis*

Présenté par
Fabio Castillo Martinez

A été évalué par un jury composé des personnes suivantes

Marylise Duperthuy
Président-rapporteur

George Szatmari
Directeur de recherche

Josée Harel
Membre du jury

Résumé

L'état circulaire du chromosome bactérien pose un problème particulier lors de la réplication. Un nombre impair d'événements de recombinaison homologue donne des chromosomes dimères concaténés qui ne peuvent pas être divisés en cellules filles. Pour résoudre ce problème, les bactéries ont mis au point un mécanisme de résolution des dimères basé sur un système de recombinaison spécifique au site.

Ceci est effectué par le système Xer/*dif*. Dans ce système, les protéines Xer effectuent une réaction de recombinaison dans le site *dif* au niveau du septum cellulaire immédiatement avant la division cellulaire. Dans la plupart des bactéries, cette réaction est effectuée par deux recombinases, XerC et XerD. Cependant, *Streptococcus suis*, un agent pathogène zoonotique important utilise un système de recombinaison différent, constitué d'une seule enzyme recombinase appelée XerS, qui catalyse la réaction de recombinaison dans un site *dif* non conventionnel. Pour caractériser le mode de clivage de XerS, des expériences EMSA ont été réalisées en utilisant des fragments de PCR marqués par HEX et des "suicide substrates". Nos données suggèrent que 1.) XerS est capable de lier la séquence entière de *difSL*; 2.) XerS lie plus efficacement le côté gauche des mutants *difSL* incomplets que le côté droit; 3.) XerS coupe les brins supérieur et inférieur du site *difSL*, avec une réaction plus efficace au bas. 4.) Modifications des nucléotides de la région la plus externe ou de la région centrale changent les préférences de clivage. 5.) XerS n'a montré aucune activité spécifique sur un autre site *dif* non conventionnel des Firmicutes, 6.) XerS interagit avec la sous-unité FtsK-γ.

L'ensemble des résultats présentés permet de mieux comprendre le fonctionnement de la recombinaison XerS dans le système de recombinase unique de *Streptococcus* et comment cette recombinaison est régulée par des facteurs de l'hôte.

Mots-clés: Recombinaison specifique du site, XerS, Tyrosine recombinases, XerC/XerD, *S. suis,* site *dif*, dimérisation.

Abstract

The circular state of the bacterial chromosome presents a specific problem during replication. An odd number of homologous recombination events results in concatenated dimer chromosomes that cannot be partitioned into daughter cells. To solve this problem, bacteria have developed a mechanism of dimer resolution based on site-specific recombination system.

This is performed by the Xer/*dif* system. In this system, the Xer proteins perform a recombination reaction in the *dif* site at the cell septum immediately prior to cell division. In most bacteria this reaction is performed by two recombinases, XerC and XerD. However, an important zoonotic pathogen; *Streptococcus suis* harbors a different recombination system, composed by a single recombinase enzyme called XerS, that catalyzes the recombination reaction in an unconventional *dif* site; *difSL*. A region characterized by two imperfect inverted repeat regions that flank a central region of 11 bp.To characterize the mode of cleavage of XerS, EMSA experiments were performed by using HEX-labelled PCR fragments and "nicked suicide substrates". Our data suggests that; 1.) XerS is able to bind the entire *difSL* sequence; 2.) XerS binds more efficiently the left half side on incomplete *difSL* mutants than the right half side; 3.) XerS cleaves both the top and bottom strands of the *difSL* site, with a more efficient reaction at the bottom strand; 4.) Nucleotides at the outermost region of a T rich region seem to be determinant for binding selectivity and modifications of the extra spacing between the inverted repeat arms as well as length modifications of the central region change cleavage preference. 5.) XerS did not show any specific activity on another unconventional *dif* site in Firmicutes, as tested on *difH*. 6.) XerS interacts with FtsK-γ subunit.

This research aims to understand how XerS recombination works in the single recombinase system of *Streptococcus* and how this recombination is regulated by host factors. Exploration of these recombinases will provide a better understanding of the mechanisms of DNA exchange and genome stability in bacteria. It can also increase our knowledge of the evolution and speciation of recombinogenic bacteria.

Keywords: Site-specific recombination, XerS, Tyrosine recombinases, XerC/XerD, *S. suis, dif* sites, dimerization.

Table des matières

List of tables

List of figures

Chapter I: Review paper

Chapter II: Article

List of abbreviations

AMINO ACIDS

D: acid aspartic
E: acid glutamic
A: alanine
R: arginine
N : asparagine
C: cysteine
Q: glutamine
G: glycine
H: histidine
I: isoleucine
L: leucine
K: lysine
M: methionine
F: phenylalanine
P: proline
S: serine
T: threonine
W: tryptophan
Y : tyrosine

NUCLEIC ACIDS

A: Adenine
C: Cytosine
G: Guanine
N: A, G, C ou T
R: Purine
T: Thymine
P: Pyrimidine

UNITS OF MEASUREMENT

Å: angstrom
bp: base pair
cm: centimeter
Da: Dalton
°C: degree Celsius
g: gram
h: hour
kb: kilobase
kDa: kiloDalton
µg: microgram
µl: microliter
µM: micromolar
ml : milliliter
mM : millimolar
min: minute
ng: nanogram
rpm: revolutions per minute
s: second
v/cm: volt per centimeter

GREEK LETTERS

α: alpha
β: beta
γ: gamma
δ: delta
ε: epsilon
ϕ:phi
λ: lambda

OTHERS

AAA+: *ATPase associated with various activities*
5'OH : 5 prime hydroxyl end
6-HEX: 6-carboxy-2',4,4',5',7,7'-hexachlorofluorescine
AIMS: *Architecture Imparting Sequences*
ASPS: *Archaea Short Polarized Sequences*
ATP: adenosine triphosphate
ATPase: adenosine triphosphatase
BN: *Bottom nick*
BSA: *Bovine Serum Albumin*
C-terminal (e): carboxy-terminal
DNA: Deoxyribonucleic acid
DAPI: 4',6'-diamidino-2-phenylindole

DAZ: *dif activity zone*
EDTA: Ethylenediaminotetraacetic acid
EMSA: Electrophoretic mobility shift assay
FRET: *Fluorescence resonance energy transfer*
GGI: gonococcal genomic island
HEX: Single isomer fluorescein dye
IPTG: isopropyl B-D thiogalactopyranoside
KOPS: *FtsK-orienting polar sequences*
LB: Luria-Bertani
MBP: *maltose binding protein*
N-terminal (e): amino-terminal
NEB: *New England Biolabs*
PCR: Polymerase chain reaction
RNA: Ribonucleic acid
SDS: sodium dodecyl sulfate
SPR: *Surface plasmon resonance*
SRS: *SpoIIIE Recognition Sequence*
TBE: tris-borate EDTA buffer
TDMNG: Tris DTT $MgCl_2$ NaCl Glycerol
TENG: Tris EDTA NaCl Glycerol
THA: Todd-Hewitt agar
THY: Todd-Hewitt + Yeast extract
TN: *Top nick*

Para las mujeres de mi universo; mi madre, mi mami Virginia, mi hermana y mi esposa.

Acknowledgments

I am really grateful for the opportunity of having met such wonderful souls during this moment of my life. I would like to say thank you to Professor George Szatmari for all his support and valuable lessons professional and life wise. And thank you Amal Benmohamed and Asmaâ Agoussar for being there, always bringing a wonderful and positive energy to the laboratory, I will always remember your thoughtful ideas to improve the harmony in the lab. I also want to thank Ali Farrokhi for all his support and professionalism.

Thank you to the members of the jury for your time and valuable input for the completion of this thesis.

A big thank you to all members of the Department of Microbiology, professors, students, technicians, administration staff, thank you Montreal and montrealais, I have learned so much from all of you!

And a special dedication to my family, to my strong father Jorge Armando Castillo, my beautiful mother Sonia Martinez Ayala, my creative brother David Castillo Martinez, my powerful sister Luisa Castillo Martinez, my brave grand-mother Virginia Ayala and to my love and best friend Debora Parrine Sant'Ana.

Chapter I

Review paper

This published review paper focuses on the action of different site-specific recombinases to solve dimerization problems during DNA replication and cellular division in bacteria. It initially addresses causes and effects of such events, two important aspects vaguely connected by most papers. With this review, we intended to offer a broader idea of the cause of such events, the reasoning for a certain outcome and the complexity to solve them. After reviewing all the mechanisms involved in crossing over events and dimer formation, we proceeded to explain the mechanisms involved in their repair and in the acquisition of new genomic material from viruses by presenting a detailed overview of all current known site-specific recombinase systems in bacteria.

I wrote the majority of this paper. George Szatmari and Amal Benmohamed provided valuable input, precise information and very helpful comments and corrections before and during the publication process.

This article was published in 2017 in *Frontiers in Microbiology* (Volume 8, article 453, pp. 1-18)

Xer site specific recombination: double and single recombinase systems

Fabio Castillo Martinez, Amal Benmohamed, George Szatmari*

Département de microbiologie, infectiologie et immunologie

Université de Montréal

Keywords: Site-specific recombination, tyrosine recombinases, single recombinases, XerS, XerH, XerA, IMEX.

Emails: fabio.andres.castillo.martinez@umontreal.ca, amal.ben.mohamed@umontreal.ca, george.szatmari@umontreal.ca

***Correspondence:** George Szatmari; e-mail: george.szatmari@umontreal.ca

Abstract

The separation and segregation of newly replicated bacterial chromosomes can be constrained by the formation of circular chromosome dimers caused by crossing over during homologous recombination events. In *E. coli* and most bacteria, dimers are resolved to monomers by site-specific recombination, a process performed by two Chromosomally Encoded tyrosine Recombinases (XerC and XerD). XerCD recombinases act at a 28 bp recombination site *dif*, which is located at the replication terminus region of the chromosome. The septal protein FtsK controls the initiation of the dimer resolution reaction, so that recombination occurs at the right time (immediately prior to cell division) and at the right place (cell division septum). XerCD and FtsK have been detected in nearly all sequenced eubacterial genomes including Proteobacteria, Archaea, and Firmicutes. However, in S*treptococci* and *Lactococci*, an alternative system has been found, composed of a single recombinase (XerS) genetically linked to an atypical 31 bp recombination site (*difSL*). A similar recombination system has also been found in ε-proteobacteria such as *Campylobacter* and *Helicobacter*, where a single recombinase (XerH) acts at a resolution site called *difH*. Most Archaea contain a recombinase called XerA that acts on a highly conserved 28 bp sequence *dif*, which appears to act independently of FtsK. Additionally, several mobile elements have been found to exploit the *dif*/Xer system to integrate their genomes into the host chromosome in *Vibrio cholerae, Neisseria gonorrhoeae* and *Enterobacter cloacae*. This review highlights the versatility of *dif*/Xer recombinase systems in prokaryotes and summarizes our current understanding of homologs of *dif*/Xer machineries.

1. Introduction

Bacteria and archaea have developed a variety of well-regulated and coordinated mechanisms of replication and segregation of their genomes that ensure the genetic material is transmitted faithfully to the daughter cells, despite the absence of temporal separation between DNA synthesis, chromosome separation and cell division [1]. However, the circular state of their chromosomes and plasmids constitutes a constant threat to genome stability and proper segregation because of dimer formation during recombinational exchanges between sister chromatids. These rearrangements can combine their genomes into larger molecules, compromising an equal distribution of the genetic material to the daughter cells [2–5]. This topological problem was fully addressed in 1981, when Austin et al. (1981) demonstrated that the stable inheritance of the prophage P1 was due to site-specific recombination (SSR), a specialized system that catalyzes DNA exchange between two defined DNA sequences, and which plays a major role in dimer resolution by converting multimeric forms to the monomeric forms. Later studies performed with the plasmid ColE1 connected SSR with plasmid monomerization and stability [7]. In 1990, the site-specific recombinase (XerC) was identified as the first protein responsible for SSR on *cer* [8], followed by the identification of a second recombinase, XerD required for this reaction [9].

2. The disadvantage of having circular DNA

In most bacteria and some archaea, replication begins at a single origin of replication *oriC* at which DnaA binds and stimulates the assembly of the replisome [10]. Replication forks then proceed bi-directionally until the two replication forks meet in an antipodal terminus region flanked by *ter* sequences. These sequences in conjunction with the replication terminator protein (Tus) stop the replication forks to synchronize their arrival at the same time and place [11–13]. However, chromosome replication is not a continuous process and is continuously halted by different types of DNA lesions such as UV irradiation, free radicals, genotoxic agents, DNA replication errors, transcription-replication conflicts, tightly bound protein-DNA complexes, or RNA secondary structures [12,14–17]. To maintain their genomic integrity, bacteria have developed several and

sophisticated mechanisms to minimize the frequency of these DNA lesions before the occurrence of replication. The initial barrier against deleterious DNA modifications is carried out by specialized mechanisms, each one required for a given type of lesion, such as proofreading, direct reversal of DNA damage, base excision repair, nucleotide excision repair and mismatch repair [18–21]. Additional groups of mechanisms are responsible for avoiding transcription-replication encounters, equally lethal for bacteria, such as the coordination of temporal and spatial gene activation and co-orientation, modulators of RNA polymerases (RNAPs) and replicative accessory helicases [16,22–24]. Nonetheless, it is unavoidable that some of this DNA damage or conflicts will escape the initial barrier and interfere with replication fork migration, leading to the eventual inactivation of the replication machinery and formation of double-strand breaks (DSBs), interstrand cross-links and single-stranded gaps (SSG). These represent critical forms of DNA damage that must be removed for chromosome replication and transcription to proceed [25,26]. Therefore, a second barrier of repair is called into play to cope with these "evasive" damages. This second barrier is preferentially carried out by the homologous recombination repair system (HR). Estimates indicate that HR repair is required in almost every cycle of replication [27–29]. In fact, the HR system is now not only considered as a functional mechanism for generating genetic diversity but also as a decisive factor in DNA repair, the latter being the primary role of this system in the maintenance of the genome and the main source of dimer events [30]. Thus, HR plays a central role in removal and/or repair of DNA damage and rescue and/or re-assembling of replication forks that have been broken or stalled [26,31]. In the traditional HR system in *E. coli*, its mode of action consists of a multistep process of breakage and rejoining of homologous sequences (one old and one newly synthesized DNA strand). It initially involves 1) recognition of the DNA lesion by the complexes RecBCD or RecFOR, depending on the type of DNA lesion; 2) formation of 3'-ssDNA overhangs processed by the exo and endonuclease activity of the Rec proteins, and subsequent coating by RecA; 3) strand invasion of the 3-terminal ssDNA into the homologous duplex DNA molecule and search of the complementary strand; 4) formation of a D-loop intermediate, transformation into a branched intermediate and Holliday junction (HJ) formation and 5) completion of the recombination process by resolution of the HJ, catalyzed by the systems RuvABC or RecG [32–34]. HJ resolution can result in two alternative products;

1. ‘crossover’ or spliced products; where reassortment of the flanking genes of the cleavage site has occurred, obtaining one different genotype at one side compared to the former DNA duplex, and therefore, long range of genetic exchange.

2. ‘non-crossover’ or patch products where the flanking regions were not exchanged, and instead, the resulting DNA duplex contain a ‘patch’ of hybrid DNA with a shorter range of genetic exchange [35,36].

Because circular chromosomes do not have “ends”, they are vulnerable to concatenation during formation of an odd number of crossover events. Thus, swapping DNA flanking regions tangles the sister chromatids and forms larger ring chromosomes that compromise cellular division **(Figure 1)** [2,4,37]. To ensure proper chromosomal segregation, bacteria and archaea have overcome these major threats by two broad mechanisms. One is to minimize the formation of crossing-over events, and the other is to solve dimer formation by performing an additional DNA exchange, immediately prior to cell division, at a specific region called *dif* (Deletion-Induced Filamentation).

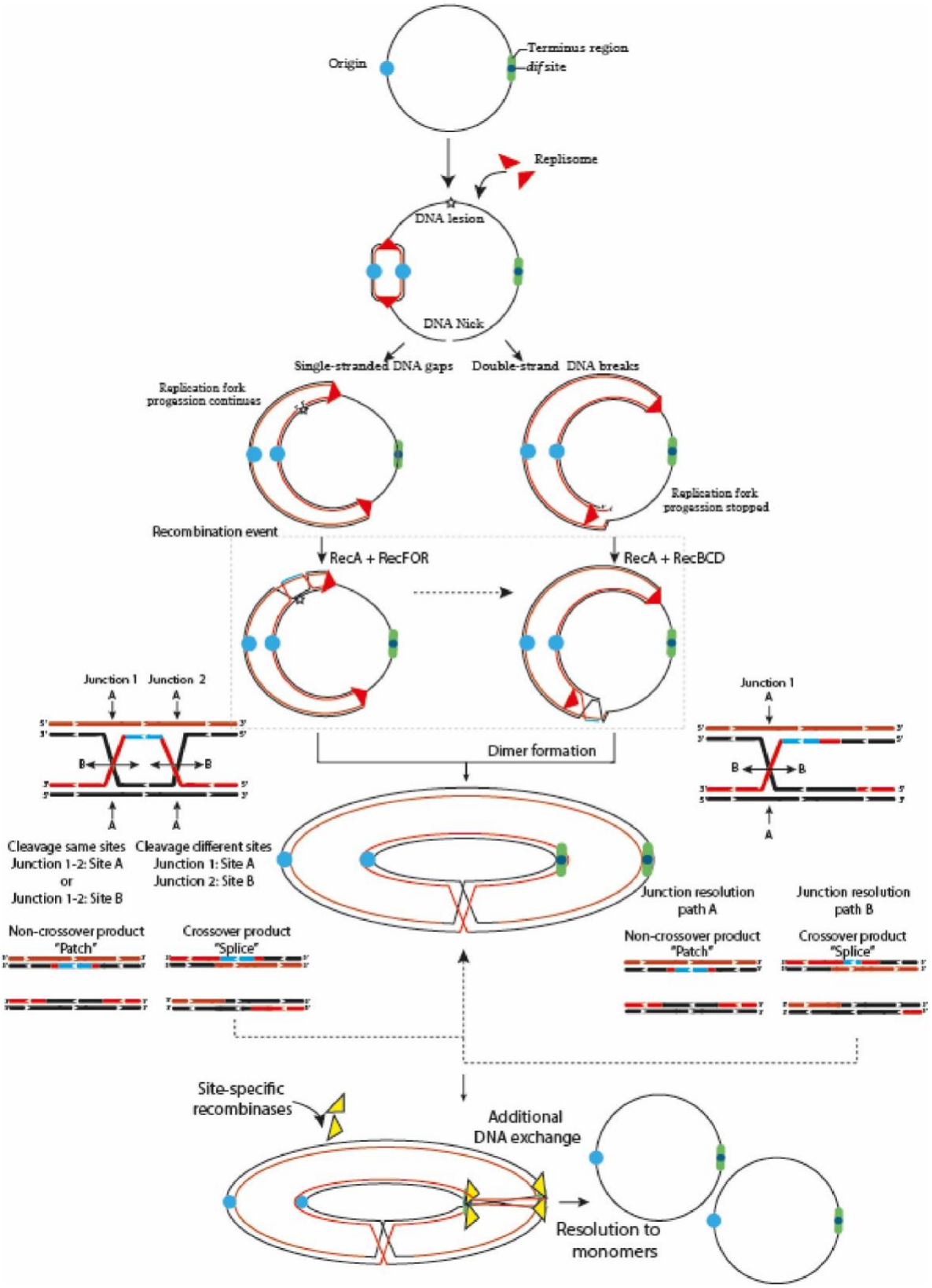


Figure 1. The two primary and most generalized pathways to solve chromosomal dimers generated by RecA-dependent repair or stalled replication forks.

If the fork encounters a non-coding lesion (oxidative damage, pyrimidine dimer or an abasic site) depicted with a yellow star, it usually generates SSG (left path), although it can also lead to DSBs ends during repair by the RecFOR system. The pathways can also diverge on the location of the DNA lesion. If the lesion is located on the lagging strand template, the replisome will be able to bypass the lesion by blocking 'Discontinuous DNA synthesis' and then resume it downstream of the lesion, leaving a gap that would be repaired by the RecFOR system. On the other hand, a lesion on the leading strand template might transiently stop the replisome, cause its dissociation and then, bind further downstream to a new leading-strand primer, although these mechanisms are still under debate [25,31]. Alternatively, when the fork encounters a nick in the template strand (unrepaired SSG) or some cases of replication fork collapse, a DSB is generated (Right path). The DSBs are processed by the RecBCD complex that catalyzes the reattachment of the damaged DNA to the sister DNA duplex, forming a D-loop structure and eventual recognition by the replication-restart PriA protein that directs replisome assembly and resumption of the replication process in an origin-independent manner. However, odd numbers of crossover events generate dimer products [38–40]. Thus, if the resolution of the HJ occurs in the same sites, it will generate monomeric chromosome (Non-crossover products). In contrast, if the resolution takes place in different sites, it will generate chromosome dimers (Crossover products) [15,41].

3. Avoiding dimer formation

One way to avoid dimer formation, as simple as it sounds, is to decrease the likelihood of dimer formation. This, however, is an intricate process of coordination and selection of the right enzymes at the right moment. Therefore, if the resolution of HJ intermediates by endonucleolytic cleavage can only result in crossover or non-crossover products, the likelihood of obtaining one or another is 50%. However, minimizing crossover events during homologous recombination repair seems to be the rule rather than the exception in organisms with circular chromosomes [42]. In *E. coli*, homologous recombination repair is processed by two predominant recombinational pathways; The RecBCD pathway associated with DSB repair, replication fork collapse, replication fork reversal and replication fork arrest, and the RecF pathway, which is mostly involved in the repair of SSG, and under certain conditions, can also repair DSBs [43,44]. Both mechanisms lead to the formation of HJs that are mostly resolved by the RuvABC complex in *E. coli* or RecU in Firmicutes and Mollicutes [41]. Deletion of the genes of the RuvABC complex eliminates non-crossover formation bias, supporting the idea that bias formation mostly depends on the action of Ruv proteins more than Rec proteins [35]. This idea was initially discussed by Van Gool et al. (1999) who demonstrated that crossover and non-crossover products are not random and, conversely, they are influenced by the positioning and orientation of the resolvasome on the HJ

intermediate, which in turn directs RuvC strand cleavage direction. Additionally, topological conditions such as DNA supercoiling, DNA catenation, adjacent HJ intermediates or the presence of double HJ intermediates can also influence assembly of the RuvABC complex on the HJ and indirectly affect resolution [45]. Subsequently, Cromie and Leach (2000) showed that RuvABC positioning may depend on the nature of the substrate caused by the type of DNA lesion, thus DSBs are more likely to result in crossover products (Frequently processed by RecBCD) whereas SSG are more likely to result in non-crossover products (frequently processed by RecF). Although, some fractions of SSG can cause DSBs [46]. However, the specific causes of crossover or non-crossover formation are still under debate, and different reactions cannot be completely dismissed. A clear example of this is the fact that RecBCD as well as RecF are not restricted to DSBs and SSG respectively and, on the contrary, the both may have interchangeable functions [47]. That would explain why some replication fork arrests generate non-crossover products even if they are mostly processed by the RecBCD pathway [48], or why RecF contributes almost equally to dimeric chromosome formation in *E. coli* despite the fact that it is responsible for SSG resolution [4].

Interestingly, the fact that DNA lesions and transcription-replication conflicts are more abundant in the leading strand than in the lagging strand in *E. coli*, and that these lesions usually generate non-crossover products, reinforces the idea that organisms with circular chromosomes favored a system that minimizes dimer formation during HR repair completion [16,35,49]. These biased reactions have also been detected in other microorganisms such as *B. subtilis* where the resolvase RecU biases homologous recombination towards non-crossover products [41]. Despite this non-crossover preference by HR system in *recA+* cells, dimer formation still occurs reaching 10 to 15% of the growing cells [50].

4. Coping with dimers

It is clear that dimer formation is regarded as a negative outcome that must be solved. Despite this, Mazin et al. (1996) proposed that under certain conditions of selective stress, plasmid dimerization could confer an advantage for the selection of adaptive mutations due to rapid accumulation and selection of plasmids carrying a specific mutation and subsequent segregation

to the daughter cells. Berza et al. (2013) also reported that plasmid dimerization greatly increased synthesis of a foreign protein and that plasmid content is unaffected by dimer formation showing some advantages for transcriptional events. However, these benefits were only considered for plasmids. Regarding circular chromosomes, dimerization must be resolved by the action of site-specific recombinases (SSRs), which are enzymes that are responsible for breaking and rejoining specific sites without requiring DNA synthesis or high energy cofactors [53]. The relevance of this system for proper chromosome segregation is supported by the high degree of conservation in Bacteria and Archaea. The Xer complex is considered one of the most conserved structural features in cells containing circular chromosomes, as well as RecA and FtsK enzymes [50,54,55]. The SSRs act on specific short DNA sequences, called recombination sites, where DNA exchange occurs in three different types of DNA rearrangements; deletion (divided into excision or resolution), insertion, or inversion. All these processes depend on the orientation and direction of the two recombination sites [56] **(Figure 1)**. All known site-specific recombinases are classified into two unrelated families, tyrosine–type or serine-type recombinases (Tyr or Ser) based on the amino acid residue that forms a covalent linkage between the protein and a phosphate at the DNA cleavage site [57]. Serine recombinases, often referred to as the resolvase/invertase family, act on a recombination site with just 2 bp separating the cleavage sites on top and bottom strands and the cleavages occur simultaneously to create a double strand break, while tyrosine recombinases, often referred to as the λ integrase family enzymes perform a two-step cleavage and rejoining process where the cleavage sites are separated by 6-11 bp. Each recombinase family possesses a distinct mechanism. Tyrosine recombinases are divided according to the recombination directionality; unidirectional or bidirectional recombinases. Whereas serine recombinases are divided according to their size; small or large recombinases [56,58].

The chromosome dimer resolution (CDR) process and heritable stability were originally elucidated in *E. coli* [8,9,59–61], where two paralogous site-specific tyrosine recombinases, XerC (298 aa) and XerD (298 aa) (Chromosomally Encoded Recombinases) were shown to act on a 28 bp DNA sequence (*dif* site), located in the *ter* region. The synaptic XerCD/*dif* complex consists of two XerC and two XerD subunits respectively bound to two *dif* sites **(Figure 2b)**. Limited structural information of some tyrosine recombinases have revealed a conserved catalytic domain fold [62],

facilitating the analysis of experimental data and allowing the development of a general model for Xer recombinases [63] consisting of; XerD [64], XerA [65], XerH [66] and other related tyrosine recombinases like Cre [67,68], HP1 integrase [69], FLP [70] and λ integrase [71,72]. The *E. coli dif* site is divided into two 11 bp half-sites that share partial dyad symmetry linked by a 6 bp central region that defines the positions of strand cleavage and exchange [73]. The initial step of site-specific recombination during dimer resolution requires the formation of a synaptic complex consisting of a tetrameric protein/DNA complex (four protomers of tyrosine recombinases and two recombination site duplexes). Once the synaptic complex is formed, two opposing and activated protomers cleave the DNA strand of each recombination site duplex. This occurs when the hydroxyl group of the nucleophilic tyrosine attacks the scissile phosphate in the central region to form a 3' phosphotyrosyl intermediate and a 5'-hydroxyl end. This intermediate conserves the energy from the phosphodiester bond cleavage to perform the first strand exchange. The recently formed 5'- hydroxyl attacks the 3' phosphotyrosyl linkage on the partner site to reseal the strand breaks creating a HJ intermediate [53]. HJ formation and isomerization activates the second pair of subunits bound to the other half of the recombination sites and inactivates the first pair of subunits. The second pair of subunits then cleaves, exchanges and rejoins the second pair of strands by the same mechanism just described; this second cleavage allows the resolution of HJ-intermediate and results in the recombinant DNA **(Figure 2B)** [63,74]. This process implies that the specific pairs of recombinases and/or active sites are continuously switched on and off to synchronize when and how recombination occurs, this coordination depends on allosteric interactions between the recombinases and external factors imposed on the synaptic complex [75–77]. In the XerCD/*dif* system, XerC normally initiates catalysis of one pair of DNA strands to form the HJ-intermediate without a subsequent resolution by XerD. Therefore, the HJs are rapidly converted back to the original DNA rearrangement. This XerC-first interaction is functionally active during the integration of certain bacteriophages that utilize Xer recombination to integrate their genomes into their host *dif* sites or in the resolution of plasmid multimers. In contrast, during chromosomal dimer resolution, pre-synapsed XerCD/*dif* complexes favor XerD activation by the FtsK protein to mediate the first strand exchange, generating a transient (XerD-HJ) intermediate,

subsequent isomerization forms a XerC-HJ intermediate that is rapidly resolved to recombinant DNA by XerC, see below and **Figure 2A, 3A** [54,78].

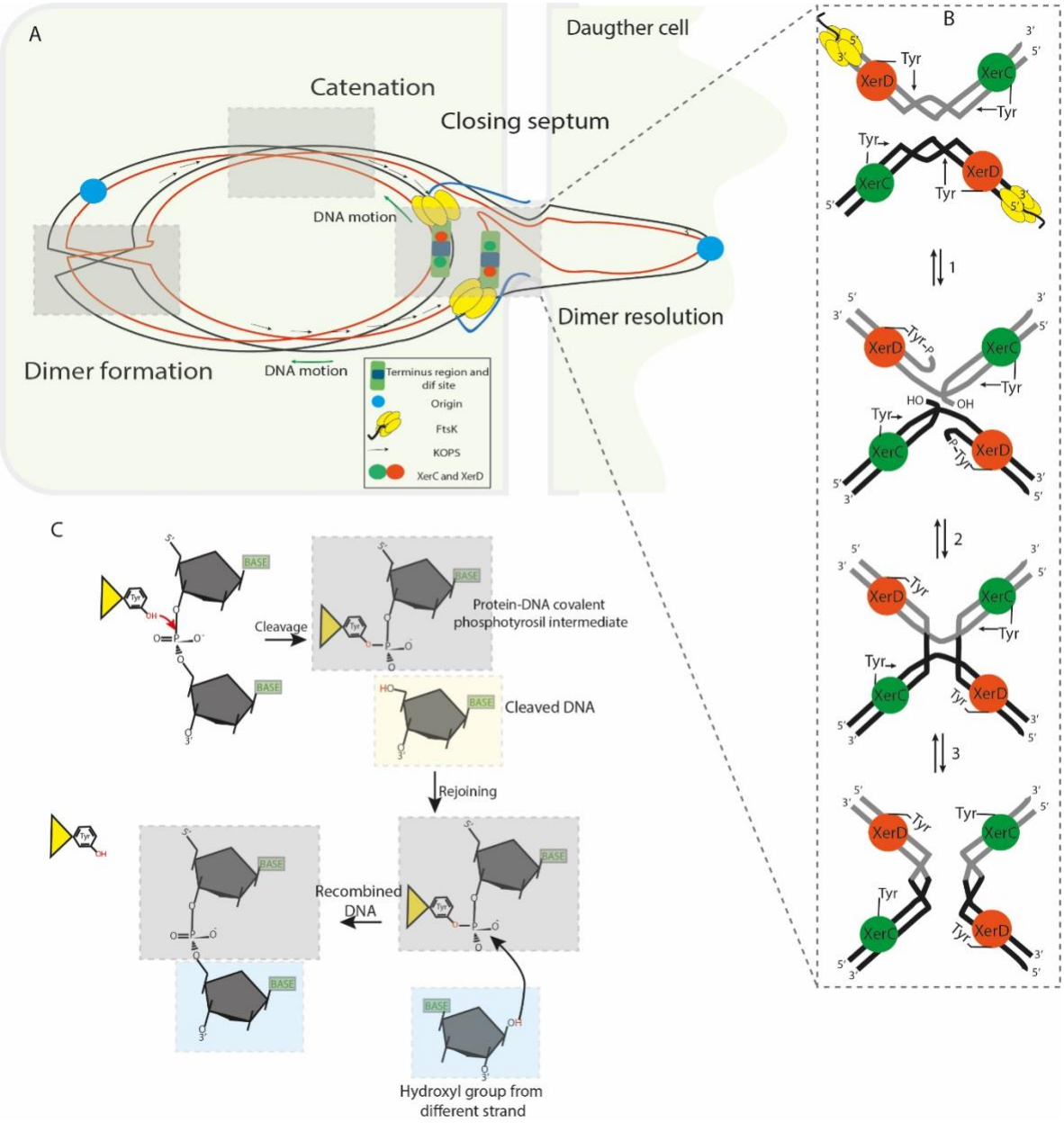


Figure 2. Segregation of the sister chromatids during chromosome dimer resolution in *E. coli.*

The illustration depicts the central part of a dividing cell in the final steps of chromosome segregation. The closing division septum, the motor domain αβ of $FtsK_C$ (yellow hexameric ring), the unstructured linker domain $FtsK_L$ (Blue ribbon), the KOPS sequences and the XerCD/*dif* synaptic complex are indicated. Concatenation prevents proper migration of the nascent chain of DNA; the origin regions move toward their respective cell poles, but the rest of the knotted DNA is stretched across and behind the septum. (B) $FtsK_C$ loads onto the KOPS sequences in an oriented manner and translocates towards XerCD/*dif* complexes. FtsK translocation allows it

to reach the XerCD/*dif* complexes and bring them into proximity; as a consequence, the γ-subdomain of the $FtsK_C$ region activates XerD (Orange sphere) to perform the first strand cleavage. Then, XerC (Green sphere) mediates the second strand cleavage, allowing separation of the sister chromatids from each other. (C) Illustration of the SSR mechanism used by tyrosine recombinases: The OH group of the active residue tyrosine attacks the scissile phosphate forming a 3'-covalent phosphotyrosyl enzyme–DNA covalent intermediate and a free 5'-hydroxyl end. The covalent intermediate is attacked in turn by the other 5'- end to reverse the cleavage reaction and obtain a recombinant product.

5. Establishing rules for dimer resolution

One of the fundamental questions about site-specific recombination at *dif* concerns how the system is controlled to ensure a proper chromosome dimer resolution (CDR) into monomers in the right place and at the right time without promoting the reverse reaction, which would generate dimers from monomers. It is understood that Xer-mediated recombination mostly depends on an active HR system because it is the major process that provides concatenated chromosomes. However, catenation problems caused by replication may require Xer-recombination system as well [3,79,80]. This reaction occurs at two polarized and specific regions of ~10kb at either side of *dif* called DAZ (dif Activation Zone), where oppositely oriented KOPS (FtsK Orienting Polar Sequences) converge and guide FtsK DNA translocation towards the *dif* locus [81]. This directional control is achieved by the interaction between the Xer recombinase system and the C-terminal domain of FtsK (Filamentous Temperature-Sensitive cell division protein K), a large division septum-associated DNA translocase, which coordinates chromosome segregation and cell division when chromosome organization has been affected (e.g. chromosome dimer formation, decatenation or delayed replication) [54,82–87]. FtsK was initially documented in 1995 due to observations in *E. coli* TOE44 (AB2497 *ftsK44*) mutant cells and their ability to form long chains of cells due to a single substitution of one amino acid in the N-terminal domain ($FtsK_N$), by then, FtsK was thought to participate in septum formation as a peptidoglycan-modifying enzyme. Then, Yu et al. (1998) demonstrated that inactivation of the C-terminal region of FtsK affected normal chromosome segregation due to the formation of long chains of cells and detected abnormal DNA distribution in some ftsK1::cat minB double mutants of *E. coli* minicells. Finally, in 1999 Steiner et al. (1999) discovered that site-specific recombination at *dif* requires $FtsK_C$ and thus, chromosome

dimer resolution only occurs in its presence**.** The ~1329aa FtsK protein can be divided into three domains; The ~279aa N-terminal domain ($FtsK_N$) is responsible for attachment of the protein to the membrane by four transmembrane segments and interaction with other proteins of the division septum such as FtsZ [90]. The linker domain $FtsK_L$, not commonly conserved in FtsK homologs, is primarily composed of proline-glutamine residues [91]. Its length and composition varies between species, being ~650 amino acids long in *E. coli* and most of Proteobacteria, ~200aa long in *Vibrios* [92] or ~125aa in *Pseudomonas* [93]. Experiments performed by Bigot et al. (2004) demonstrated that *ftsK*$_L$ mutations increased filamentation phenotypes even higher than *xer* mutants and that this filamentous formation did not correspond to problems in chromosome dimer resolution. On the contrary, it was thought to be due to a deficient positioning of the protein, reducing the possibility of contact between $FtsK_C$ and the DAZ region of the chromosome. Subsequently, Dubarry et al. (2010) revealed that different parts of the linker domain interact with other proteins of the divisome such as FtsQ, L, I and Z and these interactions help to stabilize the whole divisome at the site of septation. Interestingly, they also suggested that $FtsK_L$ domain may stop or slow down cell division during dimer resolution because of the destabilization of the divisome components when $FtsK_C$ has been pulled by the DNA during translocation, this force can separate FtsZ and delay septum constriction. There is also a proportional relation between the glutamine-proline concentration and its length, where the longest linkers are usually richer in these residues [95]. Following the linker, the highly conserved ~500aa C-terminal domain ($FtsK_C$), usually referred as the motor of the FtsK, is comprised of three separated subdomains called α, β and γ. Structural studies of the translocation module $FtsK_{\alpha\beta}$ of *P. aeruginosa* demonstrated that it assembles as a hexameric ring around double-stranded DNA forming a central channel of 30Å in diameter, where double stranded DNA (dsDNA) passes through [93,96]. Later structural studies of the orientation module $FtsK_\gamma$ of *P. aeruginosa* and *E. coli* demonstrated that six γ subdomains are loosely attached by a short linker of 10 aa to the hexameric ring $FtsK_{\alpha\beta}$ [84,97]. The $FtsK_{\alpha\beta}$ subdomains are responsible for the ATP hydrolysis-dependent DNA translocation of the protein. The 68aa $FtsK_\gamma$ subdomain is a helix-wing-helix domain that performs two main functions. The first role of this subdomain is to recognize the 8 bp KOPS sites and then directs FtsK translocation toward the *dif* site located within the *ter* region, at which, if concatenation occurs, two *dif* sites will be

brought together to form the synaptic complex XerCD/*dif* **(Figure 2A)**. KOPS are over-represented on the leading strand of replication where their concentration gradually increases as *dif* is reached; indeed, more than 90% of KOPS sequences nearby *dif* are located on the leading strand [82]; giving a possible estimate of 34 KOPS motifs located in the DAZ region [98] with a frequency of 1 motif every 13Kb [83]. The second main function of the $FtsK_{\gamma}$ subdomain is to activate the XerD catalytic function to generate the first HJ and subsequent dimer resolution [85,98,99]. How FtsK locates and assembles to initiate translocation in the correct KOPS sequence is still arguable, for this reason two models have been proposed; the loading model and the target search model, recently reviewed by Besprozvannaya & Burton (2014). New evidence strongly suggests that FtsK acts in a 350-kb region around *dif* that covers 7% of the genome where monomers of $FtsK_C$ assemble exclusively at KOPS motifs as described by the loading model. An initial interaction of a single monomer of $FtsK_{\gamma}$ will trigger a rapid and stepwise formation of the hexameric ring under high concentration of FtsK [100,101]. It is likely that FtsKγ assembles quickly and binds to KOPS as a trimer initially, with three FtsKγ modules interacting with consecutive GGG, NA, and GGG bases and then it hexamerizes gradually [84] Once a KOPS motif is detected, allosteric modifications occur leading to hexamerization of $FtsK_{\alpha\beta}$, which alters the angular conformation of $FtsK_{\gamma}$ on the DNA affecting KOPS recognition, and activates $FtsK_{\alpha\beta}$ ATP hydrolysis. As a consequence, FtsK is no longer able to recognize subsequent KOPS motifs during translocation, unless FtsK migration is impaired, and KOPS recognition is obligated to restart [84,97,98,102]. FtsK has been demonstrated to be the fastest known DNA translocase, reaching levels of 17.5 ± 3.5 kb/s at 37 °C or even faster with a striking stall force and a slight supercoiling induction, 1 positive supercoil per every 150 bp translocated [86,103,104]. It has also demonstrated a striking capacity to displace, evict or bypass different obstacles, especially proteins bound to the DNA such as RNA polymerases [98]. However, FtsK acts differently upon collision with RecBCD and XerD-XerCD/*dif* complex proteins. When FtsK collides with XerCD/*dif*, in a synapsed form, it activates XerD to create the XerD-HJ transient intermediate (structural rearrangements increase the distance between *dif* sites from about 53 to 67 A˚) [66,78] followed by a rapid dissociation from the DNA (dissociation takes to 0.5 to 1 seconds). Cleverly, May et al. (2015) demonstrated that recombination of the synaptic complex XerCD/*dif* takes 1 second longer than the FtsK dissociation time. Therefore, they suggested that

this time span can provide a regulatory control for dimer resolution because concatenated chromosomes will reform XerCD/*dif* synaptic complexes every time that resolution failed. Thus, multiple sets of FtsK hexamers colliding multiple times against XerCD/*dif* synaptic complexes will increase the likelihood of generating recombinant products. This regulatory mechanism ensures monomeric products are formed during translocation of impaired DNA.

6. Alternative dif/Xer resolution in prokaryotes

6.1 Plasmid resolution: Multicopy plasmid ColE1 and accessory proteins

Plasmid dimerization and eventual multimerization has been termed as the "Dimer catastrophe" due to its deleterious effect in cell populations [105,106]. Dimer catastrophe represents two major problems in bacteria; 1) unequal plasmid distribution among populations, in particular, multicopy plasmids that are more vulnerable to plasmid loss and 2) metabolic burden caused by the rapid accumulation of dimers into the host [106,107]. As mentioned previously, dimer resolution was originally elucidated in ColE1, resulting in the first functional characterization of XerC and subsequent identification of XerD by sequence homology to XerC [9]. These discoveries constituted a new approach for site-specific recombinases and their role in dimer resolution. Subsequent investigations led to the identification of SSR enzymes involved in dimer resolution in other plasmids of *E. coli*, and other bacteria. Current estimates have identified more than 1300 tyrosine recombinases where many of them are associated with other host proteins to regulate their activity, directionality, or processivity [63]. Large plasmids usually carry their own recombinase machineries adjacent to the recombination site. Whereas, small plasmids, like those in the ColE1 family, use the chromosomally encoded dimer resolution system of their host [108,109]. For ColE1 resolution, XerC/D proteins act on a specific site called *cer* **(Figure 3B)**, a non-codifying region of 280 bp where two additional proteins act with XerCD to catalyze SSR reactions: the arginine repressor (ArgR) (an arginine-dependent DNA binding protein originally called XerA) [59,110], and aminopeptidase A (PepA) (a bifunctional transcriptional regulatory protein that reacts to environmental signals, which was originally called XerB) [111]. SSR in *cer* is catalyzed by XerC within a sequence of 30 bp composed of two 11 bp half sides and a central region of 8 bp. XerC and XerD bind to the left and right halves cooperatively and respectively. Strand exchanges are catalyzed

by XerC to form a HJ intermediate that is eventually resolved by an uncharacterized cellular HJ resolvase to generate a recombinant product [112,113]. The *cer* site is comprised of a 30 bp core recombination site and two accessory DNA sequences of ~180 bp in length, in which one or two hexamers of PepA and one hexamer of ArgR control the reaction [114]. The two accessory proteins are necessary for recombination, since in their absence, plasmid dimer resolution cannot be completed. However, at abnormal high concentrations of PepA, recombination in vitro can proceed without the help of ArgR [115,116]. This is also seen at the recombination site *psi* of plasmid pSC101 which requires XerC, XerD and PepA but not ArgR (*psi* dimer resolution requires another accessory protein called ArcA instead of ArgR, and the cleavage reaction is performed by XerC and XerD) [113,117]. In *cer,* PepA and ArgR control recombination directionality so that dimers can only be converted into monomers and not the opposite reaction. Dimer resolution directionality caused by these two proteins involves the formation of two directly repeated *cer* sequences positioned in an antiparallel direction; this conformation is favored by negative supercoiling where the *cer* sequences are interwrapped three times around the proteins resulting in the formation of a right-handed synapse structure that brings the XerCD binding sites together. Sites in an inverted repeat position prevent right-handed formation; this ensures only dimer resolution occurs. Thus, XerC and XerD bind to the 30 bp *cer* synapse region and may interact with the N-terminal domains of the PepA hexamers. Whereas the ArgR protein, which is flanked by one or two PepA hexamers, might be involved in bending the DNA, tightening it and activating *cer* SSR by possible interaction with the C-terminal region of the Xer recombinases. Another possible function is to bring the two *cer* sites together and to allow PepA loading to form a nucleoprotein complex that promotes XerCD binding and recombination [114–116]. Additionally, *cer* also encodes for a 70 nt RNA fragment called Rcd that is only transcribed during dimer formation by the P_{cer} promoter and regulated in a sequence-specific manner by the FIS protein [118]. Rcd binds to the enzyme tryptophanase and induces a quiescent state by increasing indole production within the cell. The quiescent state permits the cell to arrest cell division and chromosomal replication but still be active metabolically. This process is thought to be part of a dimer formation checkpoint that allows the XerCD/*cer* system to resolve dimer formation during this pause [106].

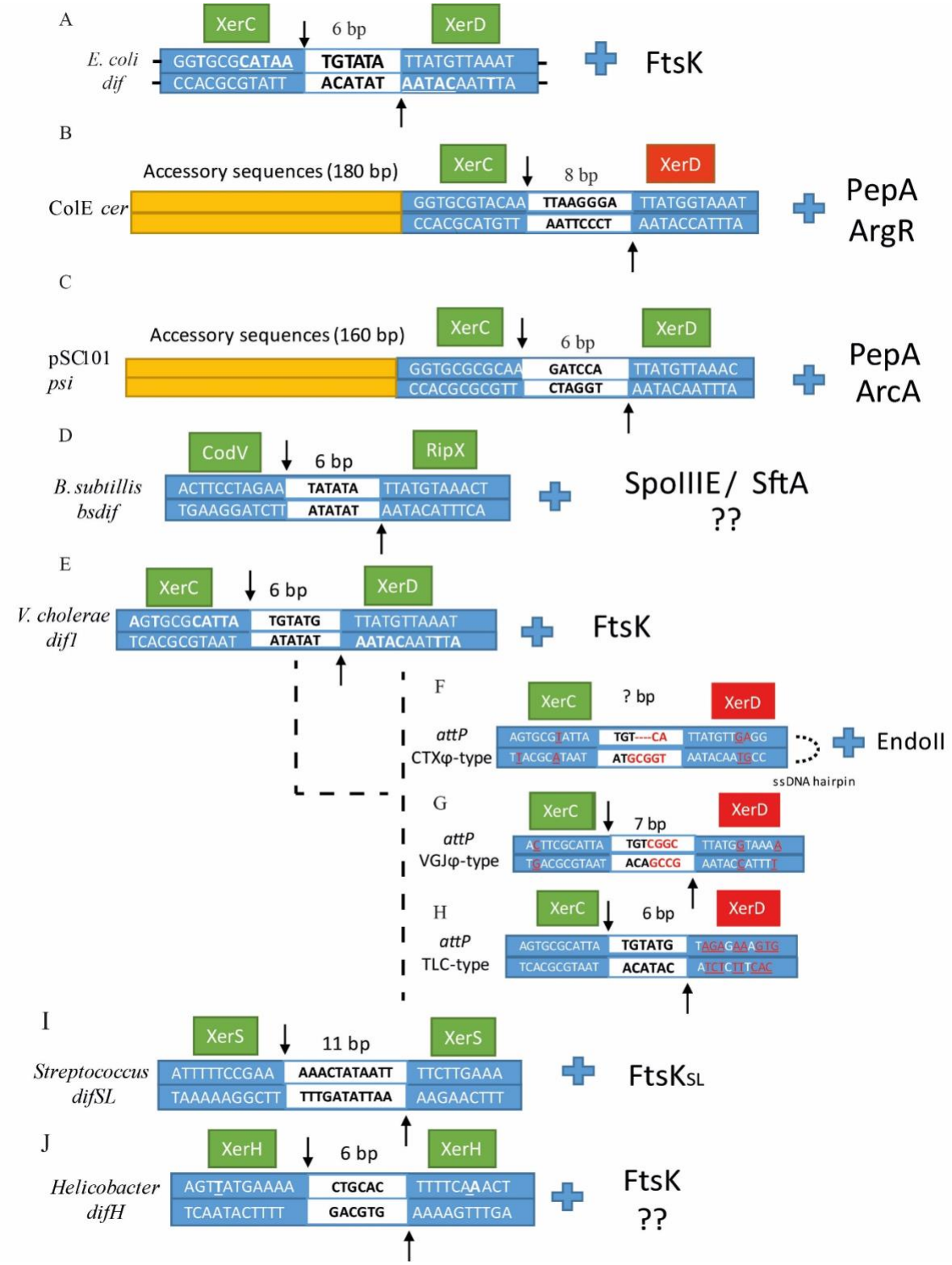


Figure 3. Sequence alignment of (A) *dif*, (B) *cer*, (C) *psi*, (D) *Bsdif,* (E) *dif1*, (F) *attP*^{CTX}, (G) *attP*^{VGJ}, (H)*attP*^{TLC}, (I) *difSL* and (J) *difH*.

Dyad bases in each arm are underlined in *dif*. XerC and XerD cleavage points are indicated by arrows, the central region is depicted in the middle of the sequence as a white box with the number of base pair corresponding to

each *dif* site above. Left and right arms are depicted as blue boxes with their corresponding sequences for each *dif* site. The catalytic unit is depicted as a green box, whereas the inactivated unit is depicted as a red box. For (E), (F), (G) and (H), bases that differ from *dif1* in *V. cholerae* are underlined and colored in red. Next to each *dif* alignment is the corresponding accessory protein that coordinates/activates dimer resolution.

6.2 The *Bacillus subtilis* model and the effect of two translocases

The capacity to perform site-specific recombination to resolve chromosome dimers is highly distributed among bacteria and archaea. Thus, homologs of XerC and XerD have been sequenced in a variety of species [108,119]. In *B. subtilis*, two homologs of XerC and XerD called CodV and RipX perform dimer resolution at a 28 bp *dif* site (*Bsdif*) close to the terminus region **(Figure 3D)**. The *Bsdif* region is comprised of two 11 bp half-sites with imperfect dyad symmetry where CodV and RipX bind simultaneously and a 6 bp central region where DNA exchange occurs. Both CodV and RipX share a 37 and 44% identity with the XerC and XerD respectively, and 39% between them [120]. CodV binds preferentially to the left half-site and preferentially cleaves the top strand whereas RipX is able to bind to both sides with preferential binding to the right-half-site and preferential cleavage of the bottom strand. Cleavage by CodV is more efficient than cleavage by RipX, which suggests that CodV performs the first strand cleavage followed by RipX in in vitro experiments [121]. Sciochetti et al. (1999) also demonstrated that RipX could interact effectively with the *E. coli dif* site, unlike CodV which showed a weaker interaction with this substrate. However, addition of XerC to RipX/$dif_{E.\ coli}$ or XerD to CodV/$dif_{E.coli}$ generated larger complex formation in gel retardation analysis, demonstrating protein-protein interactions between these four proteins, which confirms some conserved features of tyrosine recombinases among bacteria. This is supported by the fact that the right half-site presents highly conserved features with respect to other *dif* sites among some bacteria, whereas the left-half site is less conserved, which could explain why RipX can bind $dif_{E.coli}$ [121]. In contrast to *E. coli*, the synaptic complex can be brought together by the action of two DNA translocases: the membrane-associated SpoIIIE protein (Stage III Sporulation Protein E) and the soluble SftA protein (Septum-associated FtsK-like Translocase of DNA). Both translocases harbor AAA+ATPase and C-terminal domains with 56% of sequence similarity between them. SftA exhibits 50% identity with respect to the *E. coli* $FtsK_\gamma$ domain whereas SpoIIIE exhibits a 50% of similarity to the $FtsK_{\alpha\beta}$ subdomain and 42% of similarity to the

$FtsK_\gamma$ subdomain of *E. coli* [122]. The N-terminal domains of these proteins are more divergent; SpoIIIE and FtsK share 36% of identity to respect to the four transmembrane helix whereas SftA lacks the transmembrane spanning domain [123,124]. N-terminal domain variations coincide with their different location and activation in the genome. FtsK and SpoIIIE share a similar mechanism to anchor to the inner membrane of the dividing cell by their N-terminal regions; this transmembrane interaction is possibly reinforced by interactions with FtsZ or other cell division components such as FtsA or ZapA [91]. During vegetative growth, SpoIIIE shows two predominant states: a static phase, where SpoIIIE is assembled close to future sites of cellular septation, and a mobile phase, where SpoIIIE does not occupy a specific position. Once cellular division begins, the static phase takes place when SpoIIIE is recruited by FtsZ and other division machinery proteins and is escorted to the center of the division septum [125]. SpoIIIE remains in the invaginating septum and hexamerizes independently of the cell division stages (Vegetative, division and sporulation stages) and independently of DNA interaction [126], suggesting that SpoIIIE assembly may not be restricted to the presence of impaired DNA and on the contrary, may be involved in normal DNA segregation as demonstrated by the following paper, [125]. Experiments using high-resolution microscopy revealed that under formation of asymmetric (sporulation) or symmetric (vegetative growth) septa, the SpoIIIE concentration increased 2.5-fold around the constricting septa, even without evident formation of the septa, indicating close interaction with other components of the division machinery, that in turn regulates its activity under specific conditions [125]. SftA can be localized either to the cell center or more frequently, to the forming division septum. Although SftA lacks an integral membrane domain, the FtsZ ring recruits the enzyme and attaches it to the division septum during the initiation of cellular division, which explains its localization through the cell cycle [123,127,128]. These patterns of localization suggest that both translocases (SftA and SpoIIIE) are present at the septum at various times of segregation and that they perform DNA migration independently of each other, although SftA is only involved in DNA cytokinesis in contrast to SpoIIIE that may be involved in cytokinesis and cell division processes [127]. DNA translocation is initially carried out by SftA during septation, probably by recognition of the 8-nucleotide SRS motifs (SpoIIIE Recognition Sequences) which are similar to the *E. coli* KOPS sequences. The SRS motifs are mostly located on the leading strand (up to 85%), and direct

translocation towards the *Bsdif* site [83]. Therefore, the primary function of SftA consists of moving chromosomal DNA until the *ter* regions are positioned at midcell and the origin regions migrate to each pole of the cells. The SftA may also be required for other proteins involved in cytokinesis and FtsZ positioning [129]. SpoIIIE, the second translocase in *B. subtilis*, may take over DNA translocation working synergistically but not interchangeably with SftA, it can also function as a DNA segregation checkpoint preventing membrane fusion until chromosome segregation is completed [123,125]. Cattoni et al. (2014) suggested that SpoIIIE binds non-specifically to the DNA in a pre-formed hexameric open ring conformation and then searches for SRS motifs without hydrolysis of ATP. Similarly to FtsK proteins, SRS recognition by the $SpoIIIE_{\gamma}$ domain triggers allosteric modifications that activate the ATPase activity of $SpoIIIE_{\alpha\beta}$ and therefore, DNA translocation [130]. Once it encounters SRS motifs, the hexameric ring changes to the closed and active form pumping the chromosome in an oriented manner by recognizing further SRS motifs and translocating it towards *Bsdif* [126]. As mentioned before, SpoIIIE is actively expressed in all growing cells and is essential during sporulation to translocate the remaining DNA from the mother cell into the forespore compartment, and during vegetative growth to guarantee that concatenate formation or disrupted genomes will not affect normal cellular division. Moreover, SpoIIIE is also required for septal membrane fusion after completion of chromosome translocation. During sporulation, asymmetric septation encloses the DNA and traps 25–30% of one chromosome into the forespore. SpoIIIE pumps the remaining 70-75% by an analogous mechanism used by FtsK; the reaction only takes 20 minutes demonstrating its incredible speed [96,131]. The mechanism of SpoIIIE DNA translocation through the membrane is still unclear, since recent single molecule-imaging experiments still provide valid information for two main models; the paired DNA conducting channel model [132,133] and the aqueous channel model [125]. This system of both translocases has been also detected in *Staphylococcus aureus* termed FtsK and SpoIIIE because of their amino acid homology to SpoIIIE and FtsK from *B. subtilis* and *E. coli* respectively. However, in contrast to *B. subtilis* system, in *S. aureus* both enzymes seem to present a redundant, although independent role in DNA segregation. Individual deletions of either FtsK or SpoIIIE did not exhibit major changes in chromosome segregation for *S. aureus,* however when combined together they represented a major threat for *S. aureus* genome stability [134].

Consistent with their different roles, SftA and SpoIIIE do not colocalize during vegetative-replicative stages or sporulation. Thus, SftA in concert with FtsZ and division proteins moves chromosomal DNA away from the closing division septum. Then, upon septum closure, entrapped DNA is translocated through the SpoIIIE pore or channels into the correct compartment (either a forespore or a daughter cell). However, unlike FtsK that activates XerCD recombination reactions, neither SftA nor SpoIIIE directly activate CodV or RipX recombinases. In this case, SftA and SpoIIIE affect the CodV/RipX reaction by proper positioning of the *ter* region, but there is no evidence of direct interaction between these enzymes to date [123,129].

6.3 Multichromosome bacteria and IMEX

Vibrio cholerae, as well as 10% of sequenced bacteria to date, possess a very distinct property among bacteria; it harbors more than one chromosome [135]. One ancestral chromosome I (chrI) of 2.96 Mbp and one plasmid-derived chromosome II (chrII) or 'chromid' of 1.072 Mbp, encode 2,775 and 1,115 ORFs, respectively. ChrI contains most of the housekeeping genes whereas chrII contains essential genes specialized in adaptation to new environments or pathogenicity [136–139]. Harboring two or more chromosomes have shown to be highly heritable among these bacteria, which suggests that multiple chromosomes offer a positive selective pressure to maintain them. One possible explanation is that multiple chromosomes might offer an advantageous feature against dimer formation. Val et al. (2008) showed that dimer formation increases exponentially in relation to the size of the replicons, thus, dividing a single replicon into two or more replicons may reduce this topological problem. However, genome size might not be relevant for the presence or absence of Xer/*dif* recombination machinery. Some large chromosomes do not require Xer/*dif* recombination machinery as in some *Legionellales* (genome size ranging from 2 Mb to 5 Mb) whereas some small-sized chromosomes still require Xer/*dif* recombination machinery as demonstrated by some *Rickettsiales* (genome ranging from 0.85 to 1.52 Mb in size) [50].

Homologs of XerC/XerD and FtsK have been characterized on chrI, referred as $XerC_{VC}$ and $XerD_{VC}$ with 53% and 68% of amino acid similarity to *E. coli* XerC and XerD, respectively [140,141]. Whereas chrII does not encode any Xer recombinase involved in dimer resolution. *dif*-like sequences are present in both chromosomes (*dif1* & *dif2*) located near GC skew shift-points [55,92]. Interestingly,

both *dif* sites differ from each other in their sequences, *dif2* harbors five different nucleotides compared to *dif1* and most α-proteobacterial *dif* sites, four of them in the central region, resembling *dif*-like plasmid composition [55]. Dimer resolution in *V. cholerae* requires $FtsK_{VC}$ translocation by recognition of KOPS-like motifs (GGGNAGGG) in a similar way to that found in *E. coli*. Once the *dif* sites are brought together nearby, $FtsK_{VC}$ activates $XerD_{VC}$, which is positioned to cleave the bottom strand, and perform the first strand cleavage. Then $XerC_{VC}$ cleaves the top strand and performs the second strand cleavage; these reactions are carried out on both chromosomes at their respective *dif* sites **(Figure 3E)** [92]. Additional studies demonstrated that *E. coli* FtsK was able to activate 50% of the $XerCD_{VC}$ synaptic complexes at *dif1* whereas only 20% of $XerCD_{VC}$ were activated at *dif2*, suggesting that the *dif2* recombination process requires more accurate interactions between the FtsK proteins and the XerCD complex [92]. An additional feature of multiple chromosomes is their capacity to synchronize replication termination at the same time despite their different sizes [136]. This capacity may confer an additional regulatory control against dimer formation due to the time-lapse between the replicated chromosomes and cellular division. Demarre et al. (2014) showed that *terII* sites (chrII) separate earlier than *terI* and that this early separation keeps *terII* sites at midcell by the macro domain MatP/*matS* organization system. This restriction during concatenation induces several collisions at midcell between *terII* sites, increasing the number of recombinational events and the likelihood of dimer resolution. It also ensures that *ter* sites of bacterial chromosomes remain exclusively in mid-cell to be processed by FtsK.

Although XerC and XerD recombinases normally perform dimer resolution, they are also exploited by other replicons such as plasmids, bacteriophages, and other integrative elements. Indeed, initial studies on plasmid stability in ColE1 and phage integration of bacteriophage λ led to the discovery of XerC and the mechanistic insights of the tyrosine family [63]. In *V. cholerae*, the causative agent of the potentially fatal human disease cholera, $XerC_{VC}$ and $XerD_{VC}$ are hijacked by some vibriophages to integrate their genomes into the chromosome. They are usually referred to as IMEX (Integrative Mobile Elements Exploiting Xer), and the best-known ones are VGJφ (Vibrio Guillermo Javier filamentous phage), TLCφ (Toxic Linked Cryptic), and CTXφ (Cholera Toxin Phage). CTXφ is a lysogenic [(+)ssDNA] filamentous bacteriophage that encodes the A-B type

enterotoxin CT in *V. cholerae* [143]. These three vibriophages harbor a particular attachment site (*attP*), a *dif*-like site that serves to classify the three different groups of IMEX, (CTXφ-type, VGJφ-type & TLCφ-type) [144]. Although the components to integrate their genomes are very similar, their mechanisms of integration differ from one to the other and from their host strains. Direct ssDNA integration by CTXφ-type phages is characterized by the formation of a ~150bp folded structure created by the intra-strand base pairing interaction between two palindromic *attP* sites (*attP1* & *attP2*) separated by 90 nt on the ssDNA sequence **(Figure 3F)** [143]. The two overlap regions *attP1* and *attP2* reassemble the $XerC_{VC}$ side of *dif1* and *dif2* regions but differ from the XerD-side. This lack of homology between $XerD_{VC}$ recognition site and $attP_{CTX\phi}$ limits the catalytic reaction to $XerC_{VC}$ that catalyzes the complete reaction. An additional host factor called EndoIII participates in the directionality of the reaction, which blocks further rounds of strand cleavage by $XerC_{VC}$ causing its dissociation and therefore preventing CTXφ excision [145]. Although $XerD_{VC}$ is not involved in the catalytic reaction, it is still necessary for a successful integration, probably by its role in synaptic complex formation [146]. Once the integration is completed, host DNA replication proteins resolve the formed HJ intermediate and convert it to dsDNA. Prophage CTXφ cannot be excised from its host since it loses the capacity to fold itself, which in turn prevents further base-pairing interactions between the *attP* sites, which ultimately abolishes the $XerC_{VC}$ catalytic reaction [143]. Interestingly, CTXφ integration in El Tor strains is only found in chrI, and it is generally associated with two other vibriophages, TLCφ and RS1 that enable CTXφ integration in *V. cholerae* genome by reconstituting a functional *dif* site, and by promoting CTXφ replication and transmission [147]. In the classical biotype strains, CTXφ usually targets both chromosomes [148].

Similarly, to CTXφ, VGJφ integration uses the $XerC_{VC}$ catalytic reaction at the *dif1* site, but unlike CTXφ, it only harbors one *dif*-like attachment site ($attP_{VGJ\phi}$) of 29 bp that allows its integration into the chromosome as a dsDNA. The *attP* central region contains four different nucleotides close to the XerD binding side with respect to the central region of the *dif1* site **(Figure 3G)**. The lack of homology at the $XerD_{VC}$ central region side prevents $XerD_{VC}$ participation in the catalytic reaction. Once integrated, prophage VGJφ acquires two *attP* sites (*attPL* and *attPR*), equally functional for the $XerC_{VC}$ excision reaction, in contrast to CTXφ, where Xer recombinases can process VGJφ excision from the host genome [144]. TLCφ also depends on host encoded Xer

recombinases for its integration. Its $attP_{TLC\phi}$ site possesses high homology with the $XerC_{VC}$ binding side and central region of *dif1* whereas it is highly divergent from the $XerD_{VC}$ binding site **(Figure 3H)**. The prophage form of TLCφ is almost always linked to CTXφ integration confirming the regular synergistic interactions found in most IMEX. Paradoxically, despite the lack of homology between the XerD binding sites of *dif1* and $attP_{TLC\phi}$, TLCφ integration/excision is mediated by $XerD_{VC}$ and then completed by $XerC_{VC}$ resembling dimer resolution in bacteria, but independently of FtsK participation [79].

IMEX are recombination platforms that permit bacteria to evolve and adapt through the acquisition and reordering of relevant genes. They have strengthened bacterial evolution, playing an important role in the rise of multidrug resistance, gene transfer mechanisms and virulence factors among clinically relevant bacteria [149,150]. Besides the vibriophages just described above, some other relevant IMEX have been found; the gonococcal genomic island (GGI) related to pathogenic *Neisseria* species [151] and the EludIMEX-1 found in *Enterobacter ludwigii* [152]. GGI is an unusually long IMEX (57 kb long) found in almost 80% of *N. gonorrheae* strains and is involved in the expression of type IV secretion system (T4SS) genes [153]. GGI carries a degenerate *dif* site called dif_{GGI} of 28 bp with a XerC-binding site and a central region homologous to the conserved *Neisseria dif* site (dif_{Ng}) and a divergent XerD-binding site. GGI insertion into the *Neisseria* genome follows a chromosome dimer resolution-like process where FtsK activates XerD to perform the first strand cleavage between dif_{Ng} and dif_{GGI} followed by isomerisation of the synaptic complex and activation of XerC to perform the second strand cleavage, creating a GGI integrated form with two active Xer sites. Interestingly, the GGI synapse has given important clues about how IMEX might remain integrated in the host genome despite the presence of *dif* sites. Fournes et al. (2016) revealed by experiments in vitro that a trimeric form of the *E. coli* FtsK protein (t-FtsK$\alpha\beta\gamma_{Ec}$) was unable to activate XerCD recombination at one of the two *dif* sites (the dif_{GGI} site), in fact, the XerCD/dif_{GGI} complex was unable to stop t-FtsK$\alpha\beta\gamma_{EC}$ translocation. As a consequence, the XerCD complex is dissociated from dif_{GGI} and the excision process is inhibited.

EludIMEX-1 is a 29.1-kb IMEX found in *E. ludwigii* (ECAA-01) that carries the $bla_{NMC\text{-}A}$ gene that encodes for a serine carbapenemase. It was first characterized by Antonelli et al. (2015) when they sequenced the whole genome of a NMC-A-positive isolate of *E. ludwigii*. The results indicated

the presence of a new 29-kb region with lower GC content when compared to the bacterial genome, indicating a possible gene transfer acquisition [154]. Further analysis revealed that this region is flanked by putative XerC/XerD recombination sites with high homology at the XerC-binding site. They also determined that EludIMEX-1 insertion site in the genome was the same for two distinct species of the *E. cloacae* complex suggesting a possible acquisition via a XerC/XerD dependent recombination event at a specific *dif*-like site [152]. Understanding of IMEX control and excision processes will provide us a better idea of how counteract the acquisition of antibiotic resistance genes in pathogenic microorganisms.

6.4 The *difSL*/XerS model

The *E. coli* pathway of dimer resolution has been found to be highly conserved among bacteria with circular chromosomes. It was initially demonstrated by Recchia and Sherratt in 1999 when they analyzed 16 eubacterial and five archaeal genomes for XerCD-CodV/RipX homologs. They showed that most eubacterial genomes possess two putative Xer recombinases whereas Archaea presented a single recombinase in three of the five genomes analyzed [155]. Subsequently, Carnoy and Roten in 2009 demonstrated by doing an exhaustive computational analysis of 234 chromosomes from 156 proteobacterial species, that 87.8% of the genomes analyzed presented XerCD-like and *dif*-related sequences [50]. Moreover, Kono et al. (2011) predicted by a recursive hidden Markov model method (including XerCD orthologues) that 578 out of 592 bacterial genomes with a single chromosome and 63 out of 66 genomes with multiple chromosomes presented a *dif*-like sequence. Additionally, they remarked how XerC and XerD are conserved in almost 60-70% of bacterial species, and 85% in proteobacterial species [156]. These results among many others led to the general view that the *E. coli* pathway is predominant for dimer resolution. However, dimer resolution machinery or regulation of strand exchange may differ: some processes may require or disregard accessory proteins, others may or may not require activation by translocases, some will be mediated by a XerC-first strand exchange whereas others by XerD-first strand exchange and others may need two recombinases or only one. Among these divergences and unique characteristics for each bacteria to solve dimer formation, the less studied ones are the unconventional single recombinases.

Recchia and Sherratt first mentioned the presence of single recombinases in (1999) from the identification of two eubacterial genomes harboring only one Xer homologue. It was later confirmed when Le Bourgeois et al. (2007) demonstrated that some species of *Lactococcus* and *Streptococcus* use an alternative Xer recombination machinery. This new Xer complex is based on a single tyrosine recombinase called XerS (356 aa) that acts on an atypical 31 bp recombination site called *difSL* in the presence of dimers. Unlike *E. coli*, the XerS gene is found immediately adjacent to the recombination site *difSL* acting as a single module. The *difSL* site differs from most *dif* sites because of its large central region of 11 bp as opposed to the normally found 6-8 bp in all other *dif* regions **(Figure 3I)** [158]. Thus, *difSL* consists of two imperfect inverted repeat sites of different sizes separated by the central region where DNA exchange occurs. The inverted repeat region is one nucleotide longer in *difSL* and contains an extra nucleotide in the middle of the right inverted repeat (TTTTCTTGAAA) versus the left part of the sequence (TTTCCGAAAA). This additional spacing suggests XerS/*difSL* may be biased to favor binding in one-half site over the other. It was later confirmed by Leroux et al (2011), where they also showed that XerS presented stronger interaction with the left-half site of *difSL* than the right-half site, and a preference for initiating the recombination reaction on the bottom strand of the *difSL* site. These results indicate that, although the *difSL* site is relatively symmetric and XerS is a single tyrosine recombinase, there is a bias for where the proteins initially bind to *difSL* and where they initiate the strand cleavage reaction. Thus, the left-bound monomer could activate the right-bound monomer by bending the DNA or changing the conformation of the second monomer which could explain the preferential cleavage and exchange of the bottom strand. This behavior resembles what XerC displays with weak binding but stronger strand exchange when compared to XerD [158,159]. This intrinsic bias alone cannot control the preference of the directionality of the strand cleavage reaction. The achievement of proper control requires the action of a SpoIIIE-like homolog translocase called $FtsK^{SL}$, a protein of 758 aa in length in *Streptococcus mutans* or 816aa in *S. agalactiae* with low similarity at the N-terminal region between them. This low similarity does not affect its binding preference to the division septum commonly found in most proteins of the FtsK-HerA superfamily [157]. The C-terminal domain of $FtsK^{SL}$ shows 41% similarity at the amino acid level in relation to $FtsK_{E.\ coli}$ with four of the five amino acids similar (QR-GN motif) involved in

XerD interaction [99]. On the other hand, $FtsK^{SL}$ is unable to read *E. coli* KOPS motifs as demonstrated by Nolivos et al. (2012), probably due to the lack of common skewed octamers sequences called Architecture Imparting Sequences (AIMS) in *Firmicutes*, which means that KOPS sequences in *Firmicutes* are not as conserved as in proteobacteria [160]. This would also explain the divergence between $FtsK_\gamma$ domains even among *Firmicutes.* Additionally, AIMS found in *L. lactis* differ in both in length and in sequence from traditional KOPS/SRS motifs, being A-rich heptamer motifs instead of the GC-rich octamer motifs [82]. XerS also lacks critical residues found in XerD to interact with FtsK (residues RQ-QQ). Interestingly, XerS/*difSL* recombination occurs almost in a similar fashion to that of *E. coli*. Both Xer systems require $FtsK_N$ localization at the division septum and $FtsK_C$ translocation to achieved Xer dimer resolution. Additionally, XerS/*difSL* proved to be functional in *E. coli*, despite the lack of homology in their FtsK proteins. Further analyses on FtsK-Xer interactions are required since the exact mode of action is still speculative [159].

6.5 The *Helicobacter* and *Campylobacter* (*difH*/XerH) model

Studies in *Helicobacter* sp and *Campylobacter* sp led to the discovery of another type of single recombinase called XerH that acts on a recombination site called *difH* in a FtsK-dependent manner. It was shown to be involved in chromosome segregation and possibly dimer resolution in *Helicobacter pylori* [156]. XerH (354 aa to 362 aa) differs from the traditional XerCD (298 aa) recombinases by its size and protein homology (26% of identity with respect to XerCD). It also shows more similarity to XerS (356 aa) in both the size of the protein and the high degree of homology of their recombination sites [50,161]. Another characteristic of XerH and a possible hallmark of single recombinases (XerS and XerH) is that the *difH* sequence is also located near the recombinase-encoding gene, indicating a possible individual genetic module for Xer expression [50,157]. Interestingly, most of the epsilon species of ε-proteobacteria harbor a XerH/*difH* system whereas some other ε-proteobacteria (*Sulfurimonas denitrificans* and *Sulfurovum*) possess a system analogous to the classical XerCD system. Additionally, unlike other tyrosine Xer recombinases, XerH activity appears to be affected by a second Xer recombinase called XerT in *H. pylori* (the TnPZ transposon associated recombinase) since under XerT deletion, *difH* recombination levels increased [156]. Recent structural studies showed that the *Helicobacter difH* comprises two highly conserved imperfect inverted binding sites of 11 and 10 bp (AGT**T**ATGAAAA

and AAAAGTTTGA) in the left and right sides respectively, separated by a 6-**10?** bp central region **(Figure 3J)** [66] (Unpublished data suggest a 10 bp central region, [161]). Two subunits of XerH bind cooperatively to each side with a stronger binding affinity as well as cleavage reaction efficiency in the left half site than the right half site (the outer region in *dif* appears to be determinant in the order of binding and cleavage reactions). The left half site preference is due to stronger interaction between XerH and *difH* left site ΔG = -21.3 kcal/mol compared to the right half site ΔG = -15.4 kcal/mol). The extra nucleotide thymine (T4) in the outer region of the left half site confers a specific hydrogen bond between the left arm and the lysine (K290) of XerH that favors stronger protein-DNA interaction with the other three outermost nucleotides, DNA bending and specific positioning of the nucleophilic tyrosine. Surprisingly, XerH assembly on *difH* does not induce strong DNA bending alone and it seems to require FtsK to generate the required conformational rearrangements to favor XerH DNA exchange [66]. Results obtained by Leroux et al. (2013) in *difH* of *C. jenuni* (*difH*$_{camp}$) demonstrated that XerH binding to either the left or right site of *difH*$_{camp}$ resulted in similar affinities compared to the full *difH*$_{camp}$ site possibly due to the similarity between the outer sequences of both arms. Additionally, XerH binding to *difH*$_{camp}$ appears to be less efficient than XerS which suggests that it is less cooperative than XerS/*difSL* system. Additionally, these results contradict XerH binding affinities observed in *H. pylori* by Bebel et al. (2016) and in most tyrosine recombinases involved in chromosome resolution since it did not show any binding preference. On the other hand, unlike binding activity, asymmetrical cleavage reactions by XerH were found with a higher efficiency for bottom-strand substrates than top strand, in agreement with the results of Bebel et al. XerH recombination was also observed in vivo between two *difH*$_{camp}$ sites located on the same plasmid; it is also suggested that XerH might be involved in decatenation processes because of the apparent absence of Topo IV proteins in *H. pylori* [156,161].Interestingly, despite *difH*$_{camp}$ and *difSL* similarities in the recombination sites, the recombinases do not cross-react (XerH does not bind *difSL* and XerS does not bind *difH*$_{camp}$ sites) (M Leroux, unpublished).

6.6 The Archaea *dif*/XerA model

In archaea, chromosome resolution appears to be catalyzed by a single recombinase (XerA) in a FtsK-independent manner that acts on a *dif*-like site located in the replication terminus region

[162,163]. XerA shares a conserved C-terminal domain where the active tyrosine and the conserved catalytic residues (R-K-H-R-[H/ W]-Y) reside. XerA proteins are well conserved between the archaeal species analyzed with 85% of sequence similarity. The *xerA* gene location is highly variable in archaea; some species exhibit separated *xerA/dif* sequences whereas some others harbor an individual *xerA/dif* module. Unlike most bacteria, the *dif*-like site is not normally located at 180° from *oriC* and conversely, it is located between 122° to 144° from *oriC* in the analyzed genomes, although the *Methanosphaera stadtmanae* genome showed a *dif*-like site at 180° from *oriC*. *dif*-like sequences consist of the traditional structure; two inverted repeat sequences of 11 bp separated by a central region of 6 bp. XerA catalyzes cleavage reactions without any detectable strand preference [162,163].

Although archaea do not require a FtsK homolog to perform chromosome resolution, KOPS-like motifs have been found in Archaea. These KOPS-like motifs consist of four nucleotides (GTTG OR GTTC) called ASPS (Archaea Short Polarized Sequences) that are skewed towards *dif* sites, showing a similar triangle-shaped diagram observed in Bacteria of skew inversion at *dif* sites [162]. Serre et al. (2013) have revealed the crystal structure of XerA proteins from *Pyrococcus abyssi*, and Chang et al. (2016) from *Thermoplasma acidophilum*. Both groups reinforced the idea of *cis*-cleavage reaction by XerA.

7. Future directions

Much information has been gained on site specific recombinases and dimer resolution. This review has highlighted the complexity of *dif*/Xer recombinase systems in prokaryotes and its importance for genome stability and pathogenicity factors. However, many fundamental questions remain unanswered: how do SpoIIIE and SftA from *Bacillus* activate site-specific recombination? Moreover, what is the selective advantage of having two chromosome DNA translocases? Additionally, 12% of the studied proteobacterial species do not possess the traditional Xer recombination machinery. Thus, it is still unknown whether these microorganisms lost the Xer recombination system, never acquired it or developed an alternative system to decatenate the chromosomes. It raises the question of how do bacterial cells handle chromosome decatenation without Xer recombinases and *dif*? Is there an alternative recombination system

that functions as the *dif*/Xer system? Regarding single recombinases, have they evolved from XerC/XerD recombinases or vice versa, or did they arise from an ancestral recombinase? These and other issues already considered in this review are being gradually addressed by the use of the latest techniques in real-time imaging with super-resolution microscopy. Such as; photo-activated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM), plus the use of other techniques as Förster Fluorecence Resonance Energy Transfer (FRET), tethered fluorophore motion (TFM), single-molecule Flourescence resonance energy transfer (smFRET), among others. They are providing a powerful blueprint for investigators studying short- and long-range changes in DNA, DNA/protein, and protein/protein interactions. Researchers in site-specific recombination systems and protein-protein interactions might be the most direct beneficiaries of these techniques, especially when it is becoming urgent to further understand IMEX insertion and its subsequent influence in antibiotic resistance and bacterial virulence.

Literature cited

Page 98

Chapter II

This article focuses on EMSA experiments and pull-down analysis in order to identify the determinants within the *difSL* sequence for XerS binding and cleavage and its interaction with translocase FtsK. Although the work done by Le Bourgeois and Nolivos discovered and characterized to some extent the XerS/*difSL* system, no data are available regarding *dif* sequence and its relationship with the unique characteristics of XerS. Therefore, to better understand XerS interactions with *difSL*, we have used several mutated *difSL* sequences. These pin-pointed mutations were selected to see the effect of three important modifications; First, modifications on highly conserved nucleotides among Firmicutes, second, modifications on size of the targeted DNA fragment, and third, evaluation and comparison with previous reported modifications in order to assess if our results offer similar outcomes. The differences in recombinase interaction with these mutants provides a framework for understanding the conditions for single recombinases to perform dimer resolution.

Article

Determinants of selectivity in XerS site-specific recombination

Fabio Castillo* and George Szatmari

Département de microbiologie, infectiologie et immunologie

Université de Montreal

Keywords: Site specific recombination, XerS, *Streptococcus suis*, EMSA, tyrosine recombinases.

Emails: fabio.andres.castillo.martinez@umontreal.ca, george.szatmari@umontreal.ca
***Correspondence:** George Szatmari; e-mail: george.szatmari@umontreal.ca

Abstract

Chromosome dimer resolution in most bacteria is solved by a double recombinase system of two paralogous highly conserved proteins called XerC and XerD that act on a site of 28 bp called *dif* located at the terminus region. In some *Lactococci* and *Streptococci* this process is mediated by an atypical single site-specific recombinase called XerS, which acts on a larger site called *difSL* of at least 31 bp equally located at the terminus region. Previous characterization of the XerS/*difSL* system reported that XerS has a stronger binding affinity for the left site than the right site of *difSL* and cleavage reactions reported stronger activity with the bottom strand than the top strand. Similarly, it possesses an unusually large central region of 11 bp, longer than any other central region known currently. To understand what kind of determinants drive binding and cleavage selectivity of a single recombinase, a series of *difSL* variants were designed. Our results showed that nucleotides at the outermost region of a T rich region seem to be determinant for binding selectivity and modifications of the extra spacing between the inverted repeat arms as well as reduction of the central region change cleavage preference. The XerS system selectivity mainly depends on *difSL* sequence composition even though most mutations did not abolish binding or cleavage reactions. Our results also showed an XerS-FtsKγ interaction by *pull-down* assays.

1. Introduction

Xer site specific-specific recombination is known for solving dimer formation in microorganisms with circular chromosomes. It is an essential process that guarantees faithful segregation of chromosomes during cell division. During replication, a variety of obstacles can stall or collapse a replication fork, from DNA damage caused by exogenous or intracellular genotoxic stress to collisions between replication forks and DNA binding proteins such as transcribing RNA polymerases [25]. Back in 1968, Rupp and Howard-Flanders suggested that the replication fork can bypass such barriers and reload downstream from the DNA lesion, this would create a gap that would be eventually sealed by the homologous directed gap repair (HDGR), a model extensively supported by other groups [164–166], and reviewed by the following papers [15,31]. However, rescuing of the impaired replication forks often leads to chromosome rearrangements between the newly synthesized strand and the homologous strand. McClintock first proposed this topological problem in 1932. She recognized that in linear DNA, crossover events do not modify the separation between sister chromatids, however, crossing over between circular chromosomes will form concatenated chromosomes if an odd number of these rearrangements occurred[3,4]. It is particularly interesting for circular chromosomes that this topological conformation saves the problem of instability at the DNA ends but renders initiation and termination process potentially lethal when impaired [167]. The single origin on many circular genomes makes fork progression–blocking events dangerous in the absence of replication restart proteins (RRP)[164,168,169] and the catenation bias at the end of replication plus post-replication repair mechanisms make crossing over events dangerous in the absence of Chromosomally Encoded tyrosine Recombinases (Xer) [73,170–172]. The necessity to solve this problem would explain why Xer recombinases and their sites, called *dif* sites, are highly conserved among bacteria [54,155,157,173], even more conserved than other proteins for replication termination as Tus proteins [174]. Initial characterization of site-specific recombination for plasmid stabilization begun in 1981 by two groups; Austin's group and Hamilton's group and their work with bacteriophage P1 [6,175] by reporting the importance of enzyme Cre (Causes Recombination) in plasmid intramolecular recombination. Later, Summers and Sherratt corroborated the importance of these site-specific sites in other plasmids such as

ColE1 [7]. Stirling and collaborators designated the acronym *xer* (Chromosomal ColE1 Recombination function/s) to two genes required for recombination of ColE1 called *xerA* and *xerB* [59]. Interestingly, it was later known that XerA and XerB are important accessory factors for plasmid intramolecular recombination called ArgR and PepA and not tyrosine recombinases as Sherratt's group would established later in 1990, with the identification of the tyrosine recombinase XerC [8] and the first characterization of a site-specific recombinase system involved in chromosome dimer resolution that acts on the site-specific recombination site named *dif* (Deletion-Induced Filamentation) [73,176]. A second tyrosine recombinase called XerD was rapidly identified three years later by Sherratt's group [9]. Since then, the XerC/XerD system has been found in several other families of bacteria such as Proteobacteria and Firmicutes [50,108,121,155,162,173,177–179]. It is highly conserved among bacteria as different experiments showed a positive interaction between XerCD recombinases from other bacteria and the *Escherichia coli dif* site[180,181]. The bases of the system have been extensively studied in XerCD/*dif* from *E. coli* and the homologs RipX-CodV/*bsdif* from *Bacillus subtilis* [121]. Where two tyrosine recombinases work synergistically to recombine two opposed but closed *dif* sites located in the DAZ region (*dif* Activity Zone) to be activated by translocase protein FtsK in *E. coli* [77,81] , or brought together by two translocase proteins SpoIIIE and SftA in *B. subtilis* or homologs FtsK and SpoIIIE proteins in *Staphylococcus aureus* [123,125–127,134]. However, the XerCD-*dif* dimer resolution system is not unique, single Xer recombinases have been found in some *Streptococci* and the *Lactococci,* denominated XerS [157], some ε-proteobacteria (XerH) [50,66] and Archaea (XerA) [155,162,163], summarized by the following papers [50,177,182]. In double recombinase systems, the presence of two recombinases confers an additional level of asymmetry that places the two-step DNA exchange under separate control facilitating coordination and directionality of the process, whereas in single recombinases is still unknown, and coordination or directionality might be exclusively regulated by the recombination site [76].

To understand more about how single recombinases interact with their *dif* sites, we have used several *dif* variants that examine the importance of nucleotides in the central region and outlying *dif* regions. These nucleotides were chosen based on their conservation among bacteria and their effect on other Xer systems, thus facilitating analysis of single recombinase binding and cleavage.

The differences in recombinase interactions with these variants offer a framework for understanding the mechanisms of single recombinases dimer resolution and how these modifications lead to selectivity for a particular recombination outcome.

2. Material and methods

2.1 Bacterial strains and plasmids

E. coli turbo and DH5α (NEB) were used for cloning and purification of plasmids containing *difSL* and *difSL* variants. The *difSL* gene was amplified by PCR from genomic DNA of *S. suis* S735 serotype 2 and cloned into the SmaI site of vector pUC19 by standard cloning techniques as described by Leroux and collaborators [158] using primers listed in **Table 1**. *E. coli* strain NEB T7 *lysY/ I^q* and BL21 (DE3) was used for overexpression of N-terminal MBP-fused genes and His-fused genes respectively. Full-length *xerS* from *S. suis* (NCBI: YYK_03895) and *xerH* from *Campylobacter jenuni* (accession number YP_178952) were amplified and cloned into plasmid pMalC2 in the XmnI restriction site and pEQ30 between BamHI/PstI restriction sites (NEB) respectively. This resulted in vectors pmal-XerS that express XerS with an N-terminal MBP tag and pQXerH that express XerH with an N-terminal His tag. Extraction and purification procedures were as described by Leroux et al. [158]. XerH overexpressed proteins were used as controls to corroborate retained bands on the EMSA experiments are generated by specific binding and not by protein aggregation. $XerS^{Y341F}$ mutant and *difSL* variants constructs were prepared using the Q5 site-directed mutagenesis kit (NEB) with primers listed in Table 1 and used according to the supplier's conditions. These mutations created the vector (pmal $XerS^{Y341F}$) that expresses $XerS^{Y341F}$ with an N-terminal MBP tag designated MBP-$XerS^{Y341F}$ and pUC19*difSL* variants designated as (pUC19-M1 to M11). All DNA substrates were recovered and purified by conventional mini-prep procedures using GeneJet plasmid miniprep or GeneJet PCR purification Kit from ThermoFisher. DNA concentration was determined by Nanodrop analysis (Thermo Scientific) and by QuBit fluorometry (Invitrogen). DNA-binding substrates for mobility shift assays were amplified using 5' HEX-labelled M13 universal sequencing primers (AlphaDNA, Montreal, QC). All cloned PCR products were verified

by sequencing at the IRIC genomic facility of the Université de Montréal. Restriction enzymes, Vent DNA polymerase, Phusion DNA polymerase, Q5 hot start polymerase, and T4 DNA ligase, were obtained from New England Biolabs (NEB) and used according to the supplier's conditions.

2.2 Protein Production and purification

E. coli T7 cells harboring pMal-XerS, pMal-XerS^{Y341F,} pET28-FtsK$_{\gamma}$ and pET28-XerS were grown at 37°C up to an $OD_{600\ nm}$ of~0.4 in LB broth containing Ampicillin (100 µg mL^{-1}) and Kanamycin (30 µg mL^{-1}) respectively. Protein expression was induced by the addition of 0.5 and 1mM isopropyl-D- thiogalactopyranoside (IPTG) respectively and incubated for four additional hours at 37°C, after which the cells were harvested, washed and frozen in its respective buffer; either 25mL of MBP lysis buffer (20mM Tris-HCl, 200mM NaCl, 1mM EDTA, 1mM azide, 10mM β-mercaptoethanol, pH 7.4, supplemented with Pierce protease inhibitors) or His lysis buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10mM imidazole supplemented with Pierce protease inhibitors). This was followed by six bursts of sonication and centrifugation at 20.000 x g for 20 min. MBP-XerS and MBP-XerS^{Y341F} expressed proteins were purified on individual MBP-trap HP column (GE Healthcare) according to manufacturer's directions. Similarly, FtsK$_{\gamma}$-His_6 and XerS-His_6 expressed proteins were purified using Ni-NTA affinity chromatography (Qiagen) according to the manufacturer's instructions. Protein concentration was determined using Bradford assays and QuBit fluorimetry.

2.3 DNA-binding and cleavage assays

For electrophoretic mobility shift assays (EMSAs), MBP-XerS was incubated with the respective PCR product of *difSL* site and its variants for 2 h at 37°C. They were amplified using the universal M13 primers 5' HEX labeled, giving, as a result, a PCR product of 339-bp with the corresponding *dif* site. Binding reactions were performed in 15µL containing TENg buffer (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 25 mM NaCl and 5% glycerol), 10 µg of sonicated salmon sperm DNA (average molecular weight of 500 bp). Suicide substrates were created by annealing three oligonucleotides, corresponding to one full-length site (34 bp), either top or bottom strand and the two halves (left and right), either top or bottom strands (17 bp) for the full-length oligonucleotides. For the half-

sites substrates, only two oligonucleotides were used, corresponding to the left or right side. Annealing reactions were carried out in annealing buffer (10 mM Tris, pH 8.0, 50 mM NaCl, 1 Mm EDTA) at 95°C for 2 minutes. Then they were gradually cooled down to 25°C throughout 45 minutes. The final product was a double-stranded DNA containing a nick either in the top or bottom strand at the central region. Unlabeled oligonucleotides were three times more concentrated that labeled ones **(Figure 5)**. Cleavage reactions were performed in TENg buffer containing 8µM of nicked or half substrates with 350 nM of MBP-XerS for 60 min. Reactions were stopped by adding SDS to a final concentration of 0.1% (w/v) and heated to 95 °C for 5 min; samples were then electrophoresed in 6% TBE gels in the presence of 0.1% of SDS. Gels were imaged in a Typhoon 9410 imager, and images were analyzed by IMAGEQUANT software (GE Healthcare).

2.4 *Pull-down* assay

Pull-down assay was performed by immobilizing MBP-tagged XerS (83 kDa) proteins on amylose magnetic beads (NEB) in MBP column binding buffer 200mM NaCl, 20 mM Tris-HCl pH 7.4, 1 mM EDTA and 1 mM DTT for 2 hours at 4°C under constant agitation. After series of washes to remove any unbound MBP-tagged protein, the His-tagged protein (FtsKγ, 17 KDa) was loaded under the same conditions previously described. Proteins that were not immobilized due to physical interaction with MBP-XerS were removed in the subsequent washing steps. The experiment was carried out by using directly cell lysates samples of both proteins, since it seems to add non-specific controls, increases competition and it approaches to *in vivo* conditions. The Interaction was assessed on 0.1% SDS-15% PAGE gels following a selective co-elution of the interacting proteins by adding 10mM maltose to the MBP column binding buffer. Samples were determined by staining the gel with Coomassie blue, followed by scanning in an Amersham Imager 600 and subsequent quantification.

3. Results and discussion

3.1 Characterization of essential nucleotides in *difSL* involved in binding and recombination reactions.

To determine the interaction between XerS and *difSL*, the *S. suis xerS* gene and its *difSL* site were identified by sequence analysis, amplified by PCR, and cloned into plasmid vectors (pMal-c2x, pET28 and pUC19 respectively). As described by Le Bourgeois [157], the defined minimal site of the Streptococcal *difSL* was determined as 31 bp in length with the characteristic structure of other *dif* sites (i.e., two imperfect inverted repeats separated by a central region). This was measured by calculating the integration efficiency of vector pGh9 containing different variants of the *difSL* region [157]. Subsequently, DnaseI footprinting analysis and EMSA experiments of the lactococcal XerS/*difSL* complex showed a longer protected region, presumably longer than the 31 bp[159]. Further experiments showed, by the use of suicide substrates of the Streptococcal *difSL* site, that the length of the central region is 11 bp[158], considerably longer than most site-specific recombination sites [177]. Presently, the minimal *difSL* site is defined by two imperfect inverted repeat sequences of an uncharacterized length separated by a central region of 11 bp. To determine the exact minimal site and which nucleotide(s) within this minimal site are important for XerS recombination, we performed several gel retardation assays using MBP-XerS protein with eleven different *difSL* variants. These mutations were separated into four groups **(Table 2)**, mutations at the cleavage site (M1 and M2), mutations at the central region (M3, M4, M5, and M6), mutations at the binding site (M7, M8 and M9) and mutations at the outer regions (M10 and M11). Additionally, substrates with one of the two binding sites deleted were also constructed by site-directed mutagenesis and tested. The eleven EMSA experiments were performed using the same range of protein concentration that ensured at its maximal concentration (Lane 7, 700 nM) >99% MBP-XerS binding to *difSL*-*WT*, with equal concentrations of DNA (2ng/µL) and salmon sperm (100ng/µL).

All titrations followed a similar pattern of MBP-XerS/*difSL* interaction, with almost equal formation of complexes I and II at low MBP-XerS concentration, corresponding to the occupancy

of just one arm by a single monomer of XerS (Complex I) or both arms (Complex II) by the occupancy of two monomers of XerS on *difSL* [159]. At higher MBP-XerS concentration, subsequent and gradual disappearance of complex I and stronger formation of complex II takes place. *difSL WT* and most of the *difSL* variants showed formation of additional bands under 115 nM or higher concentrations of MBP-XerS (herein referred as complex III, as they might correspond to protein-protein interactions between two or more complexes II [183]) with exception of M1 and M7, (**Figure 1**). None of the mutations tested in the full-length *difSL* site abolished XerS binding completely. However, all mutations decreased to some extent binding activity, as equally reported by Jo and colleagues[183] in the XerA/*dif2* system. Mutations in the central region (M4, M5 and M6), under XerS saturation (700 nM) showed non-free DNA, similarly to *difSL-WT* with the exemption of M3. Whereas mutations at the binding sites (M1, M2 and M7) showed persistent free-DNA complexes, as expected, especially the M7 mutant, which showed poor formation of complex II and undetectable formation of complex I and III, resulting in a higher amount of free DNA. Modifications of the central region have demonstrated to reduce the affinity of XerC for the *dif* site, in fact, altering the size of the central region seems to determine the selectivity for intramolecular or intermolecular recombination processes[184]. Similarly, Lee and Saito (1998) demonstrated how substitutions of nucleotides 2 to 7 of the *loxP* central region affected recombination processes, particularly, substitutions of nucleotides 6 and 7 block first strand exchange and substitutions of nucleotides 2 to 5 avoid resolution of Holliday junction intermediates. Lee and Jayaram equally proved that mutations at the central region of the *frt* site decreased recombination processes, pointing out the importance of the central region for resolution[185]. For phage lambda Int, the *att* central region plays an important role for Holliday junction resolution[186]. Therefore, it is not surprising that mutations M4, M5 AND M6 might have affected to a very low level binding reactions as shown in **Figure 1**. Most of free-DNA is not detected using 352 nM or 532 nM of MBP-XerS, although it is still not as proficient as wild-type *difSL* [74] with a unique exemption of M4, in which the 2 nucleotide substitution in the inner part of the central region produced a similar pattern as wild-type *difSL*. Interestingly, the M3 and M5 mutations involved modifications of the length of the central region. In M3, one nucleotide deletion in the central region proved to be more detrimental than any other modification at the

central region whereas the M5 mutation (addition) gave a similar pattern to that of wild-type *difSL*. These findings suggest the flexibility of the protein-DNA complex is considerable and that the response for space modifications is asymmetric, since adding one nucleotide had a small effect while removing one nucleotide affects cooperativity, as previously noticed by Mao et al[187] with other DNA binding proteins. Our results are corroborated by Bebel's results in which the crystal structure of the synaptic complex XerH/*difH* showed low interaction between XerH and the central region[66] and that the M3 binding deficiency might be spatial-related. On the other hand, free DNA was still detectable on the EMSA gels with *difSL* mutants (M1*, M2*, M3* and M7, *cleavage site mutants**) under high concentration of protein (700 nM). This agrees with previous investigations where modifications at the binding site of *dif* in *E. coli* decreased XerCD proficiency in recombination. Their EMSA experiments also showed an overall reduced binding affinity, obtaining up to 50% more of free DNA compared to wild-type *dif* [188,189]. Similarly, Jo and colleagues showed that 11 point mutations in *dif2* at the binding site equally decreased binding efficiency by XerA for each nucleotide at different degrees [183]. Bebel's crystal structure showed a direct interaction between XerH and *difH* at the cleavage section, up to 5 nucleotides are contacted by XerH while the outer part is mostly contacted at the phosphate backbone with the exception of one conserved thymine base and the three outermost nucleotides (AGT)[66]. Interestingly, these papers coincided in that adjacent nucleotides of the cleavage point are relevant for binding as shown here with M1 and M2 mutants. However, our experiments showed that M7 modifications resulted in the lowest complex formation with the highest amount of free DNA. The M7 mutation switches the corresponding extra space TT/26-27 from the right arm the nucleotide C/8 in the left arm to see if it drives XerS binding to the correct site at the early stages of synapsis **(Table 2).** Our result showed that this mutation decreased XerS binding to the full-length M7 site. Then, we tested if half-sites of the same mutation would drive binding preference for the right half-site (RHS) rather than the left half-site (LHS) **(Figure 2).**

Half-site M7 mutants showed stronger XerS binding interaction with M7-LHS than M7-RHS, a similar pattern previously demonstrated by Leroux and colleagues[158]. M7-LHS binding is still weaker compared to *difSL* WT, whereas M7-RHS binding was undetectable under these conditions **(Figure 2)**. Unexpectedly, M7 half-sites showed the same migrating pattern as complex

II **(Figure 2, center)**, despite the absence of the second half. The presence of the extra nucleotide seems to not be determinant for binding preference and on the contrary, its deletion exacerbated XerS binding to the right half-site when compared to *difSL* wild-type. In the XerH/*difH* system, it is the left arm that possesses preferential binding and cleavage, it also contains a single base-pair insertion. However, this single base pair insertion is not directly involved with XerH binding, which our results seem to corroborate.

Furthermore, Bebel's group noticed that a conserved thymine (T/4) and three adjacent nucleotides of the outer region are highly determinant for cleavage and binding reactions, similar results were initially noticed by Hayes and Sherratt where specificity determinants are confined to the outer ends of the *E. coli dif* site synaptic complex[189]. Therefore, we decided to create M8 and M9 mutants in which nucleotides T/5→A (M8) and nucleotides TT/2-3→AG (M9) were changed respectively. T/5 is highly conserved among 49 available streptococcal genomes as well as in *E. coli*, *Bacillus* and several proteobacteria species *dif* sites [50,157], whereas TT/2-3 is part of the outermost nucleotides of the *difSL* site.

M8 and M9 modifications reduce complex II formation and seem to inhibit complex I and complex III, **(Figure 3A)**. Interestingly, when we decided to study these mutations with their respective half sites alone, either left or right, both mutations completely inhibited binding, this was expected at the right half-site due to previous analysis of known weak interaction, however, it was the first time that XerS binding was entirely inhibited by point mutations in a half-site **(Figure 3B).** For the M9 modification, this result can be supported by the fact that the left site is more divergent than the highly conserved right site, as the consensus region 5'-AT**C**TTTC-3' in the left arm of most Streptococcal genomes and some Lactococcal species has a cytosine at the third nucleotide, whereas *S. suis dif* site has a thymine at that position 5'-AT**T**TTTC-3'. Therefore, it is reasonable to expect that this nucleotide might have mutated to increase XerS specificity for the left half-site and therefore its modification affects binding significantly. For the M8 modification, we rationalized that modification of the fifth thymine can have a deleterious effect as reported by Bebel's group (T5'), which we can now corroborate as an essential base for binding. This system obeys a similar pattern found in the XerC/XerD-*dif* system and the CodV/RipX-*bsdif* system which

consists of a highly conserved right arm (XerD/RipX binding arm) and a more divergent left arm (XerC/CodV binding arm).

Additionally, XerS seems to behave like XerC/CodV which exhibit less efficient binding but more efficient cleavage [121] as shown here **(Figure 2,4)**. Subsequently, a rich TTT region was placed at the outer end of the right half-site with the goal of resembling the outer end of the left site to prove if they are determinants for binding specificity. Therefore mutants M10 and M11 were created. M11 mutant includes removal of the extra space TT/26-27 for a C **(Table 2).** However, these mutations did not alter binding preference towards the right half-site but kept the same pattern of weak or undetectable XerS binding to this site **(Data not shown).** The precise determinants of XerS binding to *difSL* remain unclear, however, it is important to discard two important facts, the outer rich TTT region at the left site neither the extra nucleotide at the right site are determinants for XerS binding preference in in vitro experiments.

The results of cleavage preference were performed through the use of suicide substrates, initially proposed by Pargellis and collaborators[190]. The catalytic Tyr-341 of XerS attacks the scissile phosphate bond of the DNA backbone at the cleavage site, which is located upstream of a pre-existing nick. Once XerS binds and cleaves the DNA, the fragment between the cleavage site and the nick is released under denaturing conditions, taking with it the energy from the phosphodiester bond cleavage and consequently trapping the phosphotyrosine-linked intermediate with the labeled DNA. Full-length suicide substrates were constructed and designed as M1 to M7 with the respective nick position, either in the top (TN) or in the bottom (BN) strand. Our experiments agreed with previous experiments in which XerS showed higher intermediate accumulation with the bottom strand nicked over the top strand nicked oligonucleotides [158,159] **(Figure 4)**.

Further experiments with *difSL* mutants showed that M2, M4, M5 and M6 maintained a similar cleavage pattern as *difSL* wild type (Bottom strand preference). M6-TN exhibited strong inhibition of intermediate formation whereas M5-BN exhibited stronger intermediate accumulation than *difSL* wild-type, M4, M5 and M6 oligonucleotide mutants involved central region modifications and/or spatial modifications. Our results agree with other central region modifications such as; the XerA/*dif2* system that exposed the low correlation between binding and cleavage efficiency,

as some mutants exhibited low cleavage reactions but active binding activity. Furthermore, all *dif2* mutations decreased cleavage intermediate formation to some extent which corroborates why even our M1 and M2 mutations did not completely abolish cleavage. For the XerH/*difH* system, it was demonstrated that one nucleotide of the central region nearby the cleavage site has not any effect on cleavage reactions. Unexpectedly, M1, M3 and M7 mutants revert cleavage preference, they all exhibited stronger TN intermediate accumulation than BN intermediates. M1 represents a cleavage site modification, M3 one base deletion from the central region and M7 exchange of the extra space between arms. This variety of modifications with a similar outcome corroborates the idea that various factors and not only one-to-one direct contact between protein residues and DNA bases directs binding or cleavage specificity [191]. As Baldwin's group described: "The protein-DNA interface is a extend hydrogen-bonded red connected by water molecules, and the interaction of this red can change in different ways in response to mutation, affecting either, the catalytic step or the binding step of the reaction" [56]. However, we noticed that M3 and M7 are involved in spatial modifications as well as M5 and M6, while M3 and M7 increased TN intermediate formation and almost abolished BN intermediates, M5 and M6 decreased TN intermediates and increased BN intermediate formation. Therefore, it is reasonable to assume that spatial modifications might play an essential role in cleavage preference. As previously described by Mao and collaborators [187]; the addition of one nucleotide will rotate the regular positions of the two recombinase binding sites by -34 ° and separate them by 3.4Å. Such rotation and spatial separation may dictate the ability to disrupt recombinase interaction across the intervening DNA, especially for site-specific recombinase sites that are characterized for its bending capacity, flexibility and the fact that just a few protein residues directly contact the *dif* bases as previously reviewed by the following papers, [56,63]. For instance, In the XerH/*difH* system, spatial modifications abolished cleavage reactions. Half-site suicide substrates revealed a similar pattern of cleavage as full-length suicide substrates **(Figure 5).** One characteristic of these results is that intermediate formation was abolished or highly reduced, which can be caused by the absence of the other XerS monomer and its synergetic effect on cleavage reactions.

3.2 The γ subdomain of FtsKc interacts with XerS

In order to determine if FtsK interacts with XerS, a truncated FtsK derivative was constructed containing only the final 125 amino-acids of the FtsK c-terminal protein attached to a His_6 tag, here denominated His-FtsKγ. The 125 amino acids comprise the entire FtsKγ-subdomain and 64 final nucleotides of the FtsKβ-subdomain. The FtsKγ-subdomain, is responsible for tyrosine recombinases recognition and KOPS (FtsK Orienting Polar Sequences) sequences in most bacteria [99,192], among the 125 amino acids, the final 62 nucleotides are sufficient for the ATP-independent XerD recognition and interaction with the double recombinase system (XerC/XerD) [2].

The determinants for XerD recognition were initially thought to rely mostly on the specific sequence TEKRKA of FtsKγ for the double recombinase system XerCD in γ-proteobacteria, however, Keller and collaborators discovered that only the 9 adjacent amino acids (underlined) (**Q**-**R**QF**RI**G**N**R) can clearly support XerD recognition and recombination with 5 of these amino acids in direct contact with XerD protein (bold letters), the exact same residues responsible for KOPS binding [99]. Amino acid sequence alignment between the corresponding γ-subdomain of *S. suis* and *E. coli* shows 39% of level of identity and 69% when compared with *L. lactis* FtsK γ-subdomain. Both, the RQFRIGYNR and the TEKRKA sequences are weakly conserved in either *L. lactis* and *S. suis,* however, four of the 5 amino acids involved in contacting XerD in the crystal structure from *E. coli* are the same. Therefore it has been hypothesised that the λ subdomain must have co-evolved with the respective Xer recombinases of each bacterial species [2]. This hypothesis is corroborated by the corresponding interacting residues in the XerDs enzymes, in the *L. lactis* XerS, only 1 of the 5 amino acids is identical to *E. coli* XerD whereas in the *S. suis* XerS enzyme only 2 are identical of the 5 amino acids. Alignment of *E. coli* XerD and *S. suis* XerS also shows high divergence with only 24% of percentage of identity **(Figure 6)**.

Our pull down assay between immobilized MBP-XerS and His-FtsKγ showed both proteins co-eluting with maltose (10mM), lanes with sample (MBP tag and His-FtsKγ) did not show any retained protein corresponding to the bait His-FtsKγ, which proved that the bands observed in the lanes (MBP-XerS) and/or (MBP-XerS+DNA) corresponding to His-FtsKγ protein are not unbound residues from incomplete washing steps. Both proteins were also tagged differently to support our initial results, by tagging XerS with a His_6 tag and FtsKγ with a MBP tag, however,

solubility of His-XerS protein proved to be challenging and the ratio of protein purification was much lower when compared to MBP-XerS (Data not shown).

Interestingly, PCR fragments of the *difSL* region plus 120 bp adjacent nucleotides improved His-FtsKγ retention. This might be due to an aggregation effect caused by *difSL* since bound XerS subunits are brought to close proximity, making it more accessible for His-FtsKγ to interact with a higher amount of XerS subunits otherwise randomly spread in a liquid interphase. It is also important to mention that FtsK interaction with tyrosine recombinases is very rapid, May and collaborators defined the time of FtsK interaction with XerD in 0.5s, an extremely short time to assess FtsK and XerS interaction by *in vitro* experiments such as pull-down, however, it is still worth noting that these interactions could be detected in this paper. Another important feature of FtsK is its bi-functional capacity to make protein-protein and protein-DNA interactions, using the same surface to carry out both functions. This means that mechanistically FtsKγ cannot interact with KOPS DNA and XerD simultaneously [99]. Regarding protein-protein interaction between FtsK and Xer proteins, few papers have assessed this interaction directly, most papers measure this interaction by assessing and comparing percentage efficiency on DNA recombination and/or resolution by FtsK and XerD/XerC mutants versus wild type [77,85,97]. It was only in 2006 that Yates and collaborators performed pull down assays by measuring immobilized FtsKγ versus XerC and XerD proteins, and vice versa, immobilized XerC/XerD proteins versus FtsKγ protein, showing strong interactions between both FtsK and XerD c-terminal regions, later on, Keller and collaborator in 2016 showed a visual representation of what could be FtsKγ and XerD interaction, by the use of a chimera protein (Fused C-terminal domain of XerD plus FtsKγ subdomain) denominated $XerD_C$–γ, providing biochemical evidence of this interaction. The lack of information regarding direct interaction between Xer proteins and FtsK motivated us to investigate if it would be plausible to detect and measure FtsK interaction with XerS, and as shown in Figure 6, immobilized MBP-XerS proteins are capable to retain FtsK proteins after a series of washing steps, our results prove that *S. suis* FtsKγ subdomain is capable to recognize and bind to XerS, this interactions are ATP-independent.

4. Perspectives

Further experiments are needed if we want to fully understand single recombinase systems; *in vivo* plasmid integration efficiency should be used to test resolution efficiency which was not evaluated by the techniques used in this paper. An increasing number of papers have addressed protein expression and protein activity by single-molecule imaging techniques. However, just a few have used these techniques to study single recombinases. These techniques will allow a better understanding of the dynamics of single recombinases during replication and segregation. Footprinting analysis and Surface plasmon resonance (SPR) techniques should corroborate and offer more details about binding reactions and their kinetics respectively, and, SPR analysis would allow to compare the rates of the binding efficiency of each *difSL* variant. It would be interesting to test if FtsK interaction with XerS is also as rapid as demonstrated with the XerCD/*dif* system by SPR analysis, and if possible, colocalization of these two proteins using *in vivo* experiments in *S. suis*.

5. Literature cited

Page 98

Acknowledgments

The authors acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (Discovery Grant 106085-2013). FC received salary support from the Département de microbiologie infectiologie et immunologie and the Faculté de Médecine, Université de Montreal.

The authors claim no conflict of interest.

Table 1. Strains, plasmids, and primers used in this study.

Purpose	**Strain, plasmid and primer name**	**Genotype, properties and sequence**	**Source or reference**
Strain			
Cloning strain	*E. coli* DH5α	*supE44 ΔlacU169 (ϕ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	NEB
Expression strain	E. coli T7 *lysY/I^q*	MiniF *lysY lacI^q*(Cam^R) */ fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--* Tet^S*)2 [dcm] R(zgb-210::Tn10--* Tet^S*) endA1 Δ(mcrC-mrr) 114::IS10*	NEB
Plasmid			
Cloning vector; Ap^r	pUC19		NEB
Cloning vector; Ap^r	pUC19-difSL	pUC19 with *difSL* sequence of *S. suis*	This study
Cloning vector; Ap^r	pUC19-M1	pUC19 with *difSL* variant M1	This study
Cloning vector; Ap^r	pUC19-M2	pUC19 with *difSL* variant M2	This study
Cloning vector; Ap^r	pUC19-M3	pUC19 with *difSL* variant M3	This study

Cloning vector; Ap^r	pUC19-M4	pUC19 with *difSL* variant M4	This study
Cloning vector; Ap^r	pUC19-M5	pUC19 with *difSL* variant M5	This study
Cloning vector; Ap^r	pUC19-M6	pUC19 with *difSL* variant M6	This study
Cloning vector; Ap^r	pUC19-M7	pUC19 with *difSL* variant M7	This study
Cloning vector; Ap^r	pUC19-M8	pUC19 with *difSL* variant M8	This study
Cloning vector; Ap^r	pUC19-M9	pUC19 with *difSL* variant M9	This study
Cloning vector; Ap^r	pUC19-M10	pUC19 with *difSL* variant M10	This study
Cloning vector; Ap^r	pUC19-M11	pUC19 with *difSL* variant M11	This study
Overexpression vector; Ap^r	pMalC2		NEB
Overexpression vector; Ap^r	pMalC2-XerS	pMalC2 with x*erS* gene of *S. suis*	This study
Overexpression vector; Ap^r	pMalC2-$XerS^{Y341F}$	pMalC2 with a mutated x*erS* gene of *S. suis,* in which Tyr-341 is replaced by phenylalanine.	This study
Overexpression vector; Ap^r	pQE30		QIAGEN
Overexpression vector; Ap^r	pQE30-XerH	pQE30 with *xerH* gene of *Campylobacter jenuni*	This study
Primers			
Oligonucleotides for amplification processes (5'-3')			
Primers for *difSL* amplification from *S. suis* genome.	Dif-SL-F	TTCCAGTTTTGTCGTTATTAAAGTAC	
	Dif-SL-R	TTTCTTTTAGTTGATCAATTTTTTCC	
Primers for *xerS* amplification	SsuisXerCFwd	GATGAGACGCGAGTTATTATTGG	

from *S. suis* genome.	SsuisXerCRev	TCACAACTGATCCAGAGCAT
Site directed mutagenesis of *xerS* (5'-3')		
Primers for *xerS* site directed mutagenesis to obtain the $xerS^{Y341F}$ allele.	SsuisXerC-SDM-Fwd	CACCGATCTCTTTACCCATATCGTC
	SsuisXerC-SDM-Rev	ACCTGGGTATTGGCATGG
Site directed mutagenesis of pUC19-*difSL* wild type to obtain pUC19-*difSL* variants (5'-3')		
difSLM1	difSL-M1-F	TATAATTTTGTTGAAACTTATATAAGGTTATGCTATACTAC
	difSL-M1-R	GTTTTTGGGAAAAATGTTTGTGGGGATATTAGAAAG
difSLM2	difSL-M2-F	TATAATTGGCTTGAAACTTATATAAGGTTATGCTATACTAC
	difSL-M2-R	GTTTGGCGGAAAAATGTTTGTGGGGATATTAGAAAG
difSLM3	difSL-M3-F	AACTATAATTTTCTTGAAACTTATATAAGG
	difSL-M3-R	TTCGGAAAAATGTTTGTGG
difSLM4	difSL-M4-F	TAGTATTTTCTTGAAACTTATATAAGGTTATGCTATACTAC
	difSL-M4-R	GTTTTTCGGAAAAATGTTTGTGGGGATATTAGAAAG
difSLM5	difSL-M5-F	GATAATTTTCTTGAAACTTATATAAGG
	difSL-M5-R	AGTTTTTCGGAAAAATGTTTG
difSLM6	difSL-M6-F	TAGTTTTTCTTGAAACTTATATAAGG
	difSL-M6-R	TAATTTTCGGAAAAATGTTTG
difSLM7	difSL-M7-F	TATAATTTTCCGAAATTATATAAGGTTATGCTATACTAC
	difSL-M7-R	GTTTTTCAAGAAAAATTGTGGGGATATTAGAAAG
difSLM7-LHS	difSL-M7-LHS-F	TAAGGTTATGCTATACTACTC
	difSL-M7-LHS-R	AATTATAGTTTTTCAAGAAAAATTGTG
difSLM7-RHS	difSL-M7-RHS-F	AAACTATAATTTTCCGAAATTATATAAG
	difSL-M7-RHS-R	TTGTGGGGATATTAGAAAG
difSLM8	difSL-M8-F	TAATTTTCTTGACACTTATATAAGGTTATGCTATACTACTC
	difSL-M8-R	TAGTTTTTCGGATAAATGTTTGTGGGGATATTAGAAAG

difSLM8-LHS	difSL-M8-LHS-F	ATAAGGTTATGCTATACTACTC
	difSL-M8-LHS-R	AATTATAGTTTTTCGGATAAATGTTTG
difSLM8-RHS	difSL-M8-RHS-F	AAACTATAATTTTCTTGACACTTATATAAG
	difSL-M8-RHS-R	TTGTGGGGATATTAGAAAG
difSLM9	difSL-M9-F	TAATTTTCTTGAAAAATATATAAGGTTATGCTATACTACTC
	difSL-M9-R	TAGTTTTTCGGAAACTTGTTTGTGGGGATATTAGAAAG
difSLM9-LHS	difSL-M9-LHS-F	ATAAGGTTATGCTATACTACTC
	difSL-M9-LHS-R	AATTAGTAGTTTTTCGGAAAC
difSLM9-RHS	difSL-M9-RHS-F	AAACTACTAATTTTCTTGAAAATATATAAG
	difSL-M9-RHS-R	TTGTGGGGATATTAGAAAG
difSLM10	difSL-M10-F	AATATATAAGGTTATGCTATACTAC
	difSL-M10-R	TTTCAAGAAAATTATAGTTTTTTGTG
difSLM11	difSL-M11-F	TTTTCCGAAAAATGTTAAGGTTATGCTATACTAC
	difSL-M11-R	TTATAGTTGTGGGGATATTAG
Oligonucleotides for suicide substrates (5'-3')		
difSLWT-BN	difSLWT-TS	AAACATTTTTCCGAAAAACTATAATTTTCTTGAAACTT
	difSLWT-BS-L	AGTTTTTCGGAAAAATGTTT
	difSLWT-BS-R-HEX	AAGTTTCAAGAAAATTAT
difSLWT-TN	difSLWT-BS	AAGTTTCAAGAAAATTATAGTTTTTCGGAAAAATGTT
	difSLWT-TS-L-HEX	AACATTTTTCCGAAAAACT
	difSLWT-TS-R	ATAATTTTCTTGAAACTT
difSLM1-BN	difSLM1-TS	AAACATTTTTCCCAAAAACTATAATTTTGTTGAAACTT
	difSLM1-BS-L	GTTTTTGGGAAAAATGTTT
	difSLM1-BS-R-HEX	AAGTTTCAACAAAATTATA
difSLM1-TN	difSLM1-BS	AAGTTTCAACAAAATTATAGTTTTTGGGAAAAATGTTT
	difSLM1-TS-L-HEX	AAACATTTTTCCCAAAAAC
	difSLM1-TS-R	TATAATTTTGTTGAAACTT
difSLM2-BN	difSLM2-TS	AAACATTTTTCCGCCAAACTATAATTGGCTTGAAACTT
	difSLM2-BS-L	GTTTGGCGGAAAAATGTTT
	difSLM2-BS-R-HEX	AAGTTTCAAGCCAATTATA

difSLM2-TN	difSLM2-BS	AAGTTTCAAGCCAATTATAGTTTGGCGGAAAAATGTTT
	difSLM2-TS-L-HEX	AAACATTTTTCCGCCAAAC
	difSLM2-TS-R	TATAATTGGCTTGAAACTT
difSLM3-BN	difSLM3-TS	AAACATTTTTCCGAAAACTATAATTTTCTTGAAACTT
	difSLM3-BS-L	AGTTTTCGGAAAAATGTTT
	difSLM3-BS-R-HEX	AAGTTTCAAGAAAATTAT
difSLM3-TN	difSLM3-BS	AAGTTTCAAGAAAATTATAGTTTTCGGAAAAATGTTT
	difSLM3-TS-L-HEX	AAACATTTTTCCGAAAACT
	difSLM3-TS-R	ATAATTTTCTTGAAACTT
difSLM4-BN	difSLM4-TS	AAACATTTTTCCGAAAAACTAGTATTTTCTTGAAACTT
	difSLM4-BS-L	GTTTTTCGGAAAAATATTT
	difSLM4-BS-R-HEX	AAGTTTCAAGAAAATACTA
difSLM4-TN	difSLM4-BS	AAGTTTCAAGAAAATACTAGTTTTTCGGAAAAATATTT
	difSLM4-TS-L-HEX	AAACATTTTTCCGAAAAAC
	difSLM4-TS-R	TAGTATTTTCTTGAAACTT
difSLM5-BN	difSLM5-TS	AAACATTTTTCCGAAAAACTGATAATTTTCTTGAAACTT
	difSLM5-BS-L	AGTTTTTCGGAAAAATGTTT
	difSLM5-BS-R-HEX	AAGTTTCAAGAAAATTATC
difSLM5-TN	difSLM5-BS	AAGTTTCAAGAAAATTATCAGTTTTTCGGAAAAATGTTT
	difSLM5-TS-L-HEX	AAACATTTTTCCGAAAAACT
	difSLM5-TS-R	GATAATTTTCTTGAAACTT
difSLM6-BN	difSLM6-TS	AAACATTTTTCCGAAAATTATAGTTTTTCTTGAAACTT
	difSLM6-BS-L	AATTTTCGGAAAAATGTTT
	difSLM6-BS-R-HEX	AAGTTTCAAGAAAAACTAT
difSLM6-TN	difSLM6-BS	AAGTTTCAAGAAAAACTATAATTTTCGGAAAAATGTTT
	difSLM6-TS-L-HEX	AAACATTTTTCCGAAAATT
	difSLM6-TS-R	ATAGTTTTTCTTGAAACTT
difSLM7-BN	difSLM7-TS	AAACATTTTTCTTGAAAAACTATAATTTTCCGAAACTT
	difSLM7-BS-L	TTTTTCAAGAAAAATGTTT
	difSLM7-BS-R-HEX	AAGTTTCGGAAAATTATAG
difSLM7-TN	difSLM7-BS	AAGTTTCGGAAAATTATAGTTTTTCAAGAAAAATGTTT
	difSLM7-TS-L-HEX	AAACATTTTTCTTGAAAAA

	difSLM7-TS-R	CTATAATTTTCCGAAACTT
Oligonucleotides for suicide half-site substrates (5'-3')		
difSLwt left site	difSLwt-TS-L	AACATTTTTCCGAAAAAC
	difSLwt-BS-L	GTTTTTGGGAAAAATGTT
difSLwt right site	difSLwt-TS-R	ATAATTTTCTTGAAACTT
	difSLwt-BS-R	AAGTTTCAAGAAAATTAT
difSLm1 left site	difSLm1-TS-L	AAACATTTTTCCCAAAAAC
	difSLm1-BS-L	GTTTTTGGGAAAAATGTTT
difSLm1 right site	difSLm1-TS-R	TATAATTTTGTTGAAACTT
	difSLm1-BS-R	AAGTTTCAACAAAATTATA
difSLm2 left site	difSLm2-TS-L	AAACATTTTTCCGCCAAAC
	difSLm2-BS-L	GTTTGGCGGAAAAATGTTT
difSLm2 right site	difSLm3-TS-R	TATAATTGGCTTGAAACTT
	difSLm3-BS-R	AAGTTTCAAGCCAATTATA
difSLm3 left site	difSLm3-TS-L	AAACATTTTTCCGAAAACT
	difSLm3-BS-L	AGTTTTCGGAAAAATGTTT
difSLm3 right site	difSLm3-TS-R	ATAATTTTCTTGAAACTT
	difSLm3-BS-R	AAGTTTCAAGAAAATTAT
difSLm4 left site	difSLm4-TS-L	AAACATTTTTCCGAAAAAC
	difSLm4-BS-L	GTTTTTCGGAAAAATATTT
difSLm4 right site	difSLm4-TS-R	TAGTATTTTCTTGAAACTT
	difSLm4-BS-R	AAGTTTCAAGAAAATACTA
difSLm5 left site	difSLm5-TS-L	AAACATTTTTCCGAAAAACT
	difSLm5-BS-L	AGTTTTTCGGAAAAATGTTT
difSLm5 right site	difSLm5-TS-R	GATAATTTTCTTGAAACTT
	difSLm5-BS-R	AAGTTTCAAGAAAATTATC

Table 2. Schematic of *difSL* sequence and a series of mutations in the *difSL* sequence. The upper-case bases indicate the defined minimal site (31 bp), the consensus bases of the *difSL* sequences are shown in boldface and the underlined bases indicate the inverted repeat sequences that match between the left and right sites. The red bases indicate the inserted mutation. The empty squares (-) correspond to deletion of these bases.

	Left site											**Central region**										**Right site**													
Positions	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31*	32	33	34
difSL WT	**A**	**T**	T	T	T	T	**C**	C	G	A	A	A	**A**	**A**	**C**	**T**	A	**T**	**A**	**A**	T	T	T	T	**C**	T	**T**	**G**	A	A	A	c	t	t
difSL M1	**A**	**T**	T	T	T	T	**C**	C	C	A	A	A	**A**	**A**	**C**	**T**	A	**T**	**A**	**A**	T	T	T	T	**G**	T	**T**	**G**	A	A	A	c	t	t

difSL M2	A	T	T	T	T	T	C	C	G	C	C	A	A	A	C	T	A	T	A	A	T	T	G	G	C	T	T	G	A	A	A	c	t	t
difSL M3	A	T	T	T	T	T	C	C	G	A	A	A	A	C	T	A	T	A	A	T	T	T	T	C	T	T	G	A	A	A	c	t	t	a
difSL M4	A	T	T	T	T	T	C	C	G	A	A	A	A	A	C	T	A	G	T	A	T	T	T	T	C	T	T	G	A	A	A	c	t	t
difSL M5	A	T	T	T	T	T	C	C	G	A	A	A	A	A	C	T	G +	A	T	A	A	T	T	T	T	C	T	T	G	A	A	A	c	t
difSL M6	A	T	T	T	T	T	C	C	G	A	A	A	A	T	T	A	T	A	G	T	T	T	T	T	C	T	T	G	A	A	A	c	t	t
difSL M7	A	T	T	T	T	T	C	T	T	G	A	A	A	A	A	C	T	A	T	A	A	T	T	T	T	C	C	G	A	A	A	c	t	t
difSL M8	A	T	T	T	A	T	C	C	G	A	A	A	A	A	C	T	A	T	A	A	T	T	T	T	C	T	T	G	A	C	A	c	t	T
difSL M9	A	A	G	T	T	T	C	C	G	A	A	A	A	A	C	T	A	T	A	A	T	T	T	T	C	T	T	G	A	A	A	c	t	t
difSL M10	A	-	-	-	-	-	-	-	-	-	-	A	A	A	C	T	A	T	A	A	T	T	T	T	C	T	T	G	A	A	A	A	A	t
difSL M11	A	-	-	-	-	-	-	-	-	-	-	A	A	A	C	T	A	T	A	A	T	T	T	T	C	C	G	A	A	A	A	A	t	a

Figure Legends.

Figure 1. Titration of a 339 bp DNA fragment containing *difSL* wild type and *difSL* mutants (M1 to M7) by increasing concentrations of MBP-XerS ranging from 30 nM (lane 1) to 700 nM (Lane 7) in an EMSA experiment (4% polyacrylamide/0.5x TBE gels)**.** The *difSL* sites were amplified by PCR using a 5' 6-Hex-labelled primer obtaining a final concentration of 2ng/µL of purified PCR product. Binding reactions were performed in 20µL containing TENg buffer (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 25 mM NaCl and 5% glycerol) for 2 hours at 37°C. The positions of the free DNA and the XerS/DNA complexes are indicated.

Figure 2. Titration of a 315 bp DNA fragment containing *difSL*-M7 half left site, M7-LHS (A) and half right site, M7-RHS (C) by increasing concentrations of MBP-XerS ranging from 30 nM (lane 2) to 530 nM (Lane 7) in an EMSA experiment (4% polyacrylamide/0.5x TBE gels). **B.** Comparison of equal amounts of *difSL-m7* half site mutants versus wild-type *difSL* (2ng/µL) under the same concentration of protein (352 nM), with (+) or without (-) protein.

Figure 3. A. Duplicate of titration of a 339 bp DNA fragment containing *difSL, difSL* mutants M8 and M9 with two concentrations of MBP-XerS (200 nM) and (530 nM) in an EMSA experiment (4% polyacrylamide/0.5x TBE gels). **B. Left**. Comparison of equal amounts of *difSL*-M9 half site mutants, left and right (M9-LHS and M9-RHS respectively) versus wild-type *difSL* (2ng/µL) under two different concentration of protein 200 nM and 530 nM, symbols (+) and (-) represent addition or not of protein. **Right**. Comparison of equal amounts of *difSL*-M8 half site mutants, left and right (M8-LHS and M8-RHS respectively) versus wild-type *difSL* (2ng/µL) under two different concentration of protein 200 nM and 530 nM, symbols (+) and (-) represent addition or not of protein.

Figure 4. XerS activity on suicide substrates. A. Schematic representation of the suicide substrates and cleavage reactions. Substrates consisted of three annealed oligonucleotides that

mimic the *difSL* sequence, either containing a top nick (TN) (Left) or a bottom nick (BN) (Right). The top-nicked (TN) substrate was 5' 6-HEX-labelled at the first nucleotide (Table 1) of the consensus bases on the top strand; the bottom-nicked substrate was 5'-6-HEX-labelled at the residue 31 (Table 1) on the bottom strand (Indicated by the *****). To assure a complete yield of fluorescently labeled suicide substrates, a three-fold excess of unlabeled fragments was used in the annealing reaction. 25µL of the reaction mixtures containing 8µM of TN or BN substrates was incubated with 350 nM of MBP-XerS for 60 min and heated to 95 °C for 5 min before loading on a 6% polyacrylamide gel run in 1% SDS in Tris-borate buffer. Center. Cleavage activity results in the formation of a stable covalent SDS-resistant complex only detectable by the interaction between MBP-XerS and the 5'HEX-labelled DNA under denaturing (1% SDS - 100°C for 5 minutes) conditions, as indicated by the dotted arrows (Left). Undetected covalent complexes are also indicated by these arrows (Right) under the presence of a catalytically defective XerS mutant ($XerS^{Y\text{-}F}$) and absence of the protein. **B.** Comparison of cleavage products between *difSL WT* and seven *difSL* mutants corresponding to modifications at the cleavage site, central region and binding site respectively, complete images with the free DNA and the controls have been omitted for clarity.

Figure 5. XerS activity on half suicide substrates. A. Half suicide substrates consisted of two annealed oligonucleotides, either corresponding to the left site or the right site. The left half site was 5' 6-HEX-labelled at the first nucleotide (A) (number 1, Table 1) of the consensus bases on the top strand; the right half site was 5'-6-HEX-labelled at the residue 31 (A) on the bottom strand. To assure a complete yield of fluorescently labeled suicide substrates, a threefold excess of unlabeled fragment was used in the annealing reaction. 25µL of the reaction mixtures containing 8µM of TN or BN substrates was incubated with 350 nM of MBP-XerS for 60 min and heated to 95 °C for 5 min before loading on a 6% polyacrylamide gel run in 1% SDS in Tris-borate buffer. Half site mutant oligonucleotides are labeled by a lowercase 'm' to distinguish them from the complete *difSL* mutant oligonucleotide.

Figure 6. Pull-down assay between immobilized MBP-XerS and His-FtsKγ. A. Alignments showing the level of conservation of interacting residues in FtsKγ and XerD/XerS of *E. coli (Ec), L. lactis* (*Ll*)

and *S. suis (Ss)*. Top: Alignment of the amino acid sequences of the FtsKγ domain of the indicated organisms in the region where interaction with XerD and XerS occurs. Structural features of the Ec FtsKγ is depicted: the helices (H1 to 3) and the wing of the w-helix DNA-binding domain. Identical residues across the 3 species are shown in bold, identical residues between Ss and Ec are show in red and important residues for KOPS binding are indicated by stars. The regions thought to interact with XerD and the KOPS motif are indicated, being the 9 most important ones underlined. Bottom: Alignments of the C-terminal portions of XerD from *E. coli* and XerSs from *L. lactis* and *S. suis*. White boxes denote amino acids seen to interact with FtsKγ in the crystal structure from *E. coli*, whereas grey boxes represent the conserved catalytic residues distinctive of the tyrosine recombinase family. Alignments are from NCBI. **B.** Analysis by SDS-PAGE of recombinase interactions (Duplicate). First gel (left): Lanes 1 and 2. **FtsK input** contains lysed extract of the over-expressed His-FtsKγ strain (Unknown concentration), Lane 3: **MBP-XerS+FtsKγ** contains the eluted fraction of sample carrying both enzymes (Bait and prey) , Lane 4: **MBP-XerS** contains the eluted fraction of sample carrying only the bait. Lane 5: **FtsK control** contains eluted fraction of sample with only the prey (FtsK binding control). Lane 6: **MBP-FtsK** contains the eluted fraction of sample with both proteins MBP tag and His-FtsKγ (MBP control). Lane 7 and 8: **MBP-XerS+FtsK+DNA** contains the eluted fraction of sample carrying both enzymes and a PCR fragment containing the *difSL* sequence (320 bp). Second gel (Right) is a replica with only three more lanes: Lane 4: **Washing step #5** that shows the flow-through of sample containing both enzymes and DNA after the fifth washing step. Lane 8 and 9 contains the respective tagged protein purified by affinity chromatography.

Figure 1. Titration of a 339 bp DNA fragment containing *difSL* wild type and *difSL* mutants (M1 to M7) by increasing concentrations of MBP-XerS.

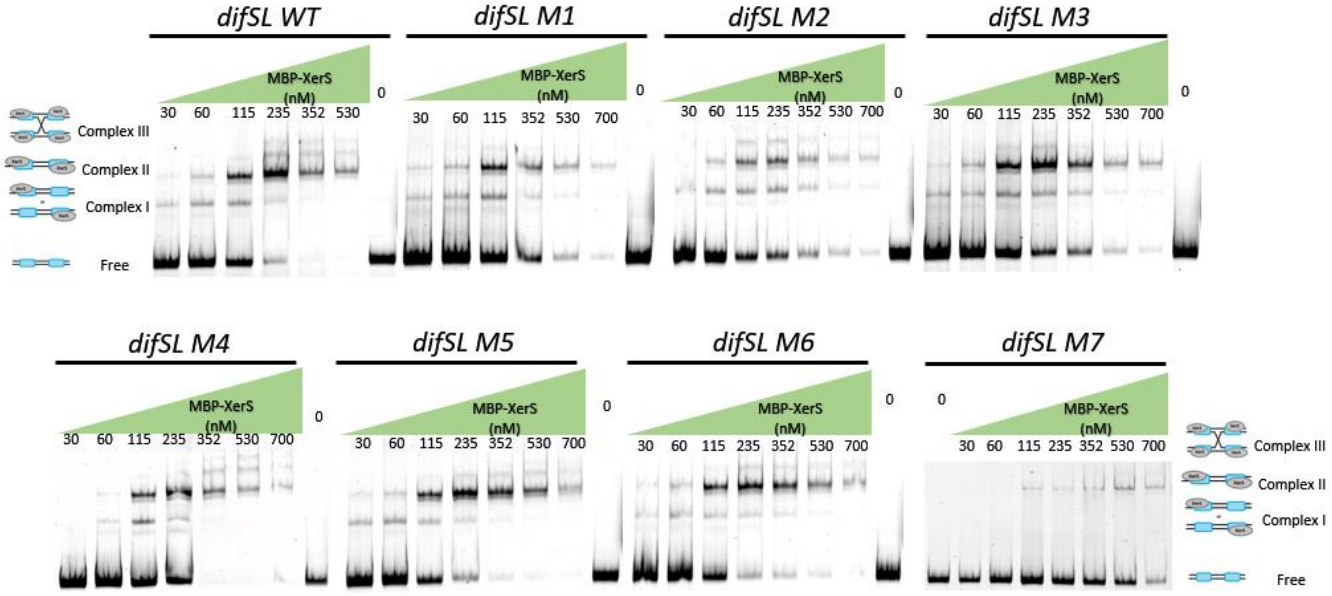


Figure 2. Titration of a 315 bp DNA fragment containing *difSL*-M7 half left site, M7-LHS **(A)** and half right site, M7-RHS **(C)**.

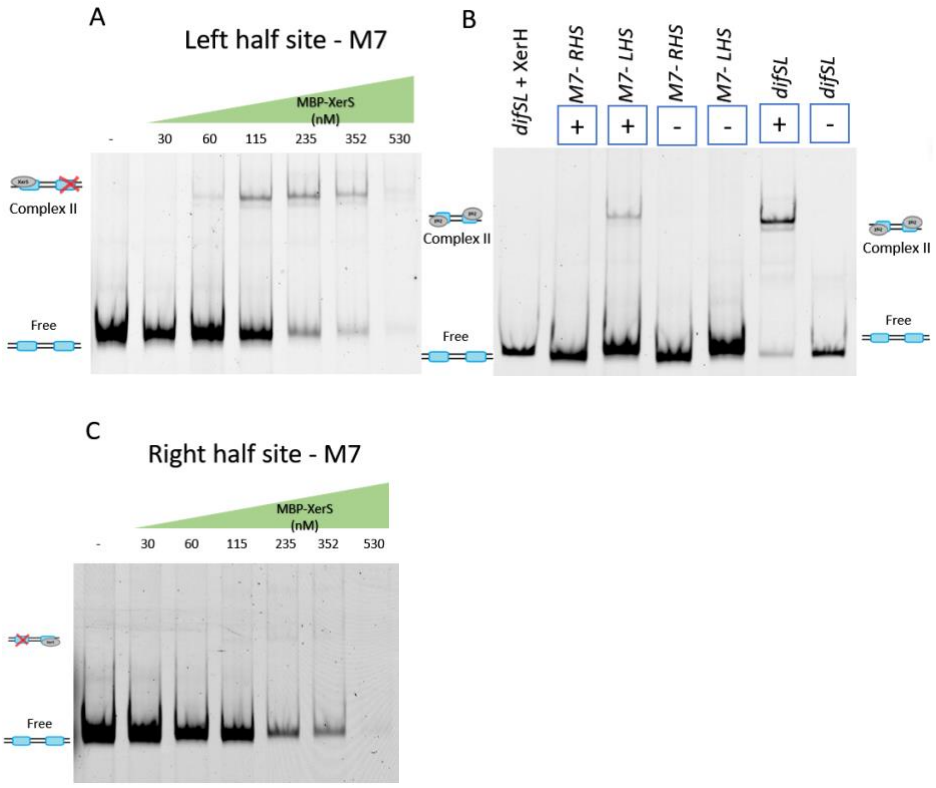


Figure 3. Titration of a 339 bp DNA fragment containing *difSL*, complete *difSL* mutants M8 and M9 and half sites.

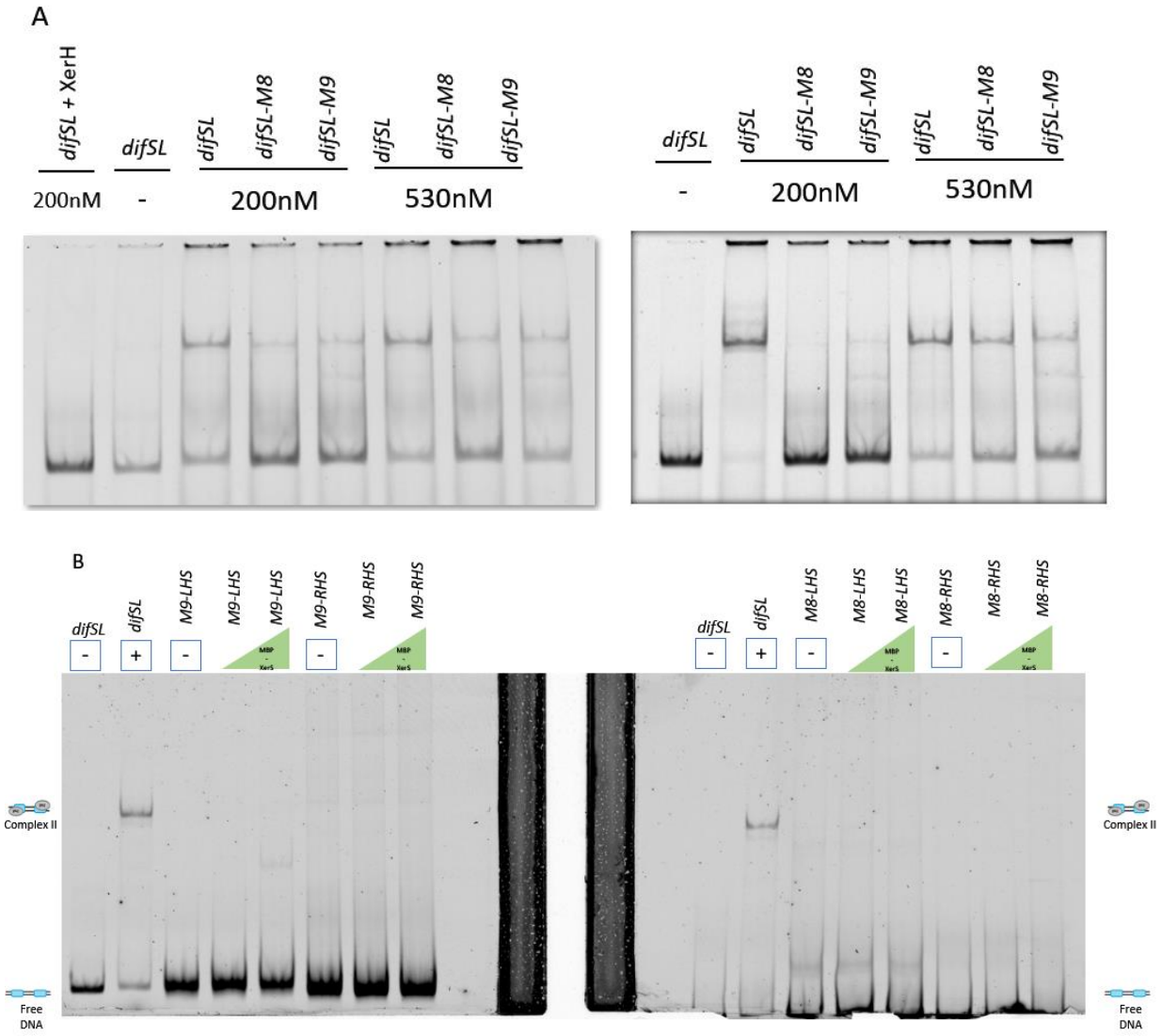


Figure 4. XerS activity on suicide substrates.

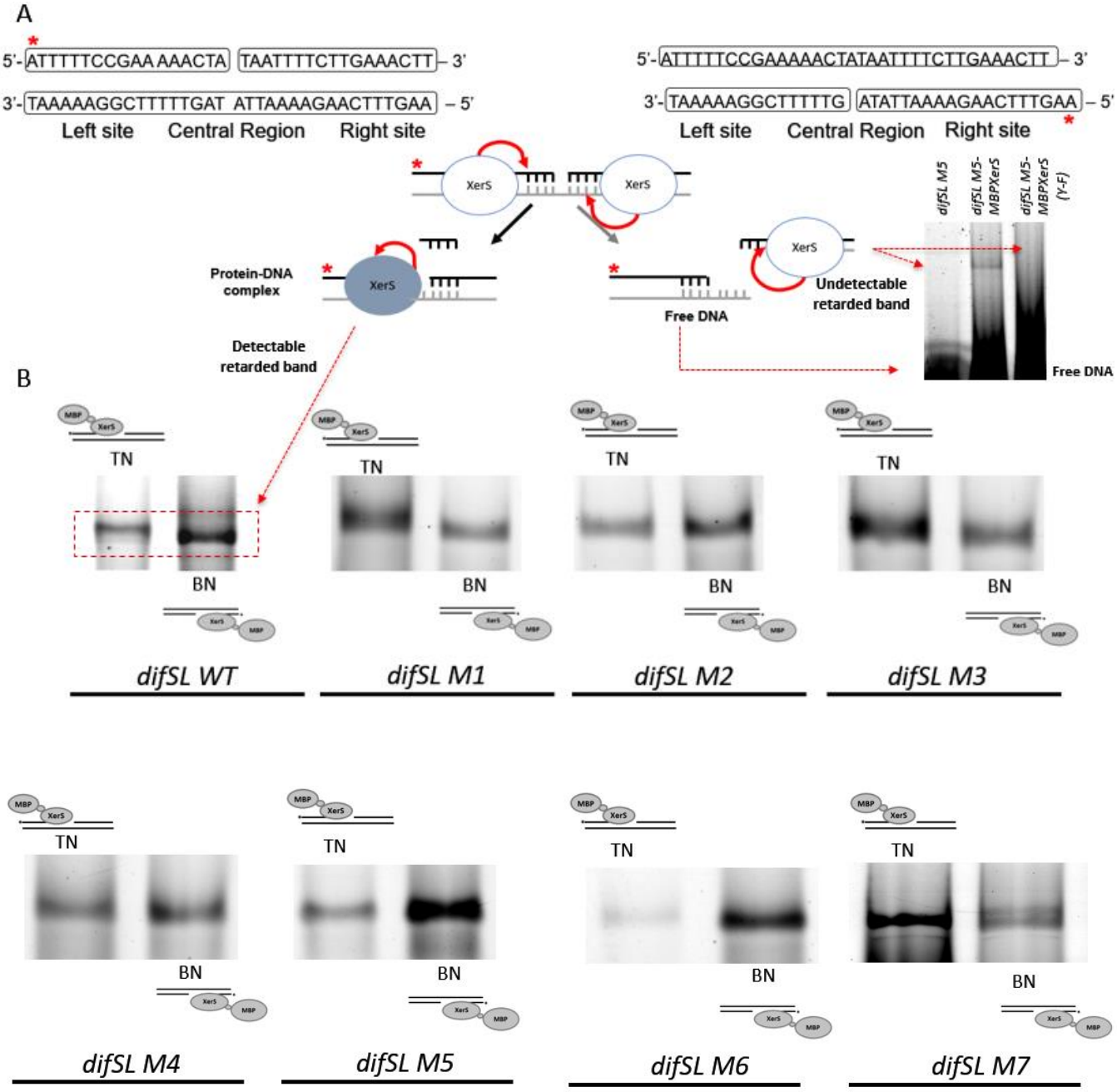


Figure 5. XerS activity on half suicide substrates.

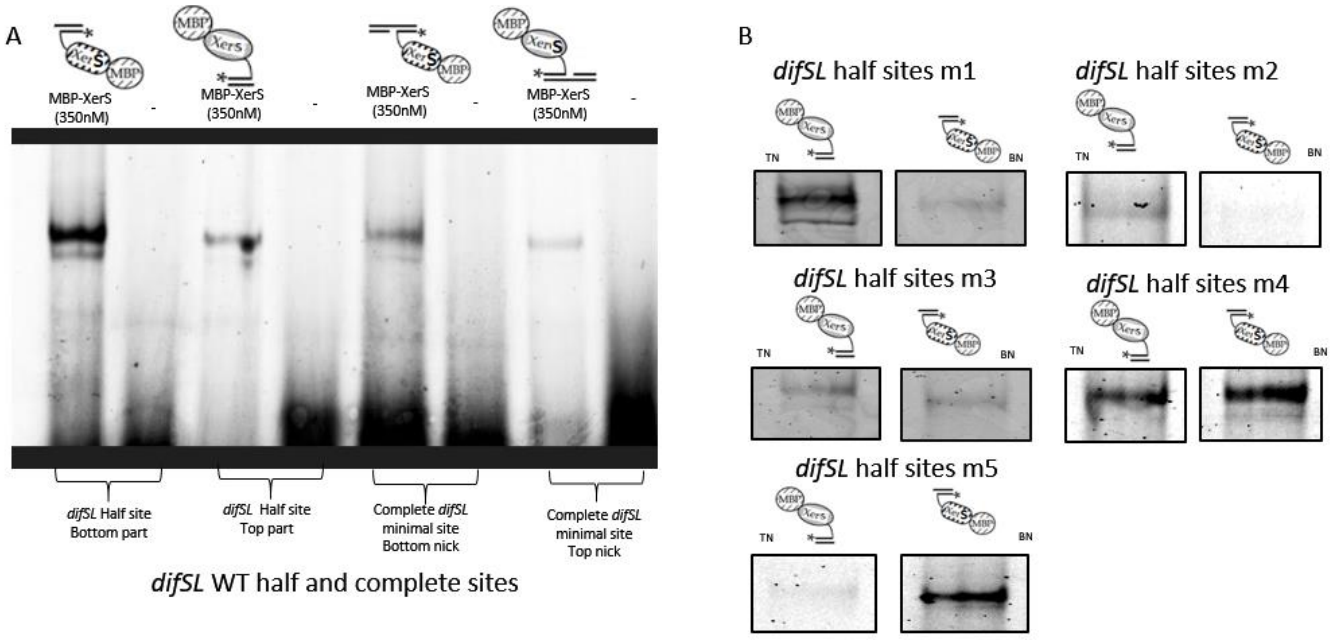
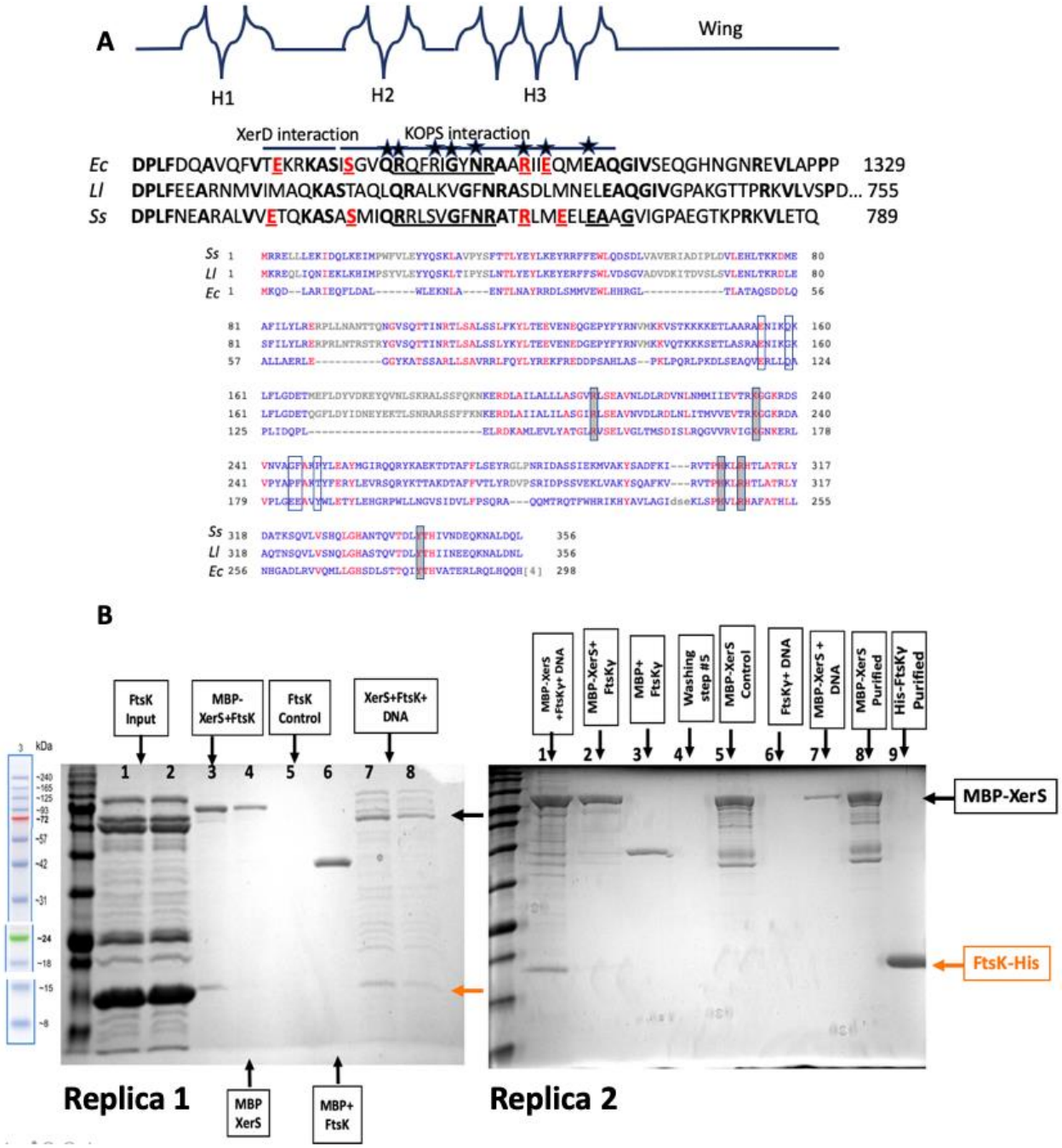


Figure 6. *Pull-down* assay between MBP-XerS and His-FtsKγ



Chapter III

General discussion

1. Results

Briefly, the experiments performed in this thesis have demonstrated that XerS binds specifically to *difSL* and that this binding is cooperative. Binding and cleavage reactions are asymmetric since XerS binds preferentially to the left half site and cleaves preferentially the bottom DNA strand. It has been previously demonstrated that the cleavage reaction is performed by the tyrosine residue (XerS position 341) directly involved in transesterification. Our experiments have also demonstrated that nucleotides at the outermost region of *difSL* seem to be determinant for binding selectivity rather than the inner ones, which show a high degree of binding impairment when mutated. Additionally, modifications of the extra spacing between the inverted repeat arms as well as reduction of the central region surprisingly change cleavage preference, which denotes the importance of the extra nucleotide on the right arm and the reason behind a larger central region within *difSL* when compared to other *dif* sites.

2. XerS/*difSL* system

The XerS/*difSL* system is a challenging system of study due to its uniqueness and similarities to most Xer/*dif* systems known so far. It is larger than most Xer recombinases with the exemption of XerH, it involves the largest *dif* site characterized so far and possibly, an asymmetric one with the right arm (12 bp) being larger than the left arm (11 bp) [50,177]. It also involves only one recombinase as do the Cre/*loxP* and Flp/FRT systems, it possesses a stronger binding preference for the left site over the right site as do XerC or CodV, and it preferentially exchanges the BSs of *difSL* as do XerD or RipX. Despite that few papers have been published on XerS and other single recombinases such as XerH or XerA, it is encouraging to see the recent discoveries done on molecular characterization of XerA and XerH by crystallography and other advanced molecular biology techniques [65,66]. This draws a promising parallel to analyze our data and formulate some

hypotheses. Do all single recombinases behave similarly despite structural differences in size and composition?

Our first results indicated that XerS binding is cooperative and it is produced in two stages, one subunit of XerS likely binds initially to the left half site of *difSL* and only then a second subunit of XerS binds to the right half site. Our EMSA experiments showed rapid formation of complex I under low concentrations of protein, and if it behaves like the XerH system, it is likely that binding to the right half site can only occur under the occupation of the left half site in which the dissociation constant is much lower than the right side [66]. These results led us to question how one subunit of XerS distinguishes between these two arms, considering that both arms are two palindromic sequences with relative symmetry (positions 4-7, 9-13 in the left arm and 21-25,28-31 in the right arm), apart from an extra nucleotide (position 26) in the middle of the right inverted repeat (TTTC**TT**GAAAA) versus the left part of the sequence (TTTC**C**GAAAA). The M7 mutation addressed this divergence by positioning the extra spacing from the right arm to the left arm (M7: TTTC**TT**GAAAA-CR-TTC**C**GAAAA). This mutation did decrease to some degree binding efficiency but did not alter binding preference, indicating that the extra nucleotide does not determine XerS stronger binding affinity for the left arm, however this mutation switched cleavage preference, favoring DNA exchange of the top strand over bottom strand, this proved that the extra spacing is indeed an additional control for the XerS/*difSL* system synchronization, once all arms are occupied by the corresponding XerS subunits. However, if the extra spacing does not directly influence XerS binding preference, it is still unclear what determines it.

In the XerH/*difH* system in *H. pylori*, XerH showed extensive direct contact with *difH* forming a tight complex in a C-shaped clamp form around each arm of the *difH* site. The majority of XerH interactions occur between helix αK and the outer part of *difH*, this result shed light on how XerS might interact with *difSL*. A similar characteristic found in the XerCD/*dif* system in *E. coli* and FLP/*FRT* system in *S. cerevisiae* with the critical nucleotides for binding being located at the outermost region of *dif* and *FRT* respectively [189,193]. Therefore, our reasoning was that the outer region of *difSL* might be determinant for XerS binding or even binding preference. Mutations M8 and M9 address this issue by modifying a conserved TTT rich region in the left arm of the *difSL* sequence (positions 2-5). Both mutations decreased XerS binding with the entire *difSL* region as

expected, but surprisingly, when we performed the same EMSA gels with their half parts, XerS binding was completely abolished. This proved that XerS requires this rich $TTT_{(2-5)}$ region at the outer part of *difSL* to bind more efficiently and agrees with previous results about the importance of the outer nucleotides for binding specificity [66,189]. Subsequently, we did place this rich TTT region on the right side to see if binding preference would shift to that arm (M10 & M11), unfortunately the experiments did not show any retarded band aside the free DNA (results not showed due to lack of repetitions). Further experiments must be done, however, if this is the case, other factors besides the TTT region must coordinate binding preference, either a combination of factors such as a specific spatial conformation or direct protein-DNA contacts. It is known that up to 6 amino acid residues of XerH alone interact with the equivalent outer region of *difH*, 6 residues of XerD with the outer region of the right side of *dif* and 8 residues of XerC with the outer region of the left side of *dif* [66,187] and that, up to one insertion or one point mutation greatly decrease Xer binding or cleavage on this region. Bebel and collaborators emphasized the importance of residue K290 of XerH which forms a specific hydrogen bond with thymine T4' whereas in the other arm this interaction is not possible due to a farther distance of the respective thymine T5'. Interestingly, the XerS/*difSL* system also possess a similar T4' nucleotide at the left side at the beginning of a highly conserved region $TTTC_{(4-7)}$ and a similar shifted thymine in the right arm, but in this case located at position T3', also in a highly conserved region $AAA_{(29-31,\ top\ strand)}$. Further experiments are necessary to determine if modifications of residue T3' to T4' do alter binding preference. It is necessary to highlight the importance of this result, since it is the first time it has been reported complete inhibition of XerS binding by a specific mutation. Suggesting that the main determinants for XerS binding are located at this region.

Another important characteristic of the *difSL* binding site is its high A/T concentration, especially more concentrated near the scissile phosphate, which facilitates bending of double-stranded DNA due to a narrow minor groove produced by these rich A/T regions [194]. However, modifications performed on these regions like M2 and M4, did not alter cleavage preference from the bottom strand to the top strand. On the contrary, adding one purine to the central region (M5 mutation) increased the intrinsic cleavage preference of XerS for the bottom strand, **(Figure 4)**. Our experiments to determine cleavage activity were performed by the use of suicide substrates as

described before, however, they must be interpreted with caution since they might not reflect real conditions nor the real structure within the bacterial genome under replication. This is because suicide substrates, due to their small length and level of freedom, may acquire different conformations which might facilitate XerS cleavage that is otherwise elusive [195]. In spite of this, the technique has proven to be reproducible (all experiments were performed by triplicate) and has been used and published in different papers [158,163,192,196]. Regarding our results, the exact reasons behind cleavage preference for the bottom strand are still unclear, mutations M1, M3 and M7 altered cleavage preference to the top strand, three mutations only related in structural modifications. Literature indicates that the correlation between DNA bending and cleavage activity, a critical factor for tyrosine recombinases and especially important for single recombinases is due to the fact that it is the level of asymmetry of the DNA that dictates cleavage preference rather than the protein structure itself [66]. Bebel's paper elucidated how *difH* asymmetry influences protein conformation on the synaptic complex, which leads to a different positioning of the segments (αN-αO) responsible for the nucleophilic attack. Therefore, it might be that our mutations reform protein conformation on the right arm making its subunit otherwise "inactive" more prone for DNA cleavage than the left arm, either by changing protein-DNA interactions or by facilitating bending of the DNA, consistent with the correlation between DNA bending and cleavage activity [56,197]. However, *difSL* asymmetry might not be the only regulatory factor for cleavage preference, especially considering the similitudes between the XerS/*difSL* system and XerH/*difH* and/or XerCD/*dif* systems, two other systems highly dependent on FtsK interaction for the obtention of productive recombinant products. Our pull-down results allowed us to infer that protein-protein interaction between XerS and FtsK indeed exists, as similarly reported by Nolivos and collaborators (2010), and that this interaction might coordinate a productive synapse between *difSL* sites as it occurs with XerD in the XerCD/*dif* system. Although further studies must be performed, it is possible to hypothesize that FtsK might prepare the top strand subunit to cleave the DNA once it comes into contact with the synaptic complex within the genome, considering the intrinsic preference of XerS for the bottom strand. If FtsK facilitates either bending of the DNA or further protein-protein interactions and/or modifications to an active conformation, holliday junction resolution will be favorized with subsequent obtention of

recombinant products [99]. Nolivos' and Yates' papers showed low homology between the residues involved in FtsK and XerD interaction and XerS and $FtsK_{sl}$ counterparts. Four of the five amino acids involved in XerD recognition of FtsK are conserved in these three systems, even though, Xer proteins are more divergent, only one of the five amino acids involved in FtsK recognition is conserved among these systems [2,159]. They proposed that FtsK/Xer protein interactions might have some degree of species specificity. It would be interesting to identify the determinants on XerS interaction with FtsK as well as using surface plasmon resonance studies to measure binding affinities between these two proteins. It is also important to understand how FtsK would play a regulatory role in single recombinase systems. In the XerCD/*dif* system, the presence of two different proteins plays an additional role on the stringency of the reaction, and it is the directionality/structure differences between XerC and XerD that determine proper recombinant outcome. For single recombinases, the reaction might follow a partial similar principle. Bebel and collaborators hypothesized that FtsK promotes rearrangement of the XerH/*difH* synaptic complex by performing two specific actions: bending of the DNA and promoting structure modification of XerH, in which the catalytic tyrosine is brought into the active position close to the cleavage site [54,66]. Therefore, it is possible that XerS subunits bound to the left arm, which present stronger binding affinity, are poised into an active cleavage state once FtsK interacts with these subunits. Bending of DNA (without the presence of FtsK) might be corroborated by previous experiments of footprinting analysis performed by Nolivos and collaborators which demonstrated both situations: protein-protein interactions between XerS subunits due to a zone of protection over the central region, and a slightly increase in DNAase I cleavage activity at each side of *difSL*, which denotes possible distortion caused by XerS binding onto *difSL*. These hypotheses with respect to cleavage activity and FtsK interaction are discussed here assuming a cis cleavage mode, but this has not yet been experimentally verified, although it is most likely. Additional experiments should be performed to validate this hypothesis. SPR analysis, and creation of chimera proteins of XerS plus γ-subunit of FtsK along holliday junction resolution experiments with either *difSL* wild type or half-site *difSL* mutants (mimicking both, the complete tetrameric complex or half of it) might answer important questions about the affinity between these two proteins and the regulatory

effect of FtsK over individual bound XerS subunits either to the left or right half sites, or the entire tetrameric complex.

3. Conclusions and perspectives

Our results allowed us to have a deeper understanding about XerS/*difSL* system, despite the requirement for more robust and extensive studies regarding not only our system but other single recombinase system for dimer resolution in bacteria. It is encouraging to study such proteins with its characteristic levels of specificity for DNA exchange, the commercial and investigative prospects are countless once the mechanistic insights will be completely understood, from genomic engineering, with wide range of applications in genome manipulation [198] to their use and fabrication by several biotechnological companies such as (NEB) with their product; CRE recombinase or System Biosciences, BIO (CA, USA) with their integrase; phiC31. Further studies involving plasmid integration efficiency as well as resolution of synthetic holliday junctions experiments must be performed to quantify binding and cleavage efficiencies. During our project we managed to create thermosensitive plasmids (pGhost9) capable of replication in *E. coli* but not in *S. suis* under non permissive conditions, harboring *difSL* wild type and *difSL* mutants. However, we couldn't obtain *S. suis recA*- mutants in order to perform integration efficiency experiments as described by Le Bourgeois' paper. Therefore, we encourage following groups to use our plasmids (pGhots line with mutations M1 to M6) if required, their modifications proved to be extremely difficult and we would be glad to know our work could save time and costs to other groups.

A second critical part to understand the dynamics of XerS dimer resolution is to elucidate the determinants for protein-protein interactions between either XerS-XerS or XerS-FtsK. A series of mutations, specially at the C-terminal region of XerS will tell us which residues are directly involved in protein-protein interactions, and it will help us to understand what determines such

species specificity for this system in some bacteria. To further provide more information *pull-down* assays and SPR experiments can be performed on the purified proteins as well as phenotypic characterization of *S. suis* Δ*xerS* strains harboring those mutants within broad host range plasmids. Initial experiments of SPR analysis were performed with the help of professor Jean-François Masson of the department of Biochemistry from the University of Montreal. However, the experiment required highly concentrated His-XerS samples with no traces of imidazole nor glycerol in them, and due to our insolubility problem with this fusion protein, we were not able to obtain the required concentration.

FRET experiments (*Fluorescence resonance energy transfer)* can also provide further information about the translocase activity of FtsK and its interaction with bound XerS proteins, within this technique, fluorophores will be attached to each protein, and the energy transfer will be measured as the variation of the emitted light by these fluorophores once interaction occurs (normally, manifested by quenching/reduction of the donor fluorophore and augmentation of the acceptor fluorescence emission) [100]. Our laboratory also intended to use different fluorescent tags (mcherry, yellow fluorescence protein (YFP) and green fluorescence protein (GFP)), plasmids kindly donated by Professor Reyes-Lamothe from the Department of Biology of McGill University, for studies of single molecule imaging. The idea would be to visualize both XerS and FtsK in real time and be capable to detect and measure times and levels of expression respectively, it would be also possible to pinpoint the exact localization of these proteins within *S. suis* in *in vivo* conditions. However, two important conditions must fit for this experiment. First, the level of autofluorescence caused by biological structures or molecules within *S. suis* must be low and secondly, the fluorescent tag must not interfere with normal cellular processes nor protein interactions. In parallel we performed some experiments to detect and isolate the genes that encode XerC, XerD and XerS from *Lactobacillus lactis* with the help of the student Sarah Tremblay. Further studies on these 3 recombinase systems will give us more information on how these systems may have evolved in bacteria, and reveal which, if any of these recombination systems is dispensable. In addition, this work could eventually lead to the discovery of new methods to integrate genes into the *Lactobacilli*, which can lead to the development of new strains for the dairy and probiotics industries. Initial experiments could involve the cloning and *in*

vitro characterization of all 3 recombinases and their sites, as well as knockouts of all three proteins to see if deleting them singly or in combination can have any effect on growth, division, and chromosome segregation.

Literature cited

1. Reyes-Lamothe, R., Nicolas, E. & Sherratt, D. J. Chromosome Replication and Segregation in Bacteria. *Annu. Rev. Genet.* **46**, 121–143 (2012).
2. Yates, J., Aroyo, M., Sherratt, D. J. & Barre, F.-X. Species specificity in the activation of Xer recombination at *dif* by FtsK. *Mol. Microbiol.* **49**, 241–249 (2003).
3. Grainge, I. *et al.* Unlinking chromosome catenanes *in vivo* by site-specific recombination. *EMBO J.* **26**, 4228–4238 (2007).
4. Barre, F. X. *et al.* Circles: the replication-recombination-chromosome segregation connection. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8189–95 (2001).
5. Massey, T. H., Aussel, L., Barre, F.-X. & Sherratt, D. J. Asymmetric activation of Xer site-specific recombination by FtsK. *EMBO Rep.* **5**, 399–404 (2004).
6. Austin, S., Ziese, M. & Sternberg, N. A novel role for site-specific recombination in maintenance of bacterial replicons. *Cell* **25**, 729–736 (1981).
7. Summers, D. K. & Sherratt, D. J. Multimerization of High Copy Number Plasmids Causes Instability : ColEI Encodes a Determinant Essential for Plasmid Monomerization and Stability. *Cell* **36**, 1097–1103 (1984).
8. Colloms, S. D., Sykora, P., Szatmari, G. & Sherratt, D. J. Recombination at ColEl *cer* Requires the *Escherichia coli xerC* Gene Product , a Member of the Lambda Integrase Family of Site-Specific Recombinases. *J. Bacteriol.* **172**, (1990).
9. Blakely, G. *et al.* Two Related Recombinases Are Required for Site-Specific Recombination at *dif* and *cer* in *E. coli* K12. *Cell* **75**, 351–361 (1993).
10. Jha, J. K., Ramachandran, R. & Chattoraj, D. K. Opening the Strands of Replication Origins—Still an Open Question. *Front. Mol. Biosci.* **3**, 1–13 (2016).

11. Pandey, M. *et al.* Two mechanisms coordinate replication termination by the *Escherichia coli* Tus–*Ter* complex. *Nucleic Acids Res.* **43**, 5924–5935 (2015).

12. Gowrishankar, J. End of the Beginning: Elongation and Termination Features of Alternative Modes of Chromosomal Replication Initiation in Bacteria. *PLoS Genet.* **11**, 1–9 (2015).

13. Elshenawy, M. M. *et al.* Replisome speed determines the efficiency of the Tus−*Ter* replication termination barrier. *Nature* **525**, 394–398 (2015).

14. Kuzminov, A. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* **63**, 751–813 (1999).

15. Ayora, S. *et al.* Double-strand break repair in bacteria: A view from *Bacillus subtilis*. *FEMS Microbiol. Rev.* **35**, 1055–1081 (2011).

16. Merrikh, H., Zhang, Y., Grossman, A. D. & Wang, J. D. Replication–transcription conflicts in bacteria. *Nat. Rev. Microbiol.* **10**, 449–458 (2012).

17. Tehranchi, A. K. *et al.* The transcription factor DksA prevents Disruption of DNA Replication upon Nutritional Stress. *Cell* **141**, 595–605 (2010).

18. Kisker, C., Kuper, J. & Houten, B. Van. Prokaryotic Nucleotide Excision Repair. *Cold Spring Harb Perspect Biol* **5**, 1–18 (2013).

19. Groothuizen, F. S. & Sixma, T. K. The conserved molecular machinery in DNA mismatch repair enzyme structures. *DNA Repair (Amst).* **38**, 14–23 (2015).

20. Lenhart, J. S., Schroeder, J. W., Walsh, B. W. & Simmons, L. a. DNA Repair and Genome Maintenance in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* **76**, 530–564 (2012).

21. van der Veen, S. & Tang, C. M. The BER necessities: the repair of DNA damage in human-adapted bacterial pathogens. *Nat. Rev. Microbiol.* **13**, 83–94 (2015).

22. Bruning, J. G., Howard, J. L. & McGlynn, P. Accessory replicative helicases and the replication of protein-bound DNA. *J. Mol. Biol.* **426**, 3917–3928 (2015).

23. Roghanian, M., Zenkin, N. & Yuzenkova, Y. Bacterial global regulators DksA/ppGpp increase fidelity of transcription. *Nucleic Acids Res.* **43**, 1529–36 (2015).

24. Pomerantz, R. T. & O'Donnell, M. What happens when replication and transcription complexes collide? *Cell Cycle* **9**, 2537–2543 (2010).

25. Yeeles, J. T. P., Poli, J., Marians, K. J. & Pasero, P. Rescuing stalled or damaged replication forks.

Cold Spring Harb. Perspect. Biol. **5**, 1–16 (2013).

26. Heller, R. C. & Marians, K. J. Replisome assembly and the direct restart of stalled replication forks. *Nat. Rev. Mol. Cell Biol.* **7**, 932–943 (2006).

27. Cox, M. M. *et al.* The importance of repairing stalled replication forks. *Nature* **404**, 37–41 (2000).

28. Vos, M. & Didelot, X. A comparison of homologous recombination rates in bacteria and archaea. *ISME J.* **3**, 199–208 (2009).

29. Darmon, E., Eykelenboom, J. K., Lopez-Vernaza, M. A., White, M. A. & Leach, D. R. F. Repair on the Go: *E. Coli* maintains a high proliferation rate while repairing a chronic DNA double-strand break. *PLoS One* **9**, 1–11 (2014).

30. Rocha, E. P. C., Cornet, E. & Michel, B. Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.* **1**, 0247–0259 (2005).

31. Costes, A. & Lambert, S. Homologous Recombination as a Replication Fork Escort: Fork-Protection and Recovery. *Biomolecules* **3**, 39–71 (2012).

32. Chen, Z., Yang, H. & Pavletich, N. P. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489–4 (2008).

33. Handa, N., Morimatsu, K., Lovett, S. T. & Kowalczykowski, S. C. Reconstitution of initial steps of dsDNA break repair by the RecF pathway of *E. coli*. *Genes Dev.* **23**, 1234–1245 (2009).

34. Zhang, J., Mahdi, A. a., Briggs, G. S. & Lloyd, R. G. Promoting and avoiding recombination: Contrasting activities of the *Escherichia coli* RuvABC Holliday junction resolvase and RecG DNA translocase. *Genetics* **185**, 23–37 (2010).

35. Cromie, G. a & Leach, D. R. Control of crossing over. *Mol. Cell* **6**, 815–826 (2000).

36. Lilley, D. M. J. & White, M. F. The junction-resolving enzymes. *Nat. Rev. Mol. cell Biol.* **2**, 433–443 (2001).

37. Steiner, W. W. & Kuempel, P. L. Sister Chromatid Exchange Frequencies in *Escherichia coli* Analyzed by Recombination at the *dif* Resolvase Site. *J. Bacteriol.* **180**, 6269–6275 (1998).

38. Yeeles, J. T. P. & Dillingham, M. S. The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes. *DNA Repair (Amst).* **9**, 276–285 (2010).

39. Đermić, D. Double-strand break repair mechanisms in *Escherichia coli*: recent insights. *Adv. Genomics Genet.* **5**, 35–42 (2015).

40. Azeroglu, B. *et al.* RecG Directs DNA Synthesis during Double-Strand Break Repair. *PLoS Genet.* **12**, 1–23 (2016).

41. Carrasco, B., Cozar, M. C., Lurz, R., Alonso, J. C. & Ayora, S. Genetic Recombination in *Bacillus subtilis* 168: Contribution of Holliday Junction Processing Functions in Chromosome segregation. *J. Bacteriol.* **186**, 5557–5566 (2004).

42. Martini, E., Diaz, R. L., Hunter, N. & Keeney, S. Crossover Homeostasis in Yeast Meiosis. *Cell* **126**, 285–295 (2006).

43. Labib, K. & Hodgson, B. Replication fork barriers: pausing for a break or stalling for time? *EMBO Rep.* **8**, 346–53 (2007).

44. Johnston, C. *et al.* RecFOR Is Not Required for Pneumococcal Transformation but Together with XerS for Resolution of Chromosome Dimers Frequently Formed in the Process. *PLoS Genet.* **11**, 1–13 (2015).

45. Van Gool, A. J., Hajibagheri, N. M. a, Stasiak, A. & West, S. C. Assembly of the *Escherichia coli* RuvABC resolvasome directs the orientation of Holliday junction resolution. *Genes Dev.* **13**, 1861–1870 (1999).

46. Spies, M. & Stephen C. Kowalczykowski. RecBCD and RecF Pathways. in *The bacterial Chromosome* (ed. N. Patrick Higgins) 389–403 (ASM press, 2005).

47. Pagès, V. Single-strand gap repair involves both RecF and RecBCD pathways. *Curr. Genet.* **62**, 519–521 (2016).

48. Michel, B., Recchia, G. D., Penel-Colin, M., Ehrlich, S. D. & Sherratt, D. J. Resolution of Holliday junctions by RuvABC prevents dimer formation in *rep* mutants and UV-irradiated cells. *Mol. Microbiol.* **37**, 180–191 (2000).

49. Fijalkowska, I. J., Schaaper, R. M. & Jonczyk, P. DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair. *FEMS Microbiol. Rev.* **36**, 1105–1121 (2012).

50. Carnoy, C. & Roten, C.-A. The *dif*/Xer recombination systems in proteobacteria. *PLoS One* **4**, e6531 (2009).

51. Mazin, Ma. V., Timchenko, T. V., Saparbaev, M. K. & Mazina, O. M. Dimerization of plasmid DNA accelerates selection for antibiotic resistance. *Mol. Microbiol.* **20**, 101–108 (1996).

52. Berza, I., Dishlers, A., Petrovskis, I., Tars, K. & Kazaks, A. Plasmid dimerization increases the

production of hepatitis B core particles in *E. coli*. *Biotechnol. Bioprocess Eng.* **18**, 850–857 (2013).

53. Warren, D., Laxmikanthan, G. & Landy, A. Integrase Family of Site-Specific Recombinases. *Brenner's Encycl. Genet.* **4**, 100–105 (2013).

54. Diagne, C. T. *et al.* TPM analyses reveal that FtsK contributes both to the assembly and the activation of the XerCD-*dif* recombination synapse. *Nucleic Acids Res.* **42**, 1721–1732 (2014).

55. Kono, N., Arakawa, K. & Tomita, M. Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. *BMC Genomics* **12**, 1–14 (2011).

56. Grindley, N. D. F., Whiteson, K. L. & Rice, P. a. Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* **75**, 567–605 (2006).

57. Hirano, N., Muroi, T., Takahashi, H. & Haruki, M. Site-specific recombinases as tools for heterologous gene integration. *Appl. Microbiol. Biotechnol.* **92**, 227–39 (2011).

58. Brown, W. R. a, Lee, N. C. O., Xu, Z. & Smith, M. C. M. Serine recombinases as tools for genome engineering. *Methods* **53**, 372–379 (2011).

59. Stirling, C. J., Stewart, G. & Sherratt, D. J. Multicopy plasmid stability in *Escherichia coli*. *Mol. Gen. Genet.* **214**, 80–84 (1988).

60. Sherratt, D. J. *et al.* Recombination and chromosome segregation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**, 61–9 (2004).

61. Sherratt, D. J. *et al.* Site-specific recombination and circular chromosome segregation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **347**, 37–42 (1995).

62. Swalla, B. M., Gumport, R. I. & Gardner, J. F. Conservation of structure and function among tyrosine recombinases: Homology-based modeling of the lambda integrase core-binding domain. *Nucleic Acids Res.* **31**, 805–818 (2003).

63. Meinke, G., Bohm, A., Hauber, J., Pisabarro, M. T. & Buchholz, F. Cre Recombinase and Other Tyrosine Recombinases. *Chem. Rev.* **116**, 12785–12820 (2016).

64. Subramanya, H. S. *et al.* Crystal structure of the site-specific recombinase, XerD. *EMBO J.* **16**, 5178–87 (1997).

65. Hwa Jo, C. *et al.* Crystal structure of *Thermoplasma acidophilum* XerA recombinase shows large C-shape clamp conformation and cis-cleavage mode for nucleophilic tyrosine. *FEBS Lett.* **590**, 848–856 (2016).

66. Bebel, A., Karaca, E., Kumar, B., Stark, W. M. & Barabas, O. Structural snapshots of Xer recombination reveal activation by synaptic complex remodeling and DNA bending. *Elife* 1–23 (2016). doi:10.7554/eLife.19706

67. Gopaul, D. N. & Duyne, G. D. Van. Structure and mechanism in site-specific recombination. *Curr. Opin. Struct. Biol.* **9**, 14–20 (1999).

68. Martin, S. S., Pulido, E., Chu, V. C., Lechner, T. S. & Baldwin, E. P. The order of strand exchanges in Cre-LoxP recombination and its basis suggested by the crystal structure of a Cre-LoxP holliday junction complex. *J. Mol. Biol.* **319**, 107–127 (2002).

69. Hickman, a B., Waninger, S., Scocca, J. J. & Dyda, F. Molecular organization in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 A resolution. *Cell* **89**, 227–37 (1997).

70. Chen, Y., Narendra, U., Iype, L. E., Cox, M. M. & Rice, P. a. Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol. Cell* **6**, 885–897 (2000).

71. Biswas, T. *et al.* A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature* **435**, 1059–1066 (2005).

72. Kwon, H. J., Tirumalai, R., Landy, A. & Ellenberger, T. Flexibility in DNA Recombination: Structure of the Lambda Integrase Catalytic Core. *Science (80-.).* **276**, 126–131 (1997).

73. Kuempel, P., Henson, J., Dircks, L., Tecklenburg, M. & Lim, D. *dif*, a recA-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. *New Biol* **3**, 799–811 (1991).

74. Rajeev, L., Malanowska, K. & Gardner, J. F. Challenging a paradigm: the role of DNA homology in tyrosine recombinase reactions. *Microbiol. Mol. Biol. Rev.* **73**, 300–309 (2009).

75. Vanhooff, V., Normand, C., Galloy, C., Segall, A. M. & Hallet, B. Control of directionality in the DNA strand-exchange reaction catalysed by the tyrosine recombinase TnpI. *Nucleic Acids Res.* **38**, 2044–2056 (2009).

76. Hallet, B., Arciszewska, L. K. & Sherratt, D. J. Reciprocal control of catalysis by the tyrosine recombinases XerC and XerD: an enzymatic switch in site-specific recombination. *Mol. Cell* **4**, 949–959 (1999).

77. Aussel, L. *et al.* FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* **108**, 195–205 (2002).

78. Zawadzki, P. *et al.* Conformational transitions during FtsK translocase activation of individual XerCD–*dif* recombination complexes. *PNAS* **110**, 17302–17307 (2013).

79. Midonet, C., Das, B., Paly, E. & Barre, F.-X. XerD-mediated FtsK-independent integration of TLCφ into the Vibrio cholerae genome. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16848–16853 (2014).

80. Recchia, G. D., Aroyo, M., Wolf, D., Blakely, G. & Sherratt, D. J. FtsK-dependent and -independent pathways of Xer site-specific recombination. *EMBO J.* **18**, 5724–5734 (1999).

81. Pérals, K., Cornet, F., Merlet, Y., Delon, I. & Louarn, J. M. Functional polarization of the *Escherichia coli* chromosome terminus: The *dif* site acts in chromosome dimer resolution only when located between long stretches of opposite polarity. *Mol. Microbiol.* **36**, 33–43 (2000).

82. Nolivos, S. *et al.* Co-evolution of segregation guide DNA motifs and the FtsK translocase in bacteria: Identification of the atypical *Lactococcus lactis* KOPS motif. *Nucleic Acids Res.* **40**, 5535–5545 (2012).

83. Besprozvannaya, M. & Burton, B. M. Do the same traffic rules apply? Directional chromosome segregation by SpoIIIE and FtsK. *Mol. Microbiol.* **93**, 599–608 (2014).

84. Löwe, J. *et al.* Molecular Mechanism of Sequence-Directed DNA Loading and Translocation by FtsK. *Mol. Cell* **31**, 498–509 (2008).

85. Grainge, I., Lesterlin, C. & Sherratt, D. J. Activation of XerCD-*dif* recombination by the FtsK DNA translocase. *Nucleic Acids Res.* **39**, 5140–5148 (2011).

86. Graham, J. E., Sivanathan, V., Sherratt, D. J. & Arciszewska, L. K. FtsK translocation on DNA stops at XerCD-*dif*. *Nucleic Acids Res.* **38**, 72–81 (2009).

87. Bisicchia, P., Steel, B., Debela, M. H. M., Löwe, J. & Sherratt, D. The N-Terminal Membrane-Spanning Domain of the *Escherichia coli* DNA Translocase FtsK Hexamerizes at Midcell. *MBio* **4**, 1–6 (2013).

88. Yu, X. C., Weihe, E. K. & Margolin, W. Role of the C terminus of FtsK in *Escherichia coli* chromosome segregation. *J. Bacteriol.* **180**, 6424–6428 (1998).

89. Steiner, W., Liu, G., Donachie, W. D. & Kuempel, P. The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers. *Mol. Microbiol.* **31**, 579–583 (1999).

90. Berezuk, A. M., Goodyear, M. & Khursigara, C. M. Site-directed fluorescence labeling reveals a revised N-terminal membrane topology and functional periplasmic residues in the *Escherichia coli* cell division protein FtsK. *J. Biol. Chem.* **289**, 23287–23301 (2014).

91. Dubarry, N., Possoz, C. & Barre, F. X. Multiple regions along the *Escherichia coli* FtsK protein are implicated in cell division. *Mol. Microbiol.* **78**, 1088–1100 (2010).

92. Val, M.-E. *et al.* FtsK-dependent dimer resolution on multiple chromosomes in the pathogen *Vibrio cholerae*. *PLoS Genet.* **4**, 1–11 (2008).

93. Massey, T. H., Mercogliano, C. P., Yates, J., Sherratt, D. J. & Löwe, J. Double-Stranded DNA Translocation: Structure and Mechanism of Hexameric FtsK. *Mol. Cell* **23**, 457–469 (2006).

94. Bigot, S., Corre, J., Louarn, J. M., Cornet, F. & Barre, F. X. FtsK activities in Xer recombination, DNA mobilization and cell division involve overlapping and separate domains of the protein. *Mol. Microbiol.* **54**, 876–886 (2004).

95. Bigot, S., Sivanathan, V., Possoz, C., Barre, F. X. & Cornet, F. FtsK, a literate chromosome segregation machine. *Mol. Microbiol.* **64**, 1434–1441 (2007).

96. Demarre, G., Galli, E. & Barre, F.-X. The FtsK Family of DNA Pumps. in *DNA Helicases and DNA Motor Proteins, Advances in experimental* 245–262 (2013). doi:10.1007/978-1-4614-5037-5

97. Sivanathan, V. *et al.* The FtsK γ domain directs oriented DNA translocation by interacting with KOPS. *Nat. Struct. Mol. Biol.* **13**, 965–972 (2006).

98. Lee, J. Y., Finkelstein, I. J., Arciszewska, L. K., Sherratt, D. J. & Greene, E. C. Single-molecule imaging of FtsK translocation reveals mechanistic features of protein-protein collisions on DNA. *Mol. Cell* **54**, 832–43 (2014).

99. Keller, A. N. *et al.* Activation of Xer-recombination at *dif*: structural basis of the FtsKγ–XerD interaction. *Sci. Rep.* **6**, 1–12 (2016).

100. May, P. F. J., Zawadzki, P., Sherratt, D. J., Kapanidis, A. N. & Arciszewska, L. K. Assembly, translocation, and activation of XerCD-*dif* recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5133–5141 (2015).

101. Graham, J. E., Sherratt, D. J. & Szczelkun, M. D. Sequence-specific assembly of FtsK hexamers establishes directional translocation on DNA. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20263–8 (2010).

102. Crozat, E. & Grainge, I. FtsK DNA Translocase: The Fast Motor That Knows Where It's Going. *ChemBioChem* **11**, 2232–2243 (2010).

103. Lee, J. Y., Finkelstein, I. J., Crozat, E., Sherratt, D. J. & Greene, E. C. Single-molecule imaging of DNA curtains reveals mechanisms of KOPS sequence targeting by the DNA translocase FtsK. *Proc. Natl. Acad. Sci.* **109**, 6531–6536 (2012).

104. Saleh, O. a, Bigot, S., Barre, F.-X. & Allemand, J.-F. Analysis of DNA supercoil induction by FtsK indicates translocation without groove-tracking. *Nat. Struct. Mol. Biol.* **12**, 436–440 (2005).

105. Summers, D. K., Beton, C. W. & Withers, H. L. Multicopy plasmid instability: the dimer catastrophe hypothesis. *Mol. Microbiol.* **8**, 1031–8 (1993).

106. Field, C. M. & Summers, D. K. Multicopy plasmid stability: Revisiting the dimer catastrophe. *J. Theor. Biol.* **291**, 119–127 (2011).

107. Million-Weaver, S. & Camps, M. Mechanisms of plasmid segregation: have multicopy plasmids been overlooked? *Plasmid* **75**, 27–36 (2014).

108. Crozat, E., Fournes, F., Cornet, F., Hallet, B. & Rousseau, P. Resolution of Multimeric Forms of Circular Plasmids and Chromosomes. *Microbiol. Spectr.* **2**, 1–16 (2014).

109. Sengupta, M. & Austin, S. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. *Infect. Immun.* **79**, 2502–2509 (2011).

110. Stirling, C. J., Szatmari, G., Stewart, G., Smith, M. C. M. & Sherratt, D. J. The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the ColE1 *cer* locus. *EMBO J.* **7**, 4389–4395 (1988).

111. Stirling, C. J., Colloms, S. D., Collins, J. F., Szatmari, G. & Sherratt, D. J. *xerB*, an *Escherichia coli* gene required for plasmid ColE1 site-specific recombination , is identical to *pepA*, encoding aminopeptidase A , a protein with substantial similarity to bovine lens leucine aminopeptidase. *EMBO J.* **8**, 1623–1627 (1989).

112. Cornet, F., Hallet, B. & Sherratt, D. J. Xer Recombination in *Escherichia coli*. *J. Biol. Chem.* **272**, 21927–21931 (1997).

113. Colloms, S. D., McCulloch, R., Grant, K., Neilson, L. & Sherratt, D. J. Xer-mediated site-specific recombination in vitro. *EMBO J.* **15**, 1172–81 (1996).

114. Colloms, S. D. The topology of plasmid-monomerizing Xer site-specific recombination. *Biochem.*

Soc. Trans. **41**, 589–94 (2013).

115. Reijns, M., Lu, Y., Leach, S. & Colloms, S. D. Mutagenesis of PepA suggests a new model for the *Xer/cer* synaptic complex. *Mol. Microbiol.* **57**, 927–41 (2005).

116. Sénéchal, H., Delesques, J. & Szatmari, G. *Escherichia coli* ArgR mutants defective in *cer*/Xer recombination, but not in DNA binding. *FEMS Microbiol. Lett.* **305**, 162–9 (2010).

117. Cornet, F., Mortier, I., Patte, J. & Louarn, J. M. Plasmid pSC101 harbors a recombination site, *psi*, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site *dif*. *J. Bacteriol.* **176**, 3188–95 (1994).

118. Blaby, I. K. & Summers, D. K. The role of FIS in the Rcd checkpoint and stable maintenance of plasmid ColE1. *Microbiology* **155**, 2676–2682 (2009).

119. Wang, X., Llopis, P. M. & Rudner, D. Z. Organization and segregation of bacterial chromosomes. *Nat. Rev. Genet.* **14**, 1–22 (2013).

120. Sciochetti, S. a., Piggot, P. J., Sherratt, D. J. & Blakely, G. The ripX locus of *Bacillus subtilis* encodes a site-specific recombinase involved in proper chromosome partitioning. *J. Bacteriol.* **181**, 6053–6062 (1999).

121. Sciochetti, S. A., Piggot, P. J. & Garry, W. Identification and Characterization of the *dif* site from *Bacillus subtilis*. *J. Bacteriol.* **183**, 1058–1068 (2001).

122. Barre, F. X. FtsK and SpoIIIE: The tale of the conserved tails. *Mol. Microbiol.* **66**, 1051–1055 (2007).

123. Kaimer, C., Schenk, K. & Graumann, P. L. Two DNA translocases synergistically affect chromosome dimer resolution in *Bacillus subtilis*. *J. Bacteriol.* **193**, 1334–1340 (2011).

124. Wu, L. J. It takes two DNA translocases to untangle chromosomes from the division septum. *Mol. Microbiol.* **74**, 773–776 (2009).

125. Fiche, J. B. *et al.* Recruitment, Assembly, and Molecular Architecture of the SpoIIIE DNA Pump Revealed by Superresolution Microscopy. *PLoS Biol.* **11**, 1–20 (2013).

126. Cattoni, D. I. *et al.* Structure and DNA-binding properties of the *Bacillus subtilis* SpoIIIE DNA translocase revealed by single-molecule and electron microscopies. *Nucleic Acids Res.* **42**, 2624–2636 (2014).

127. Kaimer, C., González-Pastor, J. E. & Graumann, P. L. SpoIIIE and a novel type of DNA translocase, SftA, couple chromosome segregation with cell division in *Bacillus subtilis*. *Mol. Microbiol.* **74**,

810–825 (2009).

128. Kaimer, C. & Graumann, P. L. Players between the worlds: Multifunctional DNA translocases. *Curr. Opin. Microbiol.* **14**, 719–725 (2011).

129. Biller, S. J. & Burkholder, W. F. The *Bacillus subtilis* SftA (YtpS) and SpoIIIe DNA translocases play distinct roles in growing cells to ensure faithful chromosome partitioning. *Mol. Microbiol.* **74**, 790–809 (2009).

130. Besprozvannaya, M., Pivorunas, V. L., Feldman, Z. & Burton, B. M. SpoIIIE protein achieves directional DNA translocation through allosteric regulation of ATPase activity by an accessory domain. *J. Biol. Chem.* **288**, 28962–28974 (2013).

131. Bose, B., Reed, S. E., Besprozvannaya, M. & Burton, B. M. Missense mutations allow a sequence-blind mutant of spoIIIE to successfully translocate chromosomes during sporulation. *PLoS One* **11**, (2016).

132. Yen Shin, J. *et al.* Visualization and functional dissection of coaxial paired SpoIIIE channels across the sporulation septum. *Elife* **4**, 1–24 (2015).

133. Burton, B. M., Marquis, K. A., Sullivan, N. L., Rapoport, T. A. & Rudner, D. Z. The ATPase SpoIIIE transports DNA across fused septal membranes during sporulation in *Bacillus subtilis*. *Cell* **131**, 1301–1312 (2010).

134. Veiga, H. & Pinho, M. *Staphylococcus aureus* requires at least one FtsK/SpoIIIE protein for correct chromosome segregation. *Mol. Microbiol.* **103**, 504–517 (2016).

135. Jha, J. K., Baek, J. H., Venkova-Canova, T. & Chattoraj, D. K. Chromosome dynamics in multichromosome bacteria. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819**, 826–829 (2012).

136. Val, M.-E. *et al.* A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.* **2**, 1–14 (2016).

137. Xu, Q., Dziejman, M. & Mekalanos, J. J. Determination of the transcriptome of *Vibrio cholerae* during intraintestinal growth and midexponential phase in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1286–1291 (2003).

138. Kirkup, B. C., Chang, L., Chang, S., Gevers, D. & Polz, M. F. *Vibrio* chromosomes share common history. *BMC Microbiol.* **10**, 1–13 (2010).

139. Harrison, P. W., Lower, R. P. J., Kim, N. K. D. & Young, J. P. W. Introducing the bacterial 'chromid':

Not a chromosome, not a plasmid. *Trends Microbiol.* **18**, 141–148 (2010).

140. McLeod, S. M. & Waldor, M. K. Characterization of XerC- and XerD-dependent CTX phage integration in *Vibrio cholerae*. *Mol. Microbiol.* **54**, 935–947 (2004).

141. Huber, K. E. & Waldor, M. K. Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* **417**, 656–659 (2002).

142. Demarre, G. *et al.* Differential Management of the Replication Terminus Regions of the Two *Vibrio cholerae* Chromosomes during Cell Division. *PLoS Genet.* **10**, 1–10 (2014).

143. Das, B. Mechanistic insights into filamentous phage integration in *Vibrio cholerae*. *Front. Microbiol.* **5**, 1–9 (2014).

144. Das, B., Martínez, E., Midonet, C. & Barre, F.-X. Integrative mobile elements exploiting Xer recombination. *Trends Microbiol.* **21**, 23–30 (2013).

145. Bischerour, J., Spangenberg, C. & Barre, F.-X. Holliday junction affinity of the base excision repair factor Endo III contributes to cholera toxin phage integration. *EMBO J.* **31**, 3757–67 (2012).

146. Val, M. E. *et al.* The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. *Mol. Cell* **19**, 559–566 (2005).

147. Hassan, F., Kamruzzaman, M., Mekalanos, J. J. & Faruque, S. M. Satellite phage TLCϕ enables toxigenic conversion by CTX phage through *dif* site alteration. *Nature* **467**, 982–5 (2010).

148. Faruque, S. M. & Mekalanos, J. J. Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence* **3**, 556–65 (2012).

149. Midonet, C. & Barre, F.-X. How Xer-exploiting mobile elements overcome cellular control. *Proc. Natl. Acad. Sci.* **113**, 8343–8345 (2016).

150. Fournes, F. *et al.* FtsK translocation permits discrimination between an endogenous and an imported Xer/*dif* recombination complex. *Proc. Natl. Acad. Sci.* **113**, 7882–7887 (2016).

151. Domínguez, N. M., Hackett, K. T. & Dillard, J. P. XerCD-mediated site-specific recombination leads to loss of the 57-kilobase gonococcal genetic Island. *J. Bacteriol.* **193**, 377–388 (2011).

152. Antonelli, A. *et al.* Characterization of a Novel Putative Xer-Dependent Integrative Mobile Element Carrying the blaNMC-A Carbapenemase Gene, Inserted into the Chromosome of Members of the *Enterobacter cloacae* Complex. *Antimicrob. Agents Chemother.* **59**, 6620–6624 (2015).

153. Christie, P. J., Whitaker, N. & González-Rivera, C. Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta* **1843**, 1578–1591 (2014).

154. Wu, H., Zhang, Z., Hu, S. & Yu, J. On the molecular mechanism of GC content variation among eubacterial genomes. *Biol. Direct* **7**, 1–16 (2012).

155. Recchia, G. D. & Sherratt, D. J. Conservation of xer site specific recombination genes in bacteria. *Mol. Microbiol.* **34**, 1146–1148 (1999).

156. Debowski, A. W. *et al.* Xer recombinase and genome integrity in *Helicobacter pylori*, a pathogen without topoisomerase IV. *PLoS One* **7**, 1–15 (2012).

157. Le Bourgeois, P. *et al.* The unconventional Xer recombination machinery of Streptococci/Lactococci. *PLoS Genet.* **3**, 1190–1199 (2007).

158. Leroux, M., Jia, F. & Szatmari, G. Characterization of the *Streptococcus suis* XerS recombinase and its unconventional cleavage of the *difSL* site. *FEMS Microbiol. Lett.* **324**, 135–41 (2011).

159. Nolivos, S., Pages, C., Rousseau, P., Le Bourgeois, P. & Cornet, F. Are two better than one? Analysis of an FtsK/Xer recombination system that uses a single recombinase. *Nucleic Acids Res.* **38**, 6477–89 (2010).

160. Hendrickson, H. & Lawrence, J. G. Selection for chromosome architecture in bacteria. *J. Mol. Evol.* **62**, 615–629 (2006).

161. Leroux, M., Rezoug, Z. & Szatmari, G. The Xer/*dif* site-specific recombination system of *Campylobacter jejuni* . *Mol. Genet. Genomics* **288**, 495–502 (2013).

162. Cortez, D. *et al.* Evidence for a Xer/*dif* system for chromosome resolution in Archaea. *PLoS Genet.* **6**, e1001166 (2010).

163. Serre, M. C. *et al.* The Carboxy-Terminal αN Helix of the Archaeal XerA Tyrosine Recombinase Is a Molecular Switch to Control Site-Specific Recombination. *PLoS One* **8**, 1–13 (2013).

164. Marians, K. J. Lesion Bypass and the Reactivation of Stalled Replication Forks. *Annu. Rev. Biochem.* **87**, 1–22 (2018).

165. Yeeles, J. T. P. & Marians, K. J. Dynamics of leading-strand lesion skipping by the replisome. *Mol. Cell* **52**, 855–865 (2013).

166. Laureti, L., Demol, J., Fuchs, R. P. & Pagés, V. Bacterial Proliferation: Keep Dividing and Don't Mind the Gap. *PLoS Genet.* **11**, 1–16 (2015).

167. Marczynski, G. T., Rolain, T. & Taylor, J. A. Redefining bacterial origins of replication as centralized information processors. *Front. Microbiol.* **6**, 1–10 (2015).

168. Windgassen, T. A., Wessel, S. R., Bhattacharyya, B. & Keck, J. L. Mechanisms of bacterial DNA replication restart. *Nucleic Acids Res.* **46**, 504–519 (2018).

169. Heller, R. C. & Marians, K. J. Replication fork reactivation downstream of a blocked nascent leading strand. *Nature* **439**, 557–562 (2006).

170. Cui, T. *et al.* Escherichia coli with a linear genome. *EMBO Rep.* **8**, 181–187 (2007).

171. Cornet, F., Louarn, J., Patte, J. & Louarn, J. M. Restriction of the activity of the recombination site dif to a small zone of the *Escherichia coli chromosome*. *Genes Dev.* **10**, 1152–1161 (1996).

172. Cebrián, J. *et al.* Direct evidence for the formation of precatenanes during DNA replication. *J. Biol. Chem.* **290**, 13725–13735 (2015).

173. Kono, N., Arakawa, K. & Tomita, M. Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. *BMC Genomics* **12**, 19 (2011).

174. Hendrickson, H. & Lawrence, J. G. Mutational bias suggests that replication termination occurs near the *dif* site, not at Ter sites. *Mol. Microbiol.* **64**, 42–56 (2007).

175. Sternberg, N. & Hamilton, D. Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites. *J. Mol. Biol.* **150**, 467–486 (1981).

176. Blakely, G., Colloms, S., May, G., Burke, M. & Sherratt, D. Escherichia coli XerC recombinase is required for chromosomal segregation at cell division. *New Biol.* **3**, 789–98 (1991).

177. Castillo, F., Benmohamed, A. & Szatmari, G. Xer Site Specific Recombination: Double and Single Recombinase Systems. *Front. Microbiol.* **8**, 1–18 (2017).

178. Neilson, L., Blakely, G. & Sherratt, D. J. Site-specific recombination at *dif* by *Haemophilus influenzae* XerC. *Mol. Microbiol.* **31**, 915–26 (1999).

179. Jensen, R. B. Analysis of the terminus region of the *Caulobacter crescentus* chromosome and identification of the dif site. *J. Bacteriol.* **188**, 6016–6019 (2006).

180. Villion, M. & Szatmari, G. The XerC recombinase of *Proteus mirabilis*: characterization and interaction with other tyrosine recombinases. *FEMS Microbiol. Lett.* **226**, 65–71 (2003).

181. Jouan, L. & Szatmari, G. Interactions of the Caulobacter crescentus XerC and XerD recombinases with the *E. coli dif* site. *FEMS Microbiol. Lett.* **222**, 257–262 (2003).

182. Midonet, C. & Barre, F.-X. Xer Site-Specific Recombination: Promoting Vertical and Horizontal Transmission of Genetic Information. *Microbiol. Spectr.* **2**, 1–18 (2014).

183. Jo, M., Murayama, Y., Tsutsui, Y. & Iwasaki, H. In vitro site-specific recombination mediated by the tyrosine recombinase XerA of Thermoplasma acidophilum. *Genes to cells* **22**, 646–661 (2017).

184. Blakely, G. & Sherratt, D. Determinants of selectivity in Xer site-specific recombination. *Genes Dev.* **10**, 762–773 (1996).

185. Lee, J. & Jayaram, M. Role of partner homology in DNA recombination: Complementary base pairing orients the 5'-hydroxyl for strand joining during Flp site-specific recombination. *Journal of Biological Chemistry* **270**, 4042–4052 (1995).

186. Nunes-Düby, S. E., Azaro, M. A. & Landy, A. Swapping DNA strands and sensing homology without branch migration in λ site-specific recombination. *Curr. Biol.* **5**, 139–148 (1995).

187. Mao, C., Carlson, N. G. & Little, J. W. Cooperative DNA-Protein Interactions Effects of changing the spacing between adjacent binding sites. *J. Mol. Biol.* **235**, 532–544 (1994).

188. Blakely, G. W. & Sherratt, D. J. Interactions of the site-specific recombinases XerC and XerD with the recombination site *dif*. *Nucleic Acids Res.* **22**, 5613–5620 (1994).

189. Hayes, F. & Sherratt, D. J. Recombinase binding specificity at the chromosome dimer resolution site *dif* of *Escherichia coli*. *J. Mol. Biol.* **266**, 525–37 (1997).

190. Pargellis, C. a., Nunes-Düby, S. E., de Vargas, L. M. & Landy, a. Suicide recombination substrates yield covalent lambda integrase-DNA complexes and lead to identification of the active site tyrosine. *J. Biol. Chem.* **263**, 7678–7685 (1988).

191. Bogdanove, A. J., Bohm, A., Miller, J. C., Morgan, R. D. & Stoddard, B. L. Engineering altered protein–DNA recognition specificity. *Nucleic Acids Res.* **46**, 4845–4871 (2018).

192. Yates, J. *et al.* Dissection of a functional interaction between the DNA translocase, FtsK, and the XerD recombinase. *Mol. Microbiol.* **59**, 1754–66 (2006).

193. Panigrahi, G. B. & Sadowski, P. D. Interaction of the NH2- and COOH-terminal domains of the FLP recombinase with the FLP recognition target sequence. *J. Biol. Chem.* **269**, 10940–10945 (1994).

194. Whiteson, K. L. & Rice, P. A. Binding and catalytic contributions to site recognition by Flp recombinase. *J. Biol. Chem.* **283**, 11414–11423 (2008).

195. Leroux, M. Les systèmes Xer à une seule recombinase. (2012). Thèse de Doctorat, Université de

Montréal.

196. Blakely, G. W., Davidson, a O. & Sherratt, D. J. Binding and cleavage of nicked substrates by site-specific recombinases XerC and XerD. *J. Mol. Biol.* **265**, 30–9 (1997).

197. Grainge, I., Lesterlin, C. & Sherratt, D. J. Activation of XerCD-dif recombination by the FtsK DNA translocase. *Nucleic Acids Res.* **39**, 5140–5148 (2011).

198. Wang, Y., Yau, Y.-Y., Perkins-Balding, D. & Thomson, J. G. Recombinase technology: applications and possibilities. *Plant Cell Rep.* **30**, 267–85 (2011).