UCLouvain

FACULTE DES SCIENCES

Faculté des sciences

Tn4430 transposase activity on Mini-Tn-containing DNA molecules *in vitro*

Author: Codemo Loïc

Supervisor: Pr. Bernard Hallet and Dr. Maricruz Fernandez

Readers: Pr. Corentin Claeys Bouuaert and Pr. Patrice Soumillon

Academic year 2021-2022

Thank you

First of all, I would like to thank my promoter, Pr. Bernard Hallet. I would like to thank him for the time he took to help and guide me during my master thesis and for all the advice he gave me. Furthermore, I would like to thank him for making me feel comfortable in the team and for the trust he gave me.

I would also like to thank Dr. Maricruz Fernandez in a special way. Thanks to her I learned to manage myself in the laboratory and I learned all the basic microbiological and biochemical techniques. I would also like to thank her for all the time she gave me to train me and also for her kindness and advice throughout my master thesis. Furthermore, I would like to thank her for helping me to improve my English!

I would also like to warmly thank the whole team of Pr. Hallet in which I had the chance to be integrated. Thank you for all your advice and for having integrated me into the group, thanks to which I was able to work in very good conditions throughout my time in the lab.

I wanted to thank all the members of the BGM team. Thank you for all your advice and for the very interesting discussions we had together.

I also wanted to thank my BGM "co-memorants", Célestin, Kevin, Valentine and Chloe. Thank you for sharing with me every day all the steps of this thesis.

I would also like to thank my family for supporting me throughout this journey and to whom I tried to explain the subject of my thesis.

Finally, I would like to thank all my friends who helped me psychologically during my master thesis. Especially to Soraya M. , my roommates and the dungeon.

I was very happy to have been able to contribute something to this project which is the study of the transposition of the Tn3 family.

Abstract

Transposon Tn4430 belongs to a widespread family of bacterial transposons, the Tn3 family. This family poses a threat to human health because it contributes to the spread of antibiotic resistance among pathogens.

Recently, new discoveries have been made in understanding the transposition mechanism of Tn4430, making this transposon a new paradigm for the study of Tn3 family transposons. Biochemical analysis of the initial steps of Tn4430 transposition has unveiled a unique pathway leading to the assembly and activation of the transposition complex. This pathway is now supported by high-resolution cryo-EM structures depicting the transposition complex at different stages of the reaction.

In addition to catalyse transposition, TnpA is responsible for a mechanism called "target immunity" which prevents the transposon from being inserted several times in the same locus. The molecular details of this mechanism are not known, but TnpA mutants that are affected in target immunity have been isolated. These mutants showed hyperactivity *in vitro*, being more prone than wild-type TnpA to catalyse DNA cleavage and strand transfer reactions.

Finally, complementary *in vivo* and *in vitro* data converge to a new "replication hijacking" mechanism according to which Tn3-family transposons integrate into replication intermediates to recruit the host replication machinery and produce a new copy of themselves during transposition.

The aim of this work is to set up a biochemical assay to determine the activity of TnpA and its hyperactive mutants ($TnpA^{3X}$ and $TnpA^{S911R}$) *in vitro*. Through this test, the impact of DNA topology on transposase activity will be determined as well as the requirements of transposon end orientation. Other factors such as the role of the cofactor or the role of glycerol will be tested.

The results of this work showed that the topology of the DNA can influence the activity of TnpA such as the presence of both ends of the transposon, the type of cofactor used and the presence of glycerol. In addition, additional bands appear on gels and hypotheses are put forward as to the origin of these bands.

Résumé

Le transposon Tn4430 appartient à une famille très répandue de transposons bactériens, la famille de Tn3. Cette famille constitue une menace pour la santé humaine car elle est impliqué dans la propagation de la résistance aux antibiotiques chez les agents pathogènes.

Récemment, des découvertes ont été réalisées dans la compréhension du mécanisme de transposition de Tn4430, faisant de ce transposon un nouveau modèle pour l'étude des transposons de la famille Tn3. L'analyse biochimique des étapes initiales de la transposition du Tn4430 a dévoilé une voie unique menant à l'assemblage et à l'activation du complexe de transposition. Cette voie est maintenant soutenue par des structures cryo-EM à haute résolution représentant le complexe de transposition à différentes étapes de la réaction.

En plus de catalyser la transposition, TnpA est responsable d'un mécanisme appelé "immunité de cible" qui empêche le transposon d'être inséré plusieurs fois dans le même locus. Les détails moléculaires de ce mécanisme ne sont pas connus, mais des mutants de TnpA qui sont affectés dans l'immunité de cible ont été isolés. Ces mutants ont montré une hyperactivité *in vitro*, étant plus enclins que la TnpA *wild-type* à catalyser les réactions de clivage de l'ADN et de transfert de brin.

Enfin, des données supplémentaires *in vivo* et *in vitro* convergent vers un nouveau mécanisme de "replication hijacking" selon lequel les transposons de la famille Tn3 s'intègrent dans les intermédiaires de réplication pour recruter la machinerie de réplication de l'hôte et produire une nouvelle copie d'eux-mêmes lors de la transposition.

Le but de ce travail est de mettre en place un test biochimique pour déterminer l'activité de TnpA et de ses mutants hyperactifs ($TnpA^{3X}$ et $TnpA^{S911R}$) *in vitro*. Grâce à ce test, l'impact de la topologie de l'ADN sur l'activité transposase sera déterminé ainsi que l'impact de l'orientation des extrémités du transposon. D'autres facteurs tels que le rôle du cofacteur ou le rôle du glycérol seront aussi testés.

Les résultats de ce travail ont montré que la topologie de l'ADN peut influencer l'activité de la TnpA comme la présence des deux extrémités du transposon, le type de cofacteur utilisé et la présence de glycérol. De plus, des bandes supplémentaires apparaissent sur les gels et des hypothèses sont émises quant à l'origine de ces bandes.

Table of content

I. Introduction

In 1950, the researcher Barbara McClintock tried to explain the variations in the colour of Zea mays grains, which were due to mutations whose origin was unknown. She noticed structural variations on chromosome 9. She described the Ds (Dissociation) locus corresponding to a break whose position seemed to move from one place to another along the chromosome. She also observed that Ds activity is dependent on another Ac (Activator) locus which was also able to move within the genome. Finally, she hypothesised that these phenomena are caused by heterochromatin transposition (McClintock 1950). She in fact discovered the first two transposable elements (Ds and Ac). As a result, she won the Nobel Prize for Medicine and Physiology in 1983.

Today, transposable elements or "transposons" are generally described as defined segment of DNA that are able to propagate from one location to another within the genome of their host (Craig et al. 2015, Hickman and Dyda 2016). Transposons are therefore part of the Mobile Genetic Elements (MGE), just like plasmids and bacteriophages in prokaryotes. MGEs share a modular organisation in addition to their ability to move. The modules can be grouped into three classes according to their function. Each of these modules is composed of several genes and other sequences related to a specific function. The first class of modules includes stability modules that ensure the persistence of the MGE within its host. The second class of modules includes the intercellular mobility modules that allow the transport of the MGE to a new host. Finally, the third class of modules includes the intracellular mobility modules that allow the movement of the EGM within its host (Toussaint and Merlin 2002). In all transposable elements, there is at least one genetic module for intragenomic mobility. Transposons of different classes can be autonomous or non-autonomous. Autonomous elements carry a complete and functional module of mobility, generally consisting of specific delimiting sequences and one or more genes for proteins generically referred to as "transposases". On the other hand, there are also non-autonomous transposable elements whose mobility module consists solely of delimiting sequences. These elements can move thanks to the presence of an autonomous element of the same family (Fedoroff et al. 1983). Transposable elements are also capable of carrying stability modules such as antibiotic resistance genes or metabolic capacity genes. In addition, their intracellular mobility gives them the possibility to take advantage of the extracellular mobility of other EGMs such as

plamids in prokaryotic cells (Touchon and Rocha 2007). This favour their dispersal into the genome of other hosts in prokaryotes.

Transposases encoding genes are by far the most abundant and ubiquitous genes in nature, which reflects the importance of transposons in the organisation and evolution of genetic material since transposases not only promote the spread of transposable elements, they also lead to mutations and rearrangements that can enhance biological diversification and thus evolution (Aziz RK Breitbart M & Edwards RA 2010). An example is the human genome, where it is assumed that sequences derived from transposable elements constitute up to 50% of the genome (Smit 1999; Yoder, Walsh, and Bestor 1997). For Zea mays, it is estimated that 60% of its genome is composed of them (SanMiguel et al. 1996).

In eukaryotes, transposons can acquire regulatory sequence as a result of recombination events or the formation of composite transposons. If a transposon acquires a regulatory sequence, multiplication of the transposon will increase the number of copies of this sequence and its movements can lead to the insertion of this sequence close to genes and ultimately cause a change in their expression (Chuong, Elde, and Feschotte 2017). Transposons thus play an important role in the function and structural organisation of chromosomes in eukaryotes, constituting a major source of genetic innovation (Kazazian 2004).

In prokaryotic organisms, transposable elements promote the movement of a pool of mobile genes referred to as the "mobilome". This is very useful for bacteria as it allows them to transfer specific genes of interest by horizontal transfer (Frost et al. 2005). They are notably involved in the dissemination of antibiotic resistances among pathogenic bacteria (Partridge 2018). The spread of antibiotic resistances is a concerning problem for human health. Today, around 700,000 people die due to resistant or multi-resistant infections every year. If we do nothing, at the horizon of 2050, almost 10 million people will die from microbial infection per year (O'Neill 2016). This is why the issue of microbial resistance is now regarded as a priority health concern by the UN (Thomson 2016).

I.1 Classification of transposons

Transposons are very diverse. There are transposons with different sizes, transposition mechanisms and hosts. In the remainder of this section we will outline two ways of classifying transposons, classification according to the transposition intermediate and classification

according to the transposition mechanism, but other classification exists depending on different criteria (Piégu et al. 2015).

I.1.1 Classification based on the transposition intermediate

Transposons can be classified according to their transposition intermediate. This intermediate state is the transient state of the transposed sequence before it integrates into a target (Finnegan 1989). The transposable elements can therefore be classified into two classes. Class 1 of transposons include those whose transposition intermediate is an RNA molecule. Class 2 of transposons include those elements whose intermediate is a DNA (Wicker et al. 2007).

In class 1 transposons, also called retrotransposons. During transposition, the transposon is transcribed into RNA and then the reverse transcriptase retrotranscribes the RNA into complementary DNA (cDNA). Finally, this cDNA is integrated into a new chromosome locus. However, within this class a distinction can be made between LTR-retrotransposons that are bounded by Long Terminal Repeats (LTRs) and non-LTR retrotransposons. The former synthesises their cDNA in viral particles while the latter produce their cDNA at the insertion site using a mechanism termed target-primed transposition (Goodier 2016).

Class II transposons or "DNA transposons" move directly from one DNA locus to another without producing an RNA intermediate. For this class, the intragenomic mobility module consists of at least one gene encoding the transposase and delimiting sequences. These sequences are often inverted repeats (IRs) which can be of varying sequences and lengths. However, transposition can be replicative or non-replicative depending on whether the transposon is duplicated during the transposition process. In non-replicative transposition, the transposon sequence is first excised and then inserted into the target by the transposase. In replicative transposition, the transposon is replicated by the host replication machinery before or after insertion into the target by the transposase (Feschotte and Pritham 2007; Mizuuchi 1992).

I.1.2 Classification by transposition mechanism and transposase

Transposons can also be classified according to their transposition mechanism and the transposase. This classification was first proposed by Curcio and Derbyshire in 2003 (Curcio and Derbyshire 2003).

During transposition, DNA sequences will move within the genome, which involves the insertion of a transposon into a target and thus the breaking and formation of phosphodiester bonds (Feschotte and Pritham 2007; Heffron et al. 1979). The transposition mechanism is often used to classify transposases. Five distinct families can thus be discerned, characterised by the amino acid residues involved in the reaction, the possible formation of a covalent protein-DNA intermediate and the possible involvement of metal ion cofactors (Hickman and Dyda 2015). There are HUH-motif transposases (Ronning et al. 2005), Tyrosine transposases (Goodwin and Poulter 2001), serine transposases (Grindley et al. 2006), casposases (Krupovic et al. 2014) as well as DDE(D)-motif transposases (Hickman et al. 2010). The mechanism of transposition of DDE(D) motif transposases will be detailed in the next section.

I.2 DDE(D) motif transposases and their diversity

I.2.1 Transposition mechanism of DDE(D) motif transposases

Transposase/integrase proteins with DDE(D) motifs are found in the majority of class II transposons and class I LTR retrotransposons. This superfamily of enzyme is characterized by the presence of a Ribonuclease H type fold catalytic domain containing the conserved triad DDE(D) (Asp, Asp, Glu or Asp). These transposases catalyse hydrolysis and transesterification reactions in DNA. They differ from serine, tyrosine and HUH-motif transposases in that they do not form a covalent intermediate during transposition (Hickman et al. 2010).

In this family of transposases, the chemical transposition mechanism consists of two steps, a cleavage step and a strand transfer step. During the cleavage step, the DDE/D residues coordinate 2 divalent metal ions with a substrate phosphate. Then a hydroxyl ion makes a nucleophilic attack on the phosphate, which leads to the formation of a planar pentavalent intermediate. After this, the phosphodiester bond breaks and this release the 3' OH end of the substrate. In the second step, the newly released 3' OH group makes a nucleophilic attack on a phosphate of the target. This leads to the breakage of the target phosphodiester bond and the formation of a new bond between the donor and recipient DNA (Fig. I.2.1) (Hickman et al. 2010).

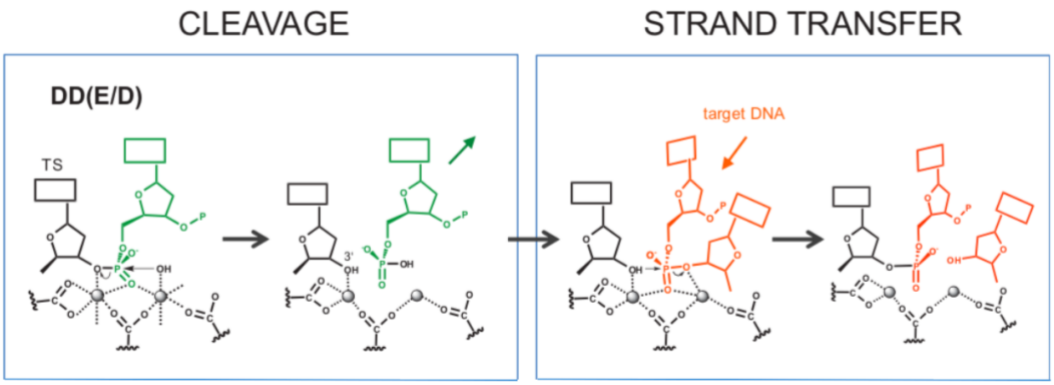


Figure I.2.1: ***Chemical reactions catalysed by DDE(D) motif transposases.*** *The green DNA shows the cleaved dinucleotide and the orange is the target strand. The spheres indicate divalent metal ions bound to the transposase. (Hickman et al. 2015)*

I.2.2 Diversity of DDE motif transposases

There is a great diversity of transposition modes within the DDE(D) motif transposases. Transposition can be non-replicative or replicative. In the first case, transposition involves the physical transfer of the transposon DNA sequence from one location to another. This is why this mechanism is known as "Cut and Paste". In this non-replicative mechanism, the 4 phosphodiester bonds at the ends of the transposon will be broken, which will generate double-strand breaks. There are different types of “Cut and Paste” mechanisms, depending on the transposon family (Fig. I.2.2C/D/E) (Hickman et al. 2015).

In the second case, replicative transposition will lead to the creation of a new copy of the transposon by replication, this mechanism is also called "Copy-paste". There are two main mechanisms of replicative transpositions. The “cut-out-copy-in” mechanism (Fig. I.2.2A) used by the Mu phage and the Tn3 family transposons and the “copy-out-paste-in” mechanism used by the IS3 family (Fig.I.2.2B) (Hickman et al. 2015; Kosek et al. 2021). In this work, only the “cut-out-copy-in” mechanism will be detailed.

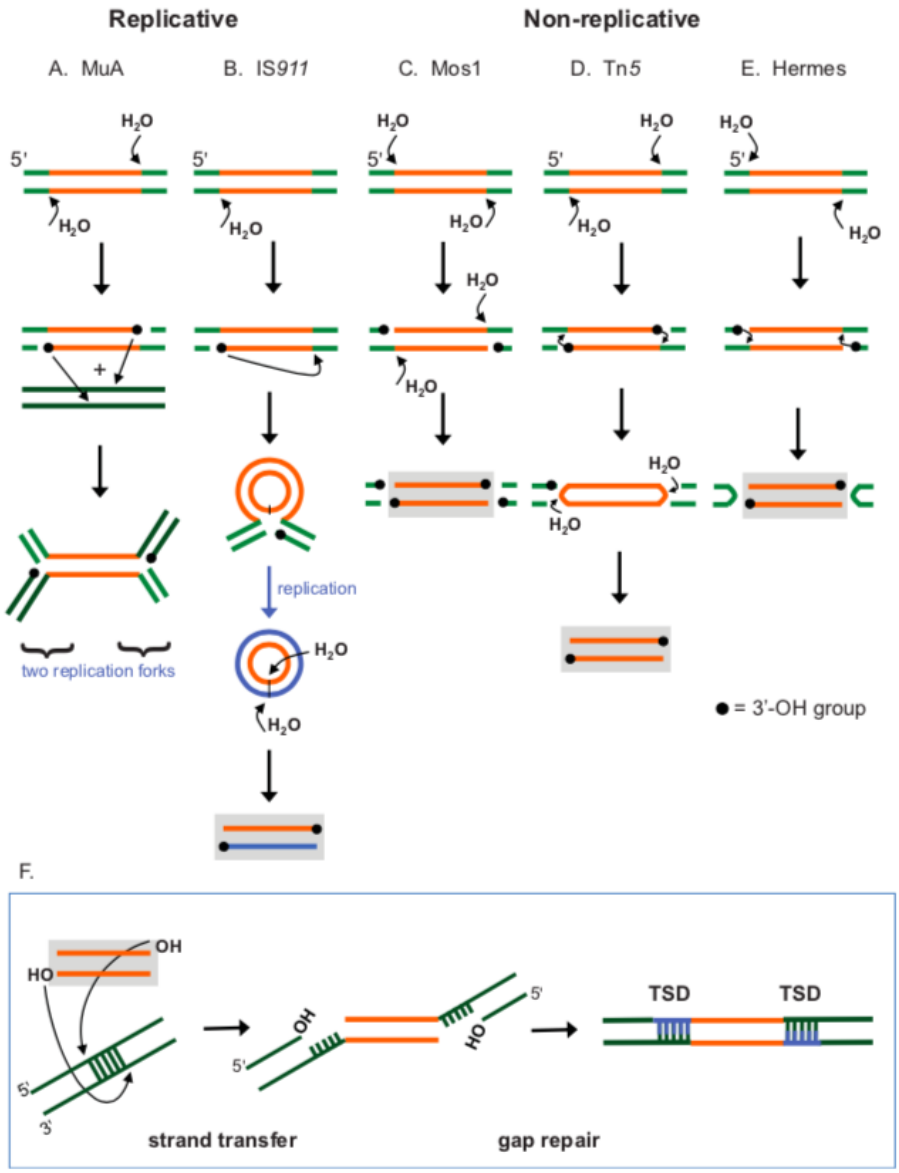


Figure I.2.2: ***Different transposition pathways of DDE(D) motif transposases.*** *Arrows show strand cleavage sites. The black dots represent the 3' -OH groups. In some cases, the pathways lead to the same form of the excised transposon (shown in grey boxes). This excised intermediate is then integrated into the target DNA, as shown in (F). When the cell repairs the gaps introduced by the strand transfer reactions, target site duplications (TSDs) are generated. (A) "Cut-out-copy-in" mechanism used by the Mu phage and the Tn3 family. (B) "Copy-out-paste-in" mechanism used by the IS3 family including IS911 transposase. (C) "Cut and paste" mechanism used by the IS630 family and the Tc1/Mariner superfamily including Mos1 transposase. (D) "Cut and paste" mechanism used by the IS4 family such as Tn10 and Tn5 and by the piggyBac. (E) "Cut and paste" mechanism used by the Hermes and V(D)J transposase. (Hickman et al. 2015)*

I.2.2.1 Replicative transposition by "paste and copy" or "cut-out-copy-in" mechanism

Replicative transposition couples transposition to DNA replication leading to the creation of a second copy of the transposon at a new target site. This mechanism allows a transposon to expand and proliferate in a genome. In this mechanism only 2 phosphodiester

bonds are broken, one at each end of the transposon. Therefore, it does not generate double-strand breaks.

One of the best documented transposable element using this mechanism is the bacteriophage Mu, which uses replicative transposition to multiply its genome during infection. During replicative transposition of this phage, the transposase encoded by the phage, called MuA, catalyses two hydrolysis reactions at each 3' end of the transposon. This nucleophilic attack reaction is carried out by water on the end of the transposon, generating a 3’ -OH group (Fig. I.2.2.1A). Then this free 3' -OH group is used to perform a transesterification reaction at the transposon insertion site (Fig. I.2.2.1B). This strand of the transposon end is therefore called the "transferred strand". Each strand of the transposon is therefore linked to both the donor at its 5' end and to the recipient at its 3' end creating a branched DNA structures at each transposon end (Fig. I.2.2.1C). This structure is then resolved by the host replication machinery to duplicate the transposable element (Hickman et al. 2015). Intermolecular transposition, transposition between 2 different molecules, by this mechanism generates a transposition product called "cointegrate". In this cointegrate, the donor and recipient molecules are linked to each other by two copies of the transposon. This mechanism is also used by transposons of the Tn3 family. The replication begins at the 3’ -OH ends that have been released by the cleavage of the target to create the complementary strands of the transposable element and thus form the cointegrate. DNA synthesis will also repair the single-stranded gaps at either end of the transposon, generating 5 bp direct repeat (DR) duplications that border the two copies of the element in the resulting cointegrate (Fig. I.2.2.1D) (Hickman et al. 2015).

Replicative transposition of the Mu phage follows a "replication hiring" mechanism in which the cellular replication machinery is recruited to the integration site after the strand transfer reaction catalysed by the MuA transposase. This mechanism is reminiscent to ectopic replication initiation taking place during DNA repair and replication re-start at stalled replication forks (Shapiro 1979, Mizuuchi 1992).

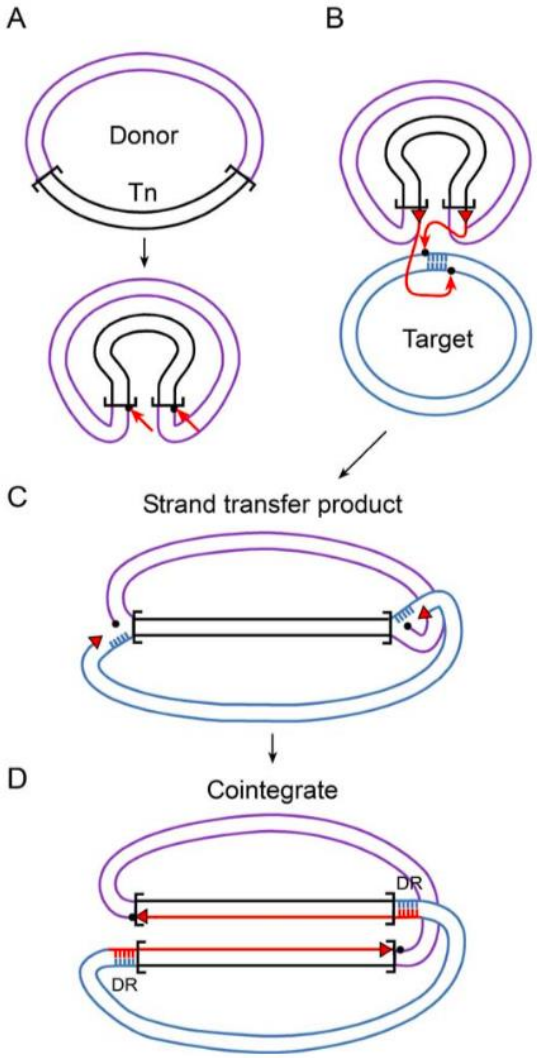


Figure I.2.2.1: ***Replicative transposition mechanism of phage Mu and Tn3.*** *The diagram illustrates the case of intermolecular movement of a transposon between a donor and a recipient DNA molecule. The transposon is represented by a double black line delimited by square brackets, the donor DNA in purple, the recipient DNA in blue, the free 3' -OHs by red triangles the target phosphates by black dots and the direct repeat by DR. (A) Transposition begins when the transposase generates single-strand specific nicks at both 3' ends of the transposon, this creates 3'- OH groups in the donor. (B) The 3'-OH ends of the transposable element will then make a nucleophilic attack on the phosphodiester bonds on each target DNA strand. (C) The reaction generates a strand transfer product in which the transposon is bound to the target and donor DNA. (D) Replication begins at the 3' -OH released by target cleavage. It will synthesise the complementary strands of the transposon (red line) and finally form the cointegrate. DNA synthesis also repairs single-strand gaps at the ends of the transposon. This creates 5 bp direct repeat (DR) duplications that border the two replicates of the element in the final cointegrate (Nicolas et al. 2015).*

I.3 Structure of DDE(D) motif transposases

Although DDE(D)-motif transposases have transposition mechanisms that can be very different from each other (Hickman et al. 2015), these transposases still require the same types of interactions and therefore often share similar functional domains. However, there are large differences in molecular structure depending on the transposase family (Hickman et

al. 2010). DDE(D) transposases typically contain a ribonuclease H (RNase H) catalytic domain with one or more DNA binding domains. The transposase organises and catalyses the DNA cut and joining reactions within a nucleoprotein complex called "transpososome". This may consist of the ends of the transposon bound to a transposase multimer. This minimal configuration is called the synaptic or paired-end complex (PEC). To perform strand transfer, the transpososome can assemble the Strand Transfer Complex (STC). In the latter, the transposase also contacts the target via a specific DNA-binding interface (Hickman et al. 2010).

I.3.1 RNase H catalytic domain

In the RNase H catalytic domain is often composed of a 5-strand β-sheet surrounded by α-helices. These strands and helices are usually organised as follows, β1-β2-β3-α1-β4-α2/3-β5-α4-α5 (Fig. I.3.1). Furthermore, in the domain, the position of the amino acids of the catalytic triad is conserved. The first aspartate residue is always in β1, the second aspartate on or just after β4 and the glutamate (aspartate) on or just before of α4. It is also observed that in the majority of transposases whose structure is known, a sequence called "insertion domain" is present between the β5 strand and the α4 helix. This sequence is of variable size and its function varies according to the transposon family (Hickman et al. 2010).

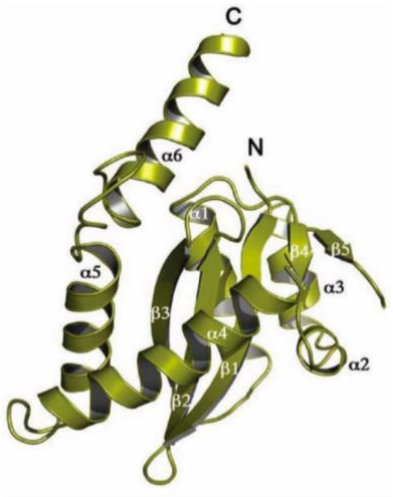


Figure I.3.1 ***: Ribbon diagram of the RNase H catalytic domain of HIV-1 integrase****. Standard secondary structure elements are highlighted. β indicates the presence of a strand of the β-sheet and the number indicates the strand number. The α indicates the presence of an α-helix and the number indicates the number of the helix (Hickman et al. 2010).*

I.3.2 DNA binding domain

The DNA binding domain of the transposase allows to recognise specific DNA sequences at the ends of the transposon (Hickman et al. 2010). In most cases, a transposase monomer can only bind one end of the transposon and the transpososome therefore always contains at least one transposase dimer. However, it can also take more complex forms such as in the Mu transpososome where the active transposase consists in a hexamer of MuA (Montano et al. 2012).

I.4 Transpososome and regulation

One of the most important steps in the regulation of transposition is certainly the assembly of the transpososome. In fact, the catalytic activity of transposases can only be triggered once the various partners of the reaction have been assembled. The reason for this is to coordinate the reactions and thus avoid incomplete transposition which could be detrimental to the cell. Coordination is generally feasible because the hydrolysis reactions performed by the transposase are catalysed in *trans*. This means that one transposase monomer is bound to one end of the transposon by its DNA binding domain but its RNase H domain cleaves the other end of the transposon (Hickman et al. 2010).

I.4.1 Synaptic complex assembly

During the formation of the transpososome, the transposase will bring the two ends of the transposon together to create the synaptic complex. This process can take place in two ways, depending on the oligomeric state of the transposase in solution.

For transposases that are monomeric in solution, oligomerisation of the protein takes place in the presence of substrate. First, a monomer will attach to each end of the transposon, forming the Single-End Complex (SEC). After this, the transposase undergoes multimerisation which leads to the formation of the PEC. This mechanism is called “Synapsis by Protein Dimerization” (S-PD). Here, the ends of the transposable element are therefore recruited simultaneously (Fig. I.4.1.1A) (Bouuaert et al. 2013; Blundell-Hunter et al. 2018).

For transposases which are in oligomeric form in solution the mechanism is different. This mechanism is called Synapsis by Naked End Capture (S-NEC). First, one end of the transposon is contacted by an oligomer of transposase, forming the SEC. Then it contacts the

second end of the transposon, forming the PEC. The ends of the transposon are thus recruited sequentially (Fig. I.4.1.1B) (Bouuaert et al. 2013; Blundell-Hunter et al. 2018).

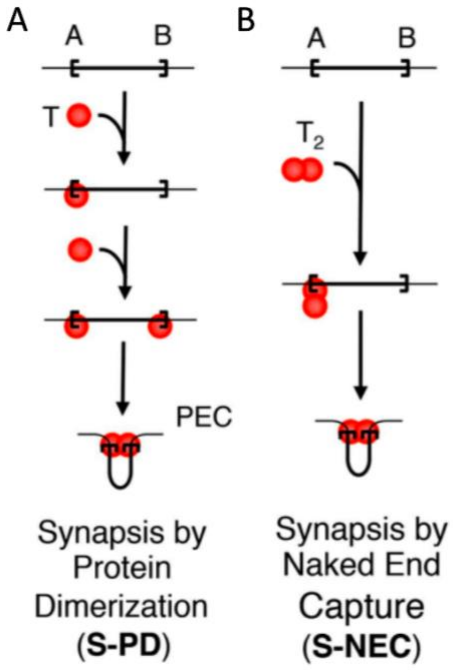


Figure I.4.1.1: ***Two pathways of transpososome assembly.*** *Transposase monomers are red spheres. The transposon ends are half-rectangles. The flanking DNA is omitted for clarity. (A) A transposase monomer binds to each end of the transposon and the Paired End Complex is formed by dimerisation. (B) A transposase oligomer, here a dimer, first contacts one end of the transposon and after recruits the other end to form the Paired End Complex. (Bouuaet et al. 2013)*

A mechanism called "overproduction inhibition" (OPI) is a self-regulating mechanism of transposition. When the concentration of transposase is high, it prevents the formation of PEC and thus inhibits transposition. The S-NEC mechanism is thought to be a mechanism that can explain OPI. If the concentration of transposase increases the number of free transposon ends decreases. This leads to a decrease in the rate of transposition because if both ends of the transposon are bound to a transposase dimer, transposition cannot take place. This mechanism is called “Assembly-site occlusion” model (ASO) (Fig. I.4.1.2) (Bouuaert et al. 2013).

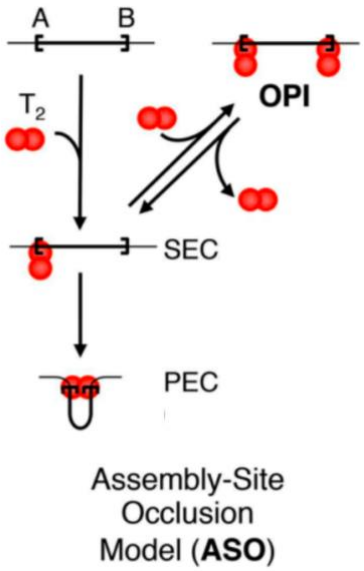


Figure I.4.1.2: ***The assembly-site occlusion model.*** *Transposase monomers are red spheres. The transposon ends are half-rectangles. The flanking DNA is omitted for clarity. Only one of the two ends of the transposon is occupied by a transposase dimer when the concentration of transposase is low. An active synapse is thus obtained (bottom left). On the other hand, both ends of the transposon are bound to a transposase dimer when the concentration of transposase is high, resulting in an overproduction inhibition (OPI, top right). (Bouuaet et al. 2013)*

For certain transposition reactions, the topology of the substrate also plays a role in the assembly of the transpososome, so the transposase will be selective according to the topology of the substrate. This selectivity is actually a result of the negative supercoiling of the substrate. Thanks to this supercoiling, the sites recognised by the transposase will be in a favourable geometry which will facilitate the assembly of the transpososome. This mechanism is called “topological filter” and it disadvantages the establishment of the transpososome between ends located on different molecules and between misoriented ends. In the S-NEC model, recruitment of the second end of the transposon during the establishment of the transposition complex appears to be a rudimentary topological filter (Bouuaert et al. 2011; Bouuaert and Chalmers 2013, Harshey 2014).

For example, for the eukaryotic transposon Hsmar1, a member of the Tc1-mariner superfamily, the topology of the donor substrate will impact on the assembly of the synaptic complex. When the donor substrate is negatively supercoiled, it accelerates the formation of the transpososome (Bouuaert and Chalmers 2013).

I.4.2 Target recruitment

The strand transfer complex is the final complex that makes transposon insertion into the target possible. As explained before, this complex consists in at least the transposase in oligomeric form, the two ends of the transposon and the target (See section I.3).

Like the assembly of the synaptic complex, the recruitment of the target into the transposition complex is a major regulatory step. This recruitment can occur at different times during transposition. In some families of transposons, recruitment of the target is necessary for the establishment of the active transpososome (Peters 2015). In other families, the target is the last element recruited before the transfer reaction (Sakai and Kleckner 1997).

I.4.3 DNA targeting and transposition immunity

Some transposable elements, such as the transposons of the Tn3, Tn7 and Mu phage family, have a very complex regulatory mechanism in terms of target selection. The main consequence of this is the transposition or target immunity mechanism. This mechanism prevents a transposon from inserting itself several times in the same locus. This process is thought to provide a considerable advantage for the propagation and stability of the transposable element. This can be said because an insertion of the transposon within itself will probably inactivate its mobility. Furthermore, if two homologous transposon sequences are close to each other, this may lead to deletions by homologous recombination (Mahillon and Chandler 1998). In particular, in the Tn3 family, the role of immunity is also to prevent inversions or deletions of the donor molecule resulting from intramolecular replicative transposition. Finally, this mechanism promotes a greater dispersion of the transposon between different replicons in the cell (Nicolas et al. 2017).

The best-known mechanisms are those of Tn7 and the bacteriophage Mu. In both cases, immunity requires the intervention of transposition helper proteins directly encoded by the transposon (Peters 2015; Harshey 2012). In contrast, the mechanism of immunity of Tn3 is not yet fully described. This mechanism does not require proteins other than transposase (Nicolas et al. 2017).

I.5 The Tn3 family

The Tn3 family is a family of replicative transposons. This family of transposons is particularly involved in the spread of antibiotic resistance genes (Nicolas et al. 2015). Members of this family are for example involved in the recent epidemic of carbapenem-resistant enterobacteria and the even more recent spread of colistin resistance. These two antibiotics are known to be the "last chance" antibiotics. Transmission of resistances against them thus seriously compromises their use against multi-drug resistant bacteria (Cerqueira et al. 2017; Snesrud et al. 2018).

In the Tn3 family of transposable elements there is an intragenomic mobility module consisting of a DDE(D) motif transposase and 38 bp IR sequences. They also contain a resolvase module composed of a resolvase and an internal recombination sequence called *res* or IRS (Nicolas et al. 2015). This resolvase may be a serine recombinase (Nicolas et al. 2015) or a tyrosine recombinase (Vanhooff et al. 2006). Finally, some transposons may also contain passenger modules (Nicolas et al. 2015).

Tn3 family transposons exhibit target immunity mechanism as Tn7 and bacteriophage Mu. In contrast to Tn7 and phage Mu, which use auxiliary proteins for their immunity mechanisms (Peters 2015; Harshey 2015). Immunity of Tn3-family transposons only requires the TnpA transposase and the presence of a single IR transposon end within the target. This immunity is quite specific to each member of the Tn3 family (Athur et al. 1984; Lambin et al. 2012).

Transposases from the Tn3 family are the largest members of the DDE(D) super-family, with a size ranging from 950 to 1020 aa (Nicolas et al. 2015).

Despite their biological and societal importance, the molecular mechanism of Tn3-family transposons is still poorly understood biochemically.

I.6 Tn4430 as a model of the Tn3 family

Our laboratory uses Tn4430 from Bacillus thuringiensis as a model representative of Tn3-family transposons. Tn4430 was discovered in *Bacillus thuringensis* but is fully functional in *Escherichia coli* (Mahillon and Lereclus 1988). The transposon contains two coding sequences, one for the transposase TnpA (987 aa) and the other for the tyrosine integrase TnpI (348 aa). There is also an IRS sequence at the level of which TnpI catalyse a DNA site

specific recombination to resolve transposition cointegrate products (Mahillon and lereclus, 1988; Vanhooff et al. 2006). (Fig. I.6). Furthermore, it does not contain any passenger genes.

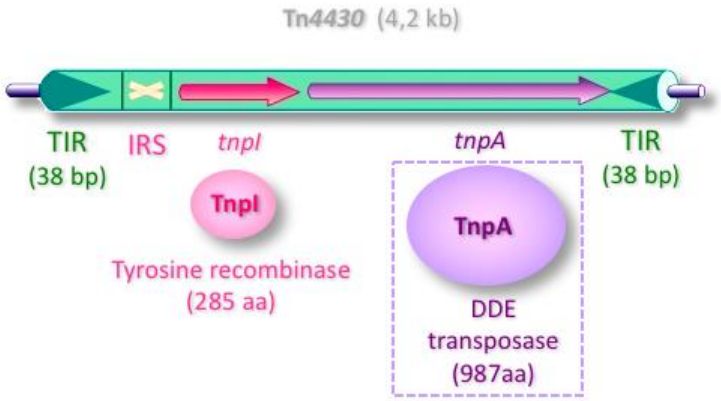


Figure I.6: ***Structure of Tn4430.*** *The 2 TIRs are represented by green triangles, the white cross represents the IRS sequence, the pink arrow represents the gene coding for tyrosine recombinase TnpI and the purple arrow represents the gene coding for the transposase TnpA.*

I.6.1 Replisome hijacking model

Studies on DNA targeting by Tn4430 have revealed that the replicative transposition mechanism proposed for the Mu bacteriophage is not the same as that of the Tn3 family. Firstly, *in vivo* experiments showed that Tn4430 preferentially inserts into genomic regions where replication forks are likely to be slowed, stopped or restarted (Nicolas et al. in prep). Moreover, if target replication is blocked, transposition does not proceed. In contrast, if donor replication is blocked, this has no effect on transposition. This shows that replication is a necessary condition for the transposition of Tn4430 (Nicolas et al. in prep). Secondly, *in vitro*, it has been shown that TnpA-mediated strand transfer can occur with synthetic DNA replication forks and that this substrate is effective (Nicolas et al. in prep). All these results lead to a new model of transposition, called "replication hijacking". In this model, the Tn3 transposons would jump into the replication intermediates and directly recruit the host replication machinery at the time of integration into the target (Fig. I.6.2D). The classical "replication hiring" mechanism, described in the bacteriophage Mu, is different from this new model since in this classical replication mechanism was recruited after the integration of the transposon into its target (Nicolas et al. in prep).

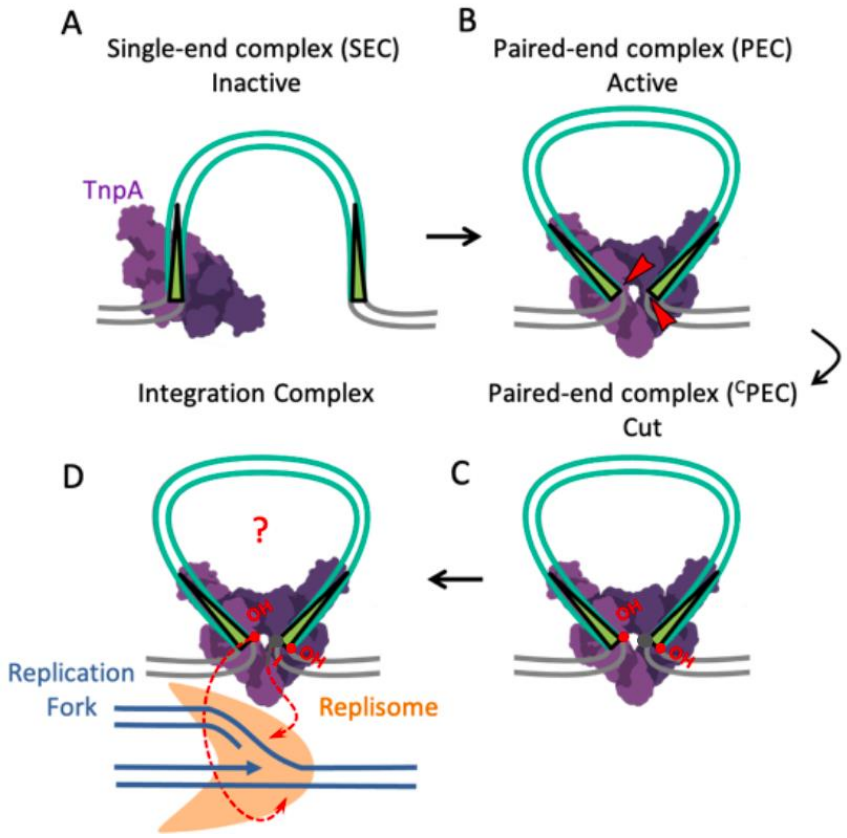


Figure I.6.2: ***Model of the replication hijacking mechanism****. The TnpA dimer is shown in purple, the transposon in green, the transposase cleavage sites in red arrowheads, the target DNA in blue, the replication machinery in orange, the transposon ends by green triangles and the transfer of the transposon ends by red arrows. (A) The TnpA dimer binds to one end of the transposon and assembles the SEC in which the transposase is inactive. (B) TnpA undergoes a conformational change that allows it to form the PEC, which can cleave DNA. (C) The cleavage generates a 3'-OH group at both ends of the transposon that will serve for the strand transfer reaction (D) Strand transfer occurs through the assembly of the strand transfer complex including PEC and target DNA at the moment of DNA replication. Joining of the replication machinery is likely to occur early in the pathway to initiate catalysis. The transposase transfers the transposon ends to the branch point of a replication fork, allowing the transposon to be duplicated when replication resumes.*

I.6.2 The structure of TnpA and paired end complex (PEC)

Using single particle cryo-EM the structure of TnpA was determined in apo state and paired end complex (PEC) conformation (Fig I.6.1A). In both apo and PEC conformations, TnpA exists as a dimer. Binding of the transposon ends is accompanied by major conformational changes transforming the compact apo complex into an expanded V-shaped PEC with about 140 Å long arms. Each monomer can be divided into 10 structural domains (Fig. I.6.1B), including four *cis*-DNA binding domains (DBD 1-4), an RNase H fold domain, 2 dimerisation domains (DD), an arm domain, a linker domain, an RNase H fold insert. DBD1 is separated from DBD2 by a 40 Å long α-helical arm. The first dimerisation domain separates DBD2 and DBD3. The second dimerization domain, DD2, is located at the C-terminus of the protein. The

linker separates DBD3 and DBD4 and the insert splits the RNase H fold domain in two. These 2 domains, together with the region downstream of the RNase H folding, form a "scaffold" domain that encircles the active site. In fact, the structure suggests that the scaffold domain 'helps' the fold of the RNase H domain to adopt its structure upon activation. The flanking sequences are also in contact with the transposase during the formation of the transposition complex. The DDE catalytic triad (Asp679, Asp751 and Glu881) is located in the RNase H domain. (Fig. I.6.1) (Shkumatov et al. *in prep*) . In addition, the complex traps a positive DNA cross while the plasmid supercoil is negative. This could have an impact on the activity of the transposase depending on the topology of the substrate.

The conformational change from the apo state to the PEC conformation involves dramatic changes during which the protein module upstream of DBD4 shifts and rotates about 50 degrees like a rigid body. This apo-PEC transition allows each transposase monomer to bind to one end of the transposon through its DBDs in order to properly position the RNase H domain for cleavage of the other end in *trans*. The majority of interactions with DNA are in *cis* through the DBDs but there are still some interactions in trans around the RNase H domain. During the conformational change the flanking DNA is bent so that both ends of the transposon can be used by the transposase more precisely at the cleavage sites. The conformational transition also creates an opening between the monomers and the DD1 that is absent in the apo conformation, placing TnpA optimally for binding to the target DNA. This cavity could be used for target DNA capture. After cleavage of the 3' ends of the transposon, the flanking DNA undergoes rotation suggesting that the DNA contains bending and possibly twisting energy from the supercoil. This energy could be used to drive the reaction (Shkumatov et al. *in prep*).

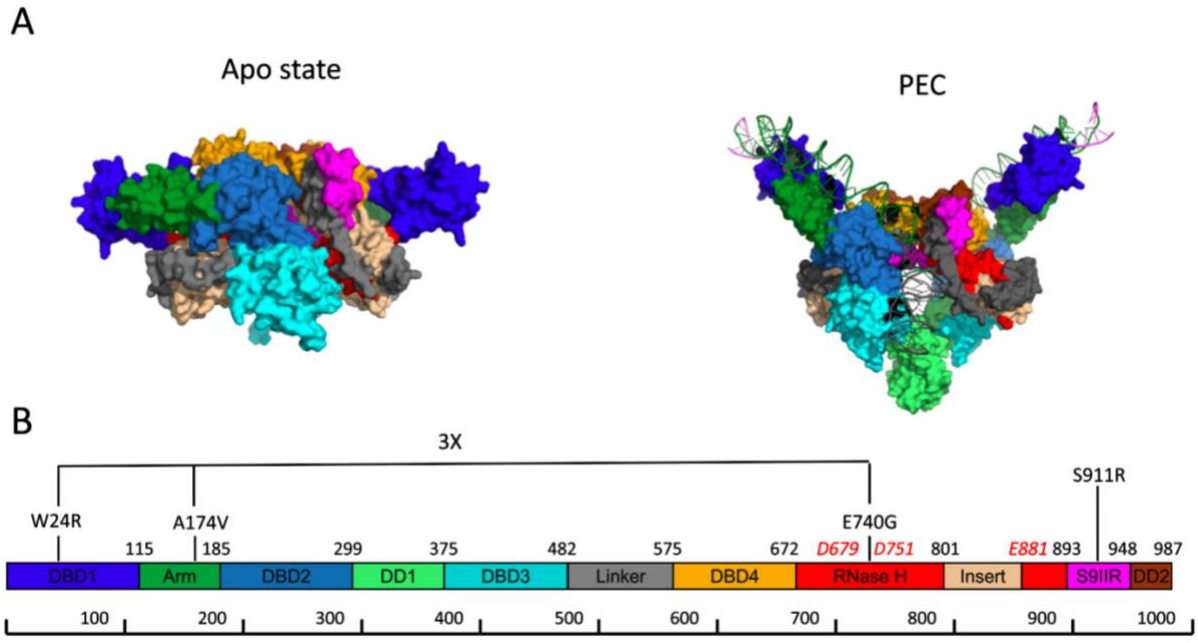


Figure I.6.1: ***Architecture of TnpA in apo and PEC conformations.*** *(A) Cryo-EM reconstructions of TnpA^{WT} apo and TnpA^{S911R} PEC conformations. The TIRs are shown in green and the flanking sequences in grey. (B) Linear diagram of structural domains constituting. Positions of the DDE active site residues and the T+/- mutations in TnpA3X (W24R, A174V, E740G) and TnpAS911R are indicated. (Shkumatov et al.* in prep*)*

I.6.3 Immunity mutants of TnpA

Although the mechanisms that regulate target immunity in Tn3 family transposons are poorly understood. It has been shown for Tn4430, that target immunity activity can be uncoupled from transposition activity (Lambin et al. 2012). Mutants of the transposase were identified that showed increased transposition into targets that were immunized by the presence of a Tn4430 TIR end. These mutants are called T+/I- because they can transpose but are impaired in immunity. There are several mutations along the transposase sequence that can lead to this phenotype (Lambin et al. 2012). Two mutants in particular, TnpA^{S911R} and TnpA^{3X} (i.e., the triple mutant TnpA$^{W24R, A174V, E740G}$) showed strongly reduced target immunity compared to wild-type TnpA (TnpA^{WT}) (Nicolas et al. 2017) (Fig. I.6.1). These deregulated mutants also showed promiscuous activity *in vitro* allowing for the first time to decipher critical biochemical steps of the Tn3-family transposition mechanism (Nicolas et al. 2017).

I.6.4 Transpososome assembly and activation

The T+/I- TnpA mutants, including TnpA^{S911R} and TnpA^{3X}, have been purified and their activity was compared to that of TnpA^{WT} using biochemical assay reproducing different steps

of the transposition reaction *in vitro*. The results showed that all three proteins can bind to the transposon TIR end specifically. However, TnpA^{WT} only formed a single-end complex (SEC), whereas the T+/I- mutants TnpA^{S911R} and TnpA^{3X} spontaneously and cooperatively assembled a paired-end complex (PEC) in which two transposon ends are abutted by the transposase. Formation of the PEC correlated with catalytic activation of DNA cleavage and strand transfer reactions by TnpA. TnpA^{WT} could only form PEC on a pre-cleaved DNA end mimicking the initial nick produced by TnpA to start transposition (Nicolas et al. 2017).

Thus, the data support an asymmetric pathway for transpososome formation. In this pathway, a TnpA dimer first binds to a single end of the transposon, forming a SEC in which TnpA is in a "locked" inactive conformation (Fig. I.6.2A). TnpA then undergoes a conformational change that enables it to capture a second end to form an activated PEC in which TnpA is competent to make single strand break on DNA and catalyse strand transfer (Fig. I.6.2B/C).

The T+/- mutations of TnpA^{S911R} and TnpA^{3X} allow to overcome the activation barrier for PEC assembly and adopt the active conformation. By contrast, the signal that normally "unclock" TnpA^{WT} is still unknown. However, this signal is likely to be provided by the target in such a way that transposition will only be triggered when all partners of the reaction, the TnpA, the two transposon ends and a permissive target are brought together in the transposase. Since TnpA^{S911R} and TnpA^{3X} were initially identified for their immunity defect, there are less demanding with respect to the target activation signal, thereby explaining their hyperactive behaviour *in vitro* (Nicolas et al. 2017; Skumatov et al. in prep).

In addition to this, the D751N/S911R mutant (TnpA$^{D751N/S911R}$), which is thus mutated in one amino acid of the catalytic triad, is completely inactive and therefore cannot form PEC and catalyse the cleavage and strand transfer reactions. It is called the "catalytic mutant".

II. Objectives

Up to now, *in vitro* investigation on Tn4430 transposition mechanism was essentially performed using simplified donor and target DNA substrates that were assembled by annealing specific oligonucleotides (Nicolas et al. 2017, Guynet et al. 2010). In this thesis, new biochemical assays will be set up based on plasmid-derived DNA substrates that more faithfully reproduces natural conditions for transposition.

The starting point of the approach is a circular DNA molecule harbouring a "Mini-Tn4430" element with properly oriented ends. The recognition sites for the single-strand DNA (ssDNA) endonuclease Nt.BspQI has been introduced opposite to the TnpA cleavage sites. Since TnpA produces ssDNA nicks at the 3' ends of Tn4430, subsequent digestion of the reactions with Nt.BspQI generates double-strand DNA breaks that can be readily identified by standard agarose gel electrophoresis. So, the first objective of this thesis is to set up a biochemical assay to study the activities of TnpA.

Next, the new plasmid-based DNA substrates will be used to decipher the conditions that allow TnpA to make single-strand nicks at one or the other end of the transposon.

Different configurations of the donor plasmid (supercoiled or open circular form) will be incubated with TnpA and its mutant derivatives ($TnpA^{3X}$, $TnpA^{S911R}$ and $TnpA^{D751N/S911R}$) to determine whether the topology of the donor plasmid influences transposase activity.

In addition, it is important to determine whether the reaction generates unexpected forms of substrates.

Finally, as the transposase cleavage reaction is in *trans* (see I.4), the last objective of this thesis is to determine the requirement for correctly oriented transposon ends on the same DNA molecule for the cleavage reaction.

III. Results

III.1 Validation of the assay

Central to this project, is a fully designed artificial 3,9-Kb plasmid, pGIMF001, carrying a Mini-Tn4430 element delineated by properly oriented left and right 38-bp terminal inverted repeats of Tn4430 in their genuine context (TIRL and TIRR, respectively) (See V.1.1).

The first experiment was carried out to confirm that the cleavage assay works correctly. In this experiment, the supercoiled plasmid pGIMF001 was first incubated with the hyperactive triple mutant TnpA^{3X} for 4 hours. In the presence of the protein, part of the supercoiled form of the plasmid was converted to an open circular form (OC) which was not observed in the absence of protein, consistent with the cleavage activity of TnpA (Fig. III.1.1, left panel). Digestion of the reactions with Nt.BspQI confirmed that TnpA introduced specific ssDNA breaks at both ends of the transposon, generating a fragment corresponding to the plasmid backbone without Mini-Tn4430 (3270 bp) and a band corresponding to Mini-Tn4430 (683 bp) (Fig. III.1.1, right panel).

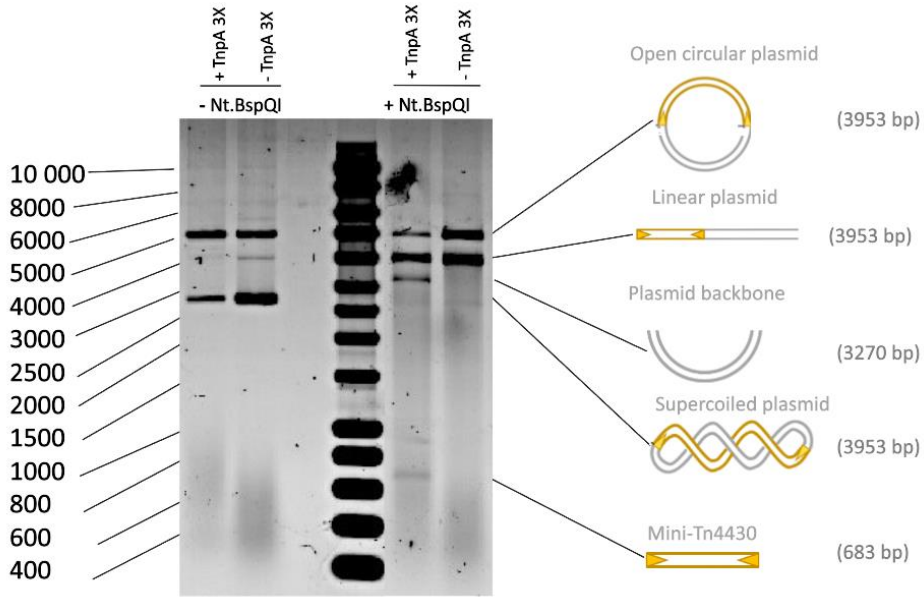


*Figure III.1.1: **Cleavage test between pGIMF001 (250 ng/µl) and TnpA^{3X} (125 nM).** 4 hours of incubation. 0.8% agarose gel. On the left panel, the plasmid is not digested with Nt.BspQI and the right panel the plasmid is digested with Nt.BspQI . The last lane of each panel is a control without TnpA^{3X}.*

After this first confirmation, it was necessary to test the activity of TnpA^{WT} as well as of the different mutants in order to determine the activity of each of them in the cleavage assay. In this experiment, the pGIMF001 was incubated with TnpA^{WT}, 2 hyperactive mutants

(TnpA^{3X}, TnpA^{S911R}) and the catalytic mutant, TnpA$^{D751N/S911R}$. Here the time of incubation was 4 hours. After digestion with Nt.BspQI, a fragment corresponding to the plasmid backbone without Mini-Tn4430 and a band corresponding to Mini-Tn4430 were clearly observed with TnpA^{3X} and TnpA^{S911R} (Fig. III.1.2, lane 2 and 3 respectively) whereas only a weak band corresponding to the backbone was detected with TnpA^{WT} (Fig.III.1.2, lane 4). As expected, the inactive mutant (TnpA$^{D751N/S911R}$) showed no cleavage activity (Fig.III.1.2, lane 5). This result confirms the "hyperactive" phenotype of TnpA^{3X} and TnpA^{S911R} compared to TnpA^{WT} as was demonstrated in previous biochemical assays (Nicolas et al; 2017). Due to its low activity, TnpA will not be used for further experiments, which will focus on the activity of TnpA^{3X} and TnpA^{S911R} to determine the requirements of the transposition reaction in terms of DNA substrates

Intriguingly, in addition to the expected cleavage products, additional bands (5 to 6) migrating between the Mini-Tn4430 fragment and the linear form of the substrate were observed in the reaction performed with TnpA^{3X}. The fact that these bands are not present in the control experiments performed with TnpA$^{D751N/S911R}$ or in the absence of transposase indicates that they arose from the specific activity of TnpA^{3X}. Although, the exact identity of these bands remains unclear at this stage, one possibility is that they resulted from strand transfer reactions catalysed by TnpA^{3X} as it will be discussed in a later section (see III.7).

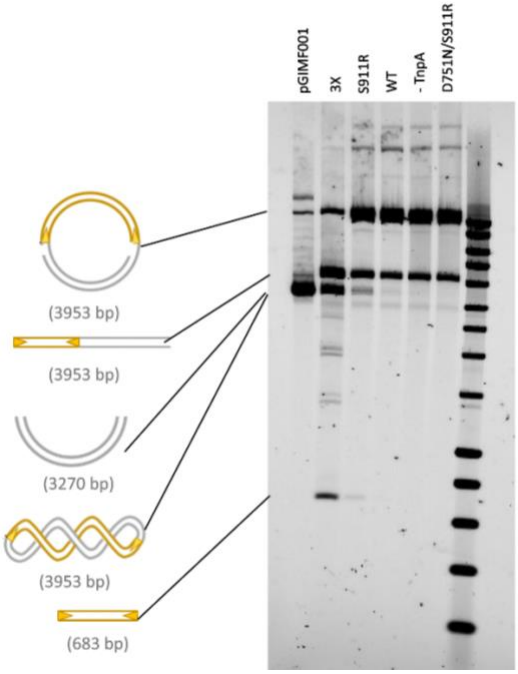


Figure III.1.2: ***Cleavage test between pGIMF001 (250 ng/µl) and TnpA^{3X} (300 nM), and TnpA^{S911R} (300 nM), and TnpA^{WT} (300 nM).*** *4 hours of incubation. 1,2% agarose gel. pGIMF001 is only the plasmid without treatment. The penultimate lane is a control without TnpA and the last lane is a control with TnpA$^{D751N/S911R}$ (300 nM).*

A time course analysis performed with TnpA^{S911R} (Fig. III.1.3) showed an increase in the intensity of the bands corresponding to the plasmid backbone (3270 bp) and Mini-Tn4430 fragment (683 bp), together with a decrease in the intensity of the OC form of the substrate a function of incubation time (Fig. III.1.3). This further demonstrates that the major product bands resulted well from TnpA-mediated specific cleavage at both transposon ends.

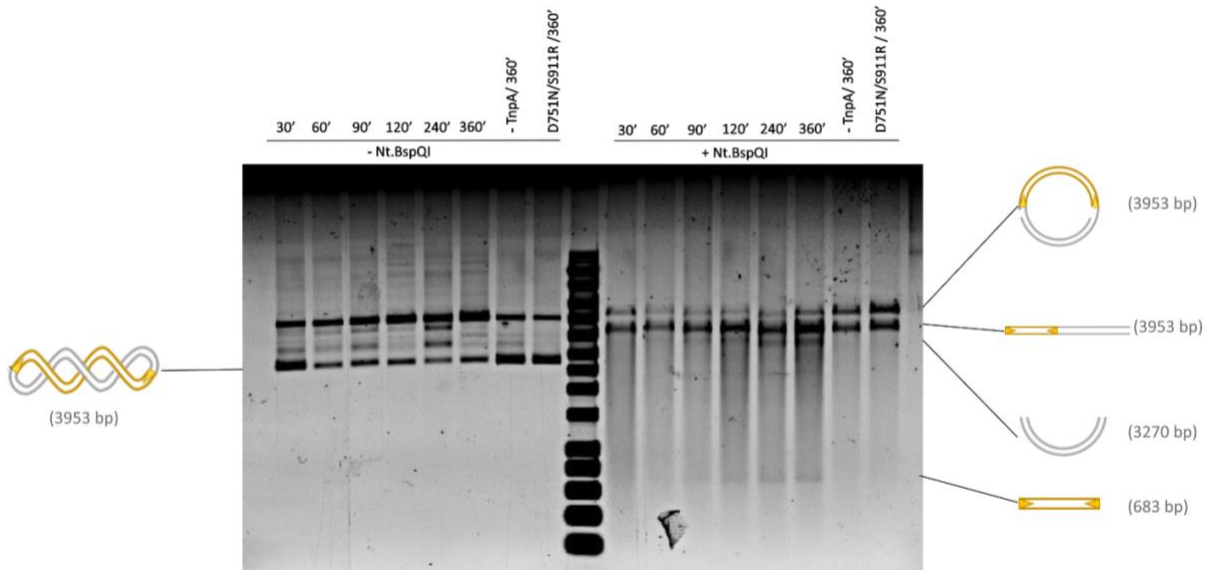


Figure III.1.3: ***Kinetic cleavage assay between pGIMF001 (250 ng/µl) and TnpA^{S911R} (300 nM).*** *Incubation time is noted in the legend. 0.8% agarose gel. In the left panel, the plasmid is not digested with Nt.BspQI and in the right panel, the plasmid is digested with Nt.BspQI . The penultimate lane of each panel is a control without TnpA and the last lane of each panel is a control with TnpA$^{D751N/S911R}$ (300 nM).*

III.2 Detection of single-end cleavage products

In the experiments of the above section, high levels of linearized plasmid DNA were observed in the reactions, as well as in the control experiments performed without transposase or with the catalytic mutant TnpA$^{D751N/S911R}$ (Fig.III.1.1 and Fig III.1.3). It was thus difficult to assess whether the linear products arose from specific TnpA cleavage at one or the other Mini-Tn4430 end, or from non-specific post-processing of Nt.BspQI products. Normally, ssDNA nicking by Nt.BspQI should only relax the plasmid giving rise to the open-circular form (OC). The problem could be due either to the DNA purification step (phenol extraction) performed after the cleavage assay, during which shear forces resulting from multiple pipetting and vortexing steps could create double-stranded breaks from single-stranded breaks, or to a problem with the Nt.BspQI enzyme that generates double-stranded breaks instead of single-stranded breaks.

To test these possibilities, supercoiled plasmid pGIMF001 was incubated with both TnpA^{3X} and TnpA^{S911R} and the reaction were cleaned with the NEB monarch purification kit instead of phenol extraction prior to Nt.BspQI treatment (Fig. III.2). In this case, the proportion of linear products generated by TnpA^{S911R} and TnpA^{3X} was clearly higher than in control experiments performed without TnpA or with the catalytic mutant TnpA$^{D751N/S911R}$, suggesting that Kit-based cleaning procedure reduce the level linear DNA and that Nt.BspQI has little or no inherent double strand cleavage activity. The Kit-based cleaning procedure was then routinely used for further experiments.

In addition, this experiment confirms that both hyperactive mutants of the TnpA proteins are capable of catalysing single-ended substrate cleavages and that cleavage at both ends are not necessarily concerted. The result also suggests that the activity of TnpA^{3X} appears to be higher than that of TnpA^{S911R} under the conditions tested.

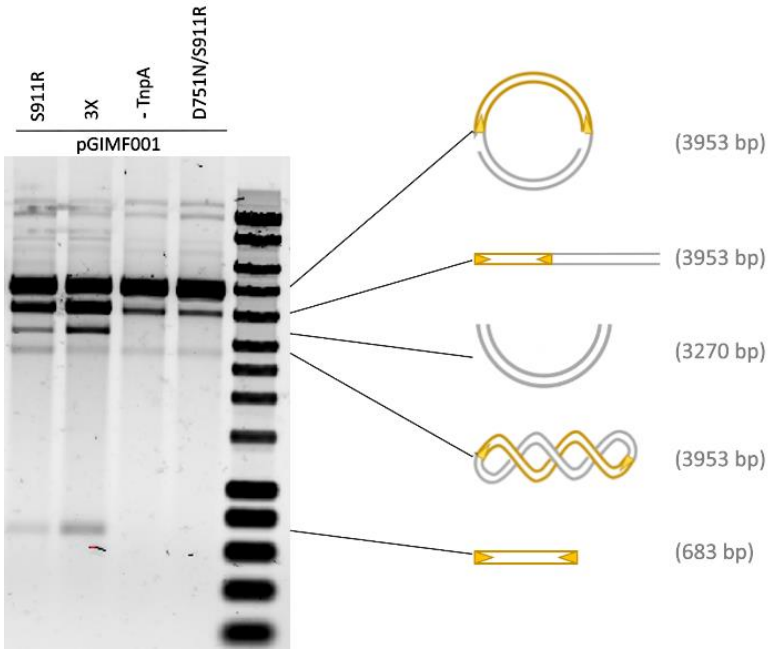


Figure III.2: ***Cleavage test between pGIMF001 (250 ng/µl) and TnpA^{S911R} (300 nM) or TnpA^{3X} (300 nM).*** *4 hours of incubation. 0.8% agarose gel. The penultimate lane is a control without TnpA. The last lane is a control with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*

III.3 Role of DNA super-coiling

For other transposable elements, DNA topology is an important player in the transposition reaction as it affects the formation of the transposition complex by enforcing correct juxtaposition of the transposon ends (see I.4.1). As a first step to investigate the role of DNA topology in Tn4430 transposition, the DNA cleavage activity of TnpA on supercoiled and relaxed DNA molecules was compared.

To this end, a relaxed form of pGIMF001 was generated by digesting the plasmid with the ssDNA endonuclease Nt.BsmAI (See V.1.1) and the resulting open-circular molecules were incubated with TnpA^{S911R} or TnpA^{3X}. Digestion of the reaction with Nt.BspQI showed that both transposases were able to cleave at one or both transposon ends, generating the corresponding linear, backbone and Mini-Tnn4430 products as previously observed (Fig. III.3.1). The O.C. form of pGIMF001 is therefore a substrate on which hyperactive TnpA mutants can be active.

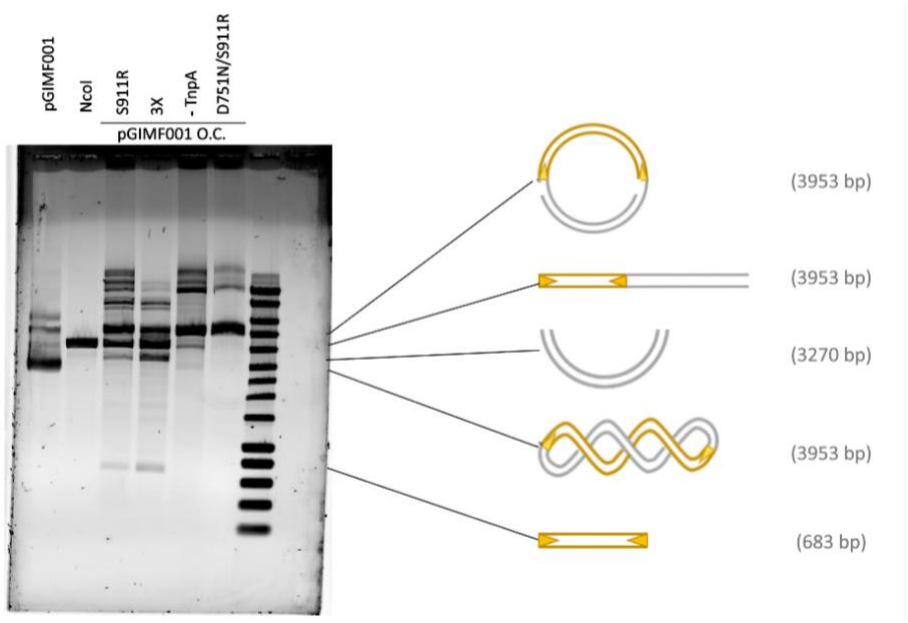


Figure III.3.1: ***Cleavage test using the open circular form (O.C.) of pGIMF001 (250 ng/µl) generated by Nt.BsmAI and different hyperactive TnpA mutants (300nM).*** *4 hours of incubation. 0.8% agarose gel. pGIMF001 is only the plasmid without treatment. NcoI is an enzyme that cuts pGIMF001 at a single site, generating a linear form of the plasmid. The penultimate lane is a control without TnpA and the last lane is a control with TnpA$^{D751N/S911R}$ (300 nM).*

To compare the activity of TnpA according to the topology of pGIMF001 a time course analysis was performed by incubating supercoiled pGIM001 or the open-circular form of the

plasmid with TnpA^{S911R} or TnpA^{3X} for increasing periods of time (from 0 to 240 min), and accumulation of the products was analysed by cutting the reaction with Nt.BspQI prior to agarose gel electrophoresis (Fig. III.3.2 and III.3.3).

For TnpA^{3X}, the analysis was repeated three times to ensure reproductivity (Fig. III.3.2, Appendix 1, Appendix 2). The gel in Figure III.3.2A was made with 1.2% agarose instead of 0.8%, which gave a better separation of the product bands, especially in the area between the cleaved backbone and the OC form of the substrate. In all cases, the same trend was observed showing that the open circular form of the plasmid reacted more than the supercoiled form, with a higher proportion of initial substrate (as visualised by Nt.BspQI-digested OC form of the plasmid) being converted into product at each time point.

For the gel shown in Fig. III.3.2A, this was confirmed by quantifying the relative amount of cleavage products compared to the substrate in each lane (Fig. III.3.2B and C). For the supercoiled substrate, products resulting from TnpA cleavage at one end (linear form) or both ends (plasmid backbone and Mini-Tn4430 fragment) accumulated progressively over time with about ~30% of the substrate being reacted after 240 min of reaction (Fig. III.3.2B). In contrast, when the O.C. form of the plasmid was used as the starting substrate, up to ~60% of the plasmid was cleaved after only 30 min of incubation. This correlated with a rapid accumulation of cleavage products at early stages of the reaction followed by a slower cleavage rate for prolonged incubation times (Fig. III.3.2C). Interestingly, quantification of the linear form of the plasmid showed a peak after 30 – 60 min of reaction when the plasmid backbone and the Mini-Tn4430 fragment reached some kind of a plateau (Fig. III.3.2C). This pattern might be interpreted as the result of sequential reaction during which TnpA first cleaves one end of the substrate and then the second end in a non-concerted manner. However, the experiment needs to be repeated on a shorter kinetics interval to verify this hypothesis.

Furthermore, additional bands that do not correspond to the supercoiled plasmid, the open circular form, the linear form, the backbone plasmid or the Mini-Tn4430 are observed on the gel (Fig. III.3.2A). The appearance of these additional bands will be discussed in a later section (see III.7).

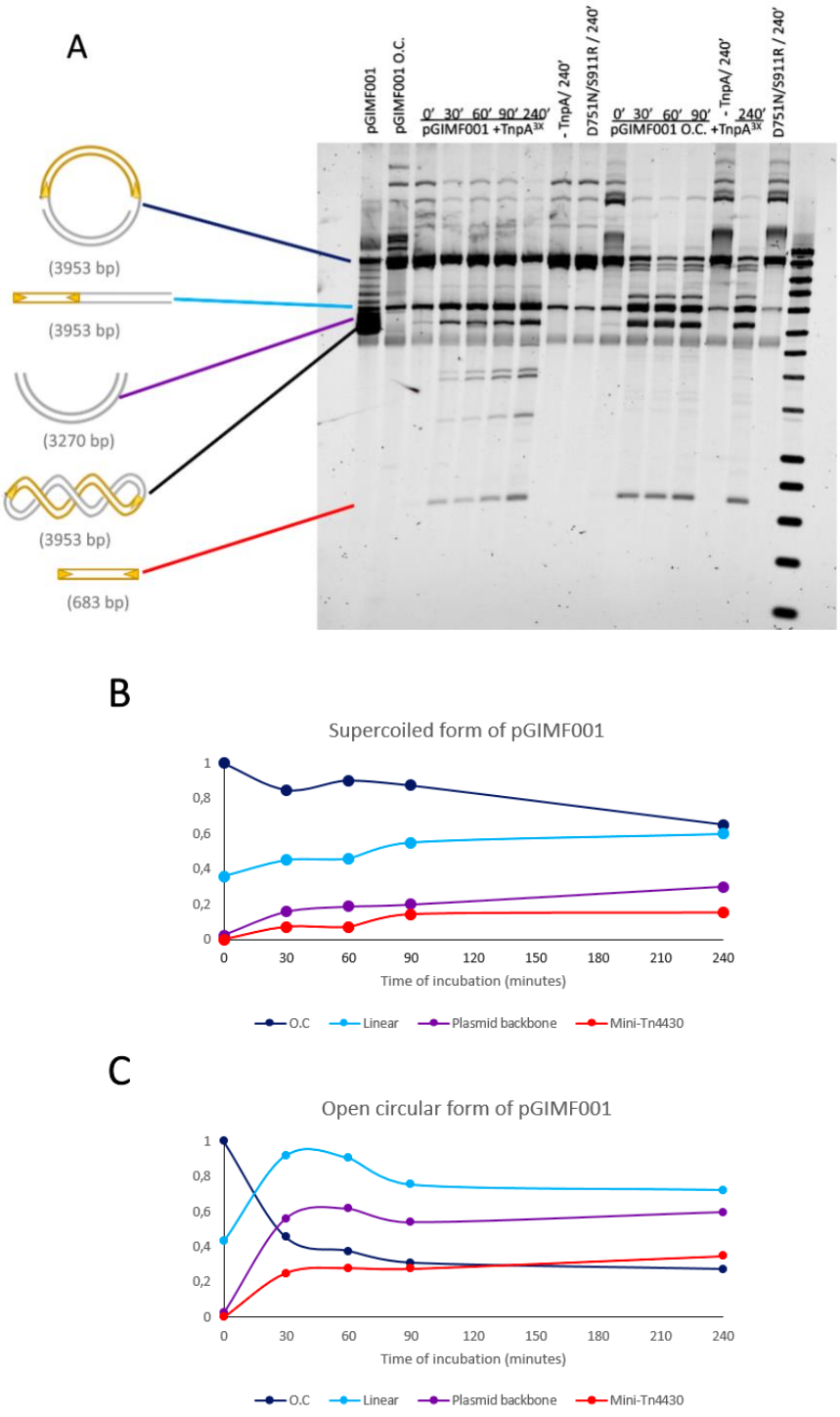


Figure III.3.2: ***(A) Kinetic cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM).*** *The incubation time is noted on the legend. 1.2% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. The eighth and fourteenth lanes are controls without TnpA. The ninth and last lanes are controls with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).* ***(B and C)*** *Quantification of TnpA^{3X} activity from the kinetic cleavage. Graph of the evolution of the quantity of the different substrates as a function of the incubation time. The graphs show the relative amount of the different cleavage products with respect to the initial quantity of the form O.C. at time 0 reported on 1. The starting substrate is the supercoiled plasmid pGIMF001 for the graph B and the starting substrate is the open circular plasmid pGIMF001 (O.C.) for the graph C. The dark blue curve represents the relative amount of the circular open form of the plasmid, the light blue curve represents the linear plasmid, the purple curve represents the backbone plasmid and the red curve represents the Mini-Tn4430. Quantification was performed with ImageJ software.*

Similar results were obtained with TnpA^{S911R} (Fig. III.3.3). As with TnpA^{3X}, time course analysis showed that the O.C. form of the plasmid was more reactive than the supercoiled form, with a greater proportion of the initial substrate being converted to product at each time point, and a faster and more extensive accumulation of TnpA-catalysed cleavage products over time. Here, with both forms of the substrate, the accumulation rate observed for the single end cleavage products (linear form of the plasmid) appeared to be faster than accumulation of products resulting from cleavage at both ends (plasmid backbone and Mini-Tn4430 fragment), which again, may suggest that cleavage at both ends may be sequential and not necessarily concerted.

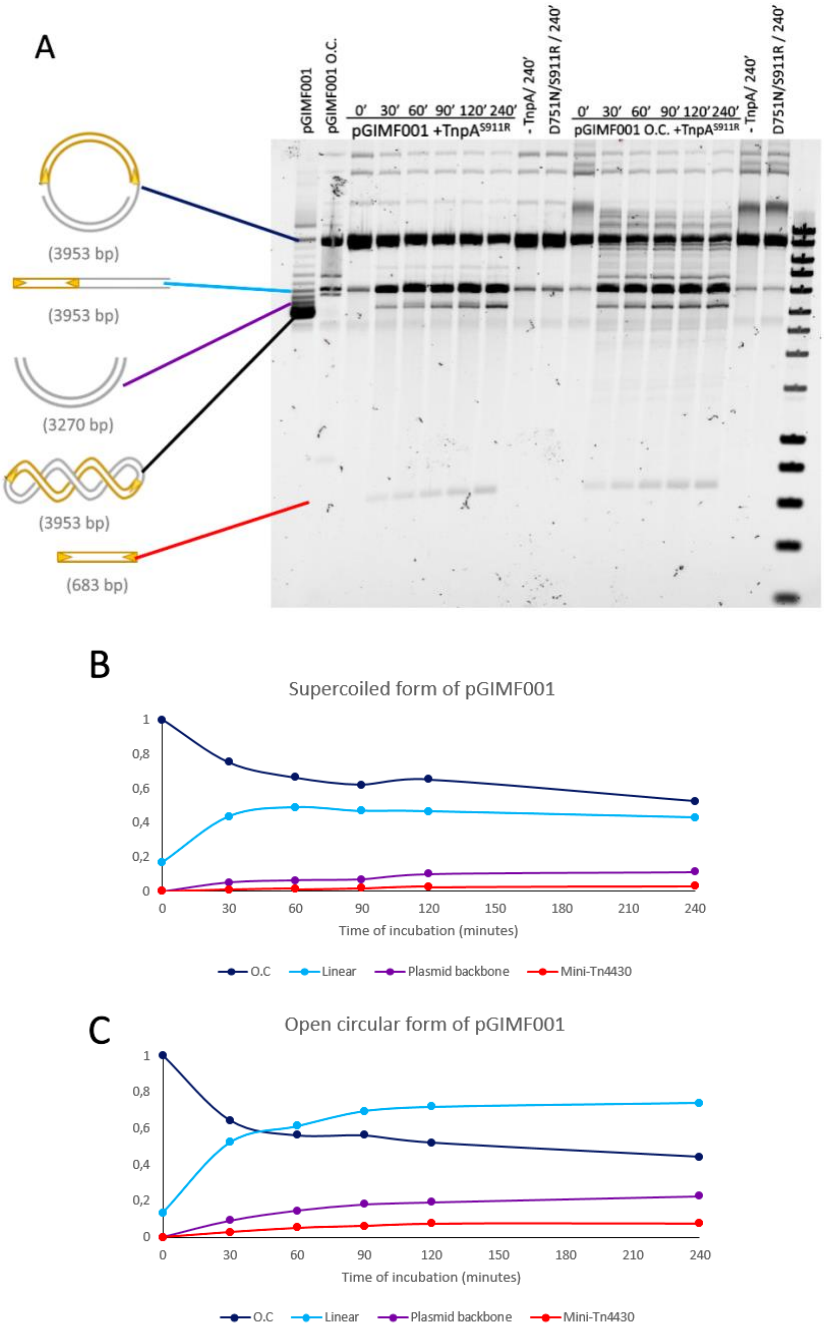


Figure III.3.3: ***(A) Kinetic cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{S911R} (300nM).*** *The incubation time is noted on the legend. 1.2% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. The ninth and penultimate lanes are controls without TnpA. The tenth and last lanes are controls with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).* ***(B and C)*** *Quantification of TnpA^{S911R} activity from the kinetic cleavage. Graph of the evolution of the quantity of the different substrates as a function of the incubation time. The graphs show the relative amount of the different cleavage products with respect to the initial quantity of the form O.C. at time 0 reported on 1. The starting substrate is the supercoiled plasmid pGIMF001 for the graph B and the starting substrate is the open circular plasmid pGIMF001 (O.C.) for the graph C. The dark blue curve represents the relative amount of the circular open form of the plasmid, the light blue curve represents the linear plasmid, the purple curve represents the backbone plasmid and the red curve represents the Mini-Tn4430. Quantification was performed with ImageJ software.*

Taken together these results show that the topology of the substrate influences the activity of the transposase, whether $TnpA^{3X}$ or $TnpA^{S911R}$, and that transposase is more active when the substrate is in open circular form rather than supercoiled. They also confirm that $TnpA^{3X}$ has a higher activity than $TnpA^{S911R}$ on both types of substrates.

III.4 The requirement for properly oriented transposon ends

In its natural configuration, Tn4430 is delineated by two IR sequences in inverted orientation. This topological organisation of the transposon ends may be important to allow TnpA to properly bind one end to form the SEC and then to undergo the required conformational change to capture the other end and form the PEC (see I.6.4). This correct orientation of the transposon ends could be even more crucial as cleavage of the ends is done in *trans,* with the TnpA monomer bound to one end catalysing the reaction at the opposite end. Therefore, if one end is missing or misaligned, this could disrupt the transposition.

To test whether the presence of 2 inversely oriented IR on the same DNA molecule is important for transposition two PGIMF001 derivatives containing a single transposon end (pGILC001, and pGILC002) were constructed by deleting a 76-bp fragment containing the left or right IR end of Mini-Tn4430, respectively (Fig. III.4.1, see also V.1.1). To this end, pGIMF001 was digested with compatible restriction enzymes XbaI and NheI to construct pGILC001, and with BamHI and BglII to construct pGILC002 (Fig. III.4.1). Both constructs were then ligated and transformed into thermo-competent cells. The cells were then cultured and plasmid candidates were purified from isolated colonies.

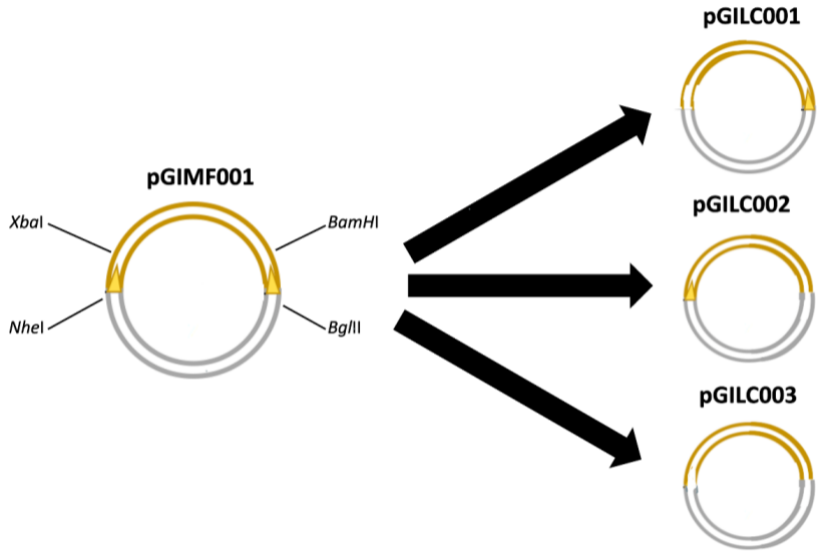


Figure III.4.1: ***Representation of the 3 constructions with a single TIR and without TIR.*** *pGILC001 was created from pGIMF001 digested with NheI and XbaI, pGILC002 was also generated from pGIMF001 but this time digested with BamHI and BglII and pGILC003 was generated from pGILC001 and digested with BamHI and BglII.*

Deletion of the 76-bp fragments containing the left and right IR was verified by digesting, pGILC001 candidates with *Sal*I and *Sph*I and pGILC002 candidates with *Nco*I and *Xho*I, respectively (Fig. III.4.2).

After deletion, digestion of pGILC001, should give a shortened *Sal*I-*Sph*I fragment of 345 bp instead of 421 bp in the native plasmid pGIMF001, whereas digestion of pGILC002 is expected to give a *Nco*I-*Xho*I fragment of 214 bp instead of 290 bp. For pGILC001, all the analysed clones had the expected profile, whereas 4 out of 5 pGILC002 candidates were as expected, clone 21 presumably containing an unmodified version of the initial pGMF001 plasmid (Fig. III.4.2)

For the following experiments, purification of pGILC001 was done from clone 5 and purification of pGILC002 was done from clone 13.

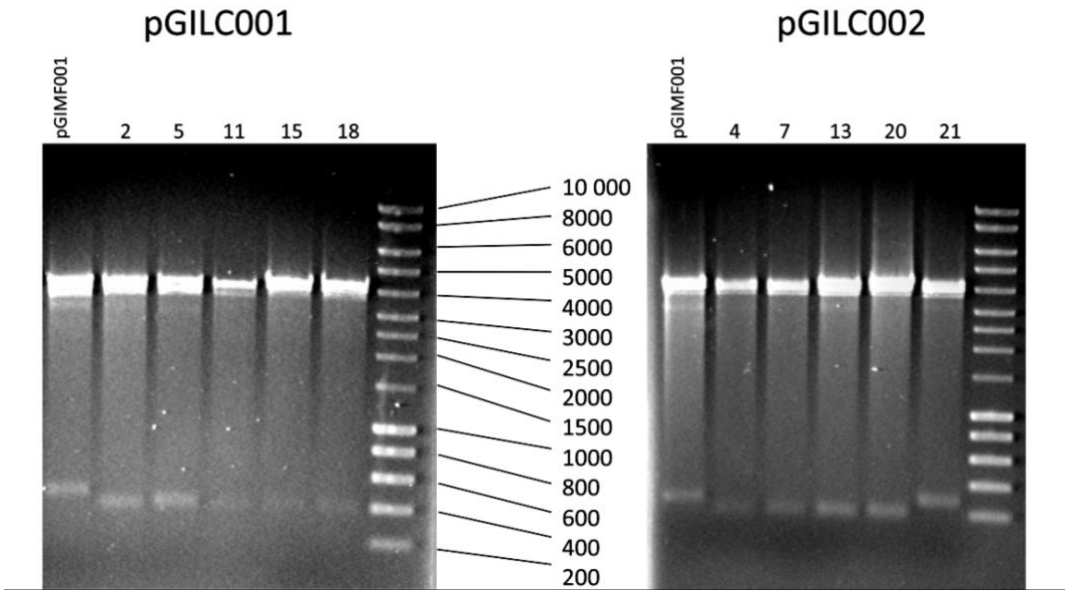


Figure III.4.2: ***Confirmation of pGILC001 and pGILC002 by restriction.*** *pGILC001 (250 ng/µl) was digested with SalI and SphI and pGILC002 (250 ng/µl) with NcoI and XhoI. The number corresponds to the number of the clone from which the purification comes. In the column pGIMF001 the starting substrate is pGIMF001 native (250 ng/µl) and it is a control.*

A cleavage test was performed by incubating the supercoiled form of both single IR end plasmids pGILC001 and pGILC002 with TnpA^{3X}, TnpA^{S911R} and TnpA^{WT} and by cutting the reactions with Nt.BspQI as was one previously for the Mini-Tn4430 containing substrate pGIMF001 (Fig. III.4.3). For both plasmids, reactions performed with the hyperactive mutants TnpA^{3X} and TnpA^{S911R} showed a marked decrease in the amount of initial substrate and a corresponding increase in the amount of the linear form of the plasmid when compared to the reactions performed with the catalytic mutant TnpA$^{D751N/S911R}$ or in absence of TnpA (Fig. III.4.3). This is consistent with TnpA cleavage occurring at a specific position of the substrates,

in front of the Nt.BspQI nicking site. In addition, both TnpA mutants produced additional bands migrating ahead of the linear form of the plasmids. Since these bands were not observed in the control experiments, they likely resulted from further processing reactions requiring the catalytic activity of TnpA such as strand transfer. The fact that the pattern of bands is different for pGILC001 and pGILC002 indicates that the specificity of these reactions is not the same with both plasmids. No apparent cleavage or other processing products were detected with $TnpA^{WT}$ consistent with its lower or less "promiscuous" activity compared to the hyperactive mutants. Thus, these results suggest that even though catalysis is thought to occur *in trans*, the presence of a single IR end on a DNA substrate is sufficient to promote specific cleavage and by TnpA.

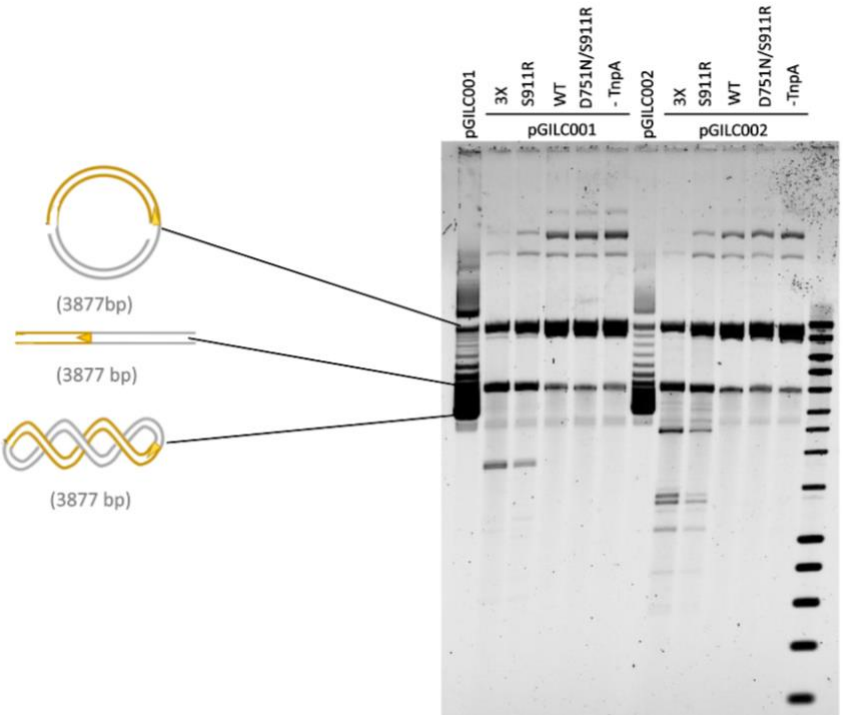


Figure III.4.3: ***Cleavage test between pGILC001 (250 ng/µl) and pGILC002 (250 ng/µl) and $TnpA^{3X}$ (300 nM), and $TnpA^{S911R}$ (300 nM), and $TnpA^{WT}$ (300 nM).*** *4 hours of incubation. 1,2% agarose gel. pGILC001 and pGILC002 are only the constructions without treatment. The sixth and last lanes are a control without TnpA and the fifth and penultimate lanes are a control with $TnpA^{D751N/S911R}$ (300 nM).*

To totally rule out the possibility that the products observed with the single end substrates resulted from non-specific activity of TnpA, a cleavage assay was performed on a plasmid equivalent to pGIMF001 from which both Tn4430 ends have been deleted. To generate this plasmid (pGILC003), pGILC001 was digested with *BamH*I and *Bgl*II to remove the 76 bp fragment containing the right IR end (Fig. III.4.1). After ligation and transformation into *E. coli*, plasmid DNA was purified from individual clones and analysed by digestion with *Sac*I and *Nco*I (Fig. III.4.4).

If the construction is correct pGILC003 should carries a 152-bp deletion compared to the original plasmid pGIMF001 and *Sac*I/*Nco*I digestion of the plasmid should give a corresponding band of 708 bp instead of 860 bp (see V.2.2 and Appendix 4). Only the plasmid extracted from clone 15 showed the expected pattern, the other candidates likely contained the undigested pGILC001 plasmid.

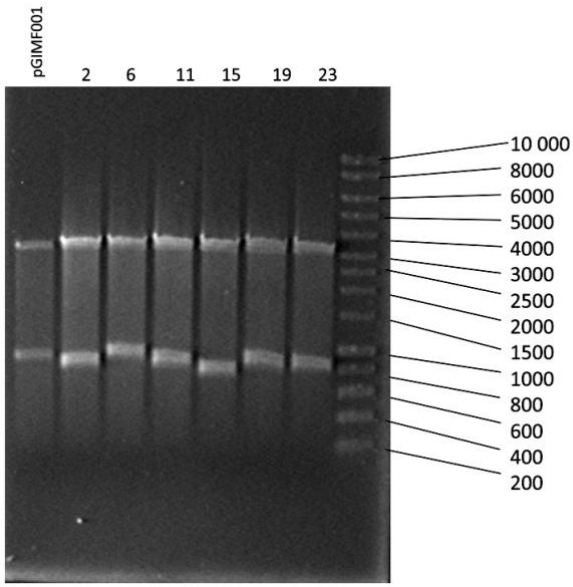


Figure III.4.4: ***Confirmation of pGILC003 by restriction.*** *pGILC003 (250 ng/µl) was digested with SacI and NcoI. The number corresponds to the number of the isolated colony from which the purification comes. In the column pGIMF001 the starting substrate is pGIMF001 native (250 ng/µl) and it is also digested with SacI and NcoI.*

Standard cleavage assay performed with the IR-less plasmid pGILC003 and TnpA^{3X}, TnpA^{S911R} and TnpA^{WT} showed no detectable difference with the control reactions performed with the TnpA$^{D751N/S911R}$ or in absence of TnpA (Fig. III.4.5). The amount of the initial substrate did not decrease and the amount of the linear form was at the same background level as detected in absence of active TnpA (Fig. III.4.5). No additional band was observed demonstrating that the processing products that were generated by TnpA^{3X} and TnpA^{S911R} on the single- or double-IR end substrates resulted well from the specific activity of TnpA.

Thus, this result shows that a IR-less plasmid cannot be recognised as a substrate by that TnpA, and that the IR sequence contains the specific determinants for TnpA-catalysed DNA cleavage and other subsequent steps of transposition.

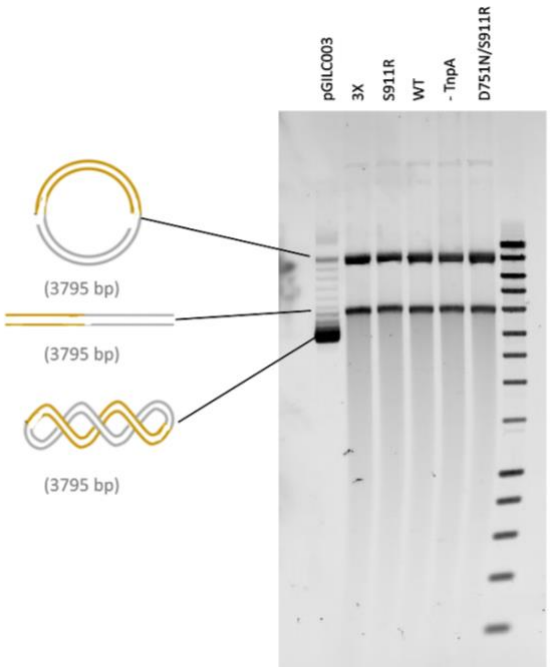


Figure V.4.5: ***Cleavage test between pGILC003 (250 ng/µl) and TnpA^{3X} (300 nM), and TnpA^{S911R} (300 nM), and TnpA^{WT} (300 nM).*** *4 hours of incubation. 1,2% agarose gel. pGILC003 is only the constructions without treatment. The penultimate lane is a control without TnpA and the last lane is a control with TnpA$^{D751N/S911R}$ (300 nM).*

III.5 The role of the cofactor

In all previous experiments, the cleavage test was performed with Mn^{2+} as cofactor. This cofactor is likely not the natural cofactor of transposase, which is thought to be Mg^{2+}. However, replacement of Mg^{2+} by Mn^{2+} was shown to hyperactivate TnpA in different biochemical assays as is often the case with other polynucleotidyl transferases (Nicolas et al., 2017). It was therefore interesting to test the activity of the transposase in our new cleavage assay according to the cofactor used. A cleavage assay was therefore performed on supercoiled and O.C. forms pGIMF001 in the presence of Mn^{2+} or Mg^{2+}. $TnpA^{3X}$ was used here as it showed the highest activity in previous experiments (*e.g.*, see III.3).

With both forms of the plasmid, the cleavage activity of $TnpA^{3X}$ was higher when the cofactor is Mn^{2+} compared to its activity in the presence of Mg^{2+} as revealed by the consumption of the Nt.BspQI-cut (O.C.) form of the substrate after 4 hours of reaction (Fig. III.5). In addition, the final amount of linear plasmid, backbone plasmid and Mini-Tn4430 fragment was also higher when the cofactor was Mg^{2+} for both the O.C. and supercoiled substrate (Fig III.5). Furthermore, additional bands that do not correspond to expected cleavage products were observed in presence of Mn^{2+}, but not in the Mg^{2+} reactions (Fig. III.5). Again, the appearance of these additional bands will be discussed in a later section (see III.7). Thus, this result confirms that the activity of TnpA can be influenced by the cofactor, and that replacing the natural co-factor Mg^{2+} by Mn^{2+} tends to hyperactivate TnpA and/or to make it more promiscuous.

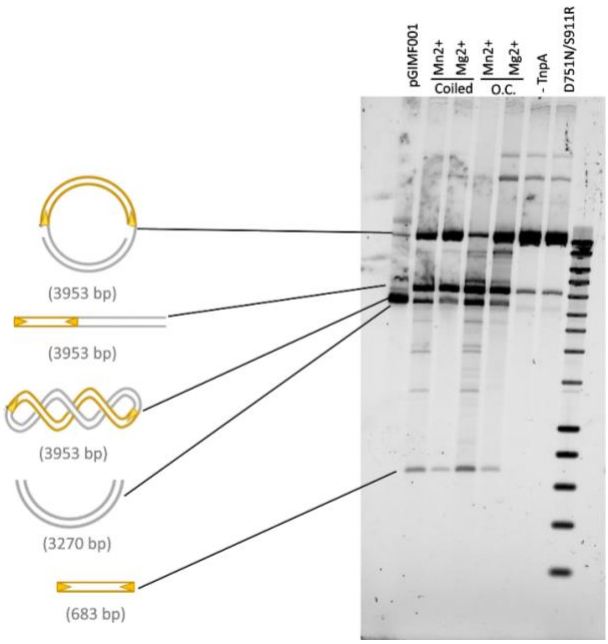


Figure III.5: ***Cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM) in presence of manganese and in presence of magnesium.*** *4 hours of incubation. When the supercoiled form is used it is indicated coiled and when the open circular form is used it is indicated O.C. 1.2% agarose gel. pGIMF001 is only the plasmid without treatment. The penultimate lane is a control without TnpA. The last lane is a control with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*

III.6 The role of glycerol

Standard cleavage assays were performed in the presence of 22% glycerol. Previous experiments have shown that the presence of glycerol increases the cleavage activity of TnpA *in vitro* (Aryanpour, unpublished results). Here, a cleavage assay was therefore performed in the presence and absence of glycerol to determine whether glycerol has an impact on TnpA activity in presence of different DNA substrates. TnpA^{3X} was again used here as it showed better activity than TnpA^{S911R} (see III.3).

Here, the activity of TnpA^{3X} appears to be higher in the presence of glycerol with the supercoiled form of pGIMF001 because the amount of O.C. form observed is lower and the amount of backbone plasmid and Mini-Tn4430 is higher (Fig. III.6). However, when the starting substrate is the plasmid in the O.C. form, no difference in activity is observed when glycerol is absent. There is no difference in intensity between the bands corresponding to the circular open form of the plasmid, the linear form of the plasmid, the plasmid backbone and Mini-Tn4430.

This shows that glycerol and therefore the test conditions will help transposase when the plasmid is supercoiled but not when the plasmid is in the O.C. form.

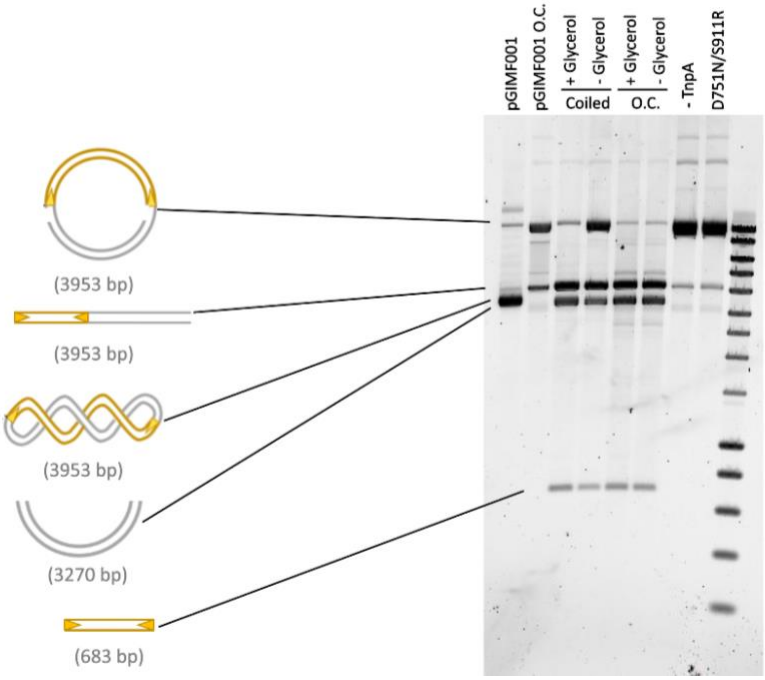


Figure III.6: ***Cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM) in presence (+) and in absence (-) of glycerol.*** *4 hours of incubation. When the supercoiled form is used it is indicated coiled and when the open circular form is used it is indicated O.C. 1.2% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. The penultimate lane is a control without TnpA. The last lane is a control with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*

III.7 Strand transfer products

In the previous experiments, some results were intriguing as additional bands appeared on some gels, for example, in figure III.1.2, figure III.3.2A and figure III.4.3. It is therefore interesting to determine whether the appearance of these bands is reproducible and what these additional bands are.

Firstly, an experiment was conducted to find out whether the appearance of these bands was reproducible. This experiment was a cleavage assay with $TnpA^{3X}$ as it has a better activity than $TnpA^{S911R}$ with pGIMF001 (see III.3) but this time the experiment was done in the presence and absence of Nt.BspQI to see if the profile of the bands changed if the single-stranded DNA endonuclease was not used (Fig. III.7.1).

Here too, additional bands appear on the gel. When Nt.BspQI is used, the additional bands present on the gel in figure III.1.2 and figure III.3.3 are also present on this gel (Appendix 4). However, here other additional bands of lower intensity also appear (Fig III.7.1). Following the use of Nt.BspQI, the band profile is almost the same regardless of the starting substrate, either the supercoiled pGIMF001 or its O.C. form. However, 3 more bands are present when the starting substrate is the O.C. form. (Appendix 4, blue rectangle). These bands were also observed previously on another gel (Fig. III.3.3). Furthermore, the intensity of the bands varies with the starting substrate. The additional bands seem to be more intense with the supercoiled form than with the O.C. form (Fig. III.7.1). In comparison, when the Nt.BspQI is not used the band profile is completely different (Fig. III.7.1). Under these conditions this is normal as TnpA only makes single strand breaks on each end of Mini-Tn4430 and therefore it is not possible to observe the plasmid backbone and Mini-Tn4430 on the gel. However, additional bands appear. In this case, the profile of the additional bands is also different depending on the topology of the substrate used (Fig. III.7.1).

These results suggest that these additional bands may be the result of specific intermolecular strand transfer. This hypothesis is supported by the fact that the appearance of these bands is reproducible and that when using Nt.BspQI the band profile does not change depending on the topology of the starting substrate, except for the 3 bands that appear with the O.C. form of the plasmid. These 3 additional bands can be explained by a strand transfer

which is more favourable with the O.C. form and which does not occur or occurs very little with the supercoiled form.

In addition, the band profile is different when the Nt.BspQI is not used. Under these conditions, the band profile varies according to the topology of the starting substrate. This is also consistent with our hypothesis that when no endonuclease is used, the strand transfer products will be different since Nt.BspQI does not make single strand breaks and therefore does not alter the form of the strand transfer products.

If it is indeed a strand transfer phenomenon, this phenomenon would be specific because here the bands observed are always the same. Some additional bands never observed and of lower intensity appear nevertheless (Fig. III.7.1). They could also be due to a specific strand transfer, but this happens less often as the transfer would be less favourable.

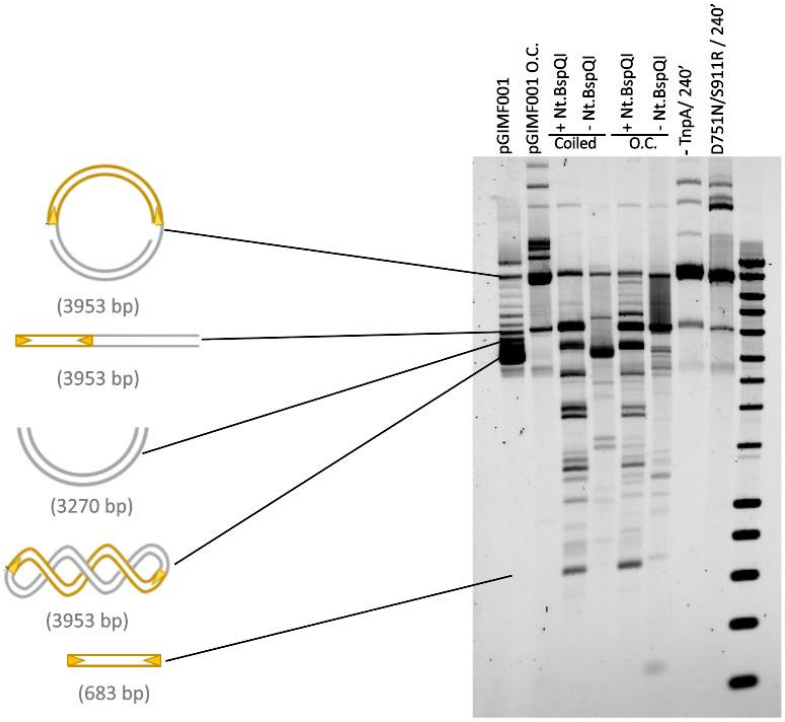


Figure III.7.1: ***Cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM) in presence (+) and in absence (-) of Nt.BspQI.*** *4 hours of incubation. When the supercoiled form is used it is indicated coiled and when the open circular form is used it is indicated O.C. 1.2% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. The penultimate lane is a control without TnpA. The last lane is a control with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*

These results support the idea that these additional bands can be due to specific strand transfer. Another experiment was conducted to provide further evidences. This experiment will be identical to the previous one but after the cleavage test, the reaction product will also be digested with 3 restriction enzymes (*Xho*I, *EcoR*I and *Sph*I). Each of these three enzymes will make a double strand break in the DNA at a specific site. The enzyme *Xho*I will make a

double strand break in Mini-Tn4430, *EcoR*I will make a break to the right of Mini-Tn4430 and *Sph*I will make a break to the left of Mini-Tn4430 (see V.1.1). This digestion is performed to determine whether or not the pattern of bands varies and thus to see if these bands could be due to specific intramolecular strand transfer.

The first thing that can be observed here is that the additional strong bands visible on the previous gels (Fig. III.7.1) are also visible here when using Nt.BspQI (Appendix 4, Fig. III.7.2). When no Nt.BspQI is used, the additional band profiles with the supercoiled and O.C. form are also the same as in Fig. III.7.1 (Fig. III.7.2). This again confirms that our results are reproducible. However, when the starting substrate is the O.C. form of the plasmid and the product is not treated with Nt.BspQI, there is a difference in intensity in the additional bands observed. This shows that although strand transfer appears to be specific, there may still be a variation in the amount of strand transfer products (sixth lane, Fig. III.7.2).

Here the *Sph*I enzyme did not work properly. This can be confirmed because on the gel, when this enzyme is used, there is still some of the O.C. form of the plasmid practically like in the control (Fig.III.7.2).

With the *EcoR*I enzyme, the same profile is observed regardless of the starting substrate. However, this profile is completely different from the profile obtained previously when using Nt.BspQI (Fig. III.7.2). This shows that the profile is independent of the topology of the starting substrate and will depend on the restriction enzyme used. As *EcoR*I will cut to the right of our Mini-Tn4430, it will generate different strand transfer products than those observed after restriction by Nt.BspQI, which explains this different profile. Nevertheless, when the starting substrate is the supercoiled pGIMF001, the enzyme has not completely digested the plasmid because here too the plasmid can still be observed in its O.C. form (Fig.7.2).

With the *Xho*I enzyme, the profile of the bands is quite similar if we compare the profile of the supercoiled form and the O.C. form. However, we can observe that there are 2 additional bands with the O.C. form. (Fig. III.7.2). These bands correspond to a fragment of more or less 800 bp and another of more or less 3500 bp. The appearance of this band can be explained by a star activity of Nt.BsmaI. We know that *Xho*I cleaves the plasmid at base 3134 and if there is a double-stranded cleavage by Nt.BsmaI at base 3950, this will generate a fragment of 816 bp and another 3137. This corresponds to the fragment observed on the gel. However, the other two most pronounced bands are the same for the supercoiled form and

the O.C. form. These two bands migrate to the same position as two bands observed when treating the cleavage products with Nt.BspQI (lane 2 and 7, Fig. III.7.2). The enzyme *Xho*I, has its restriction site in Mini-Tn4430 and if the additional bands are indeed strand transfer products and these products are cleaved by *Xho*I we obtain the same products as after cleavage with Nt.BspQI. This is another indication that these bands could correspond to specific strand transfer products because when the products are cleaved into the transposon by *Xho*I the same bands are obtained as some bands produced by cleavage with Nt.BspQI.

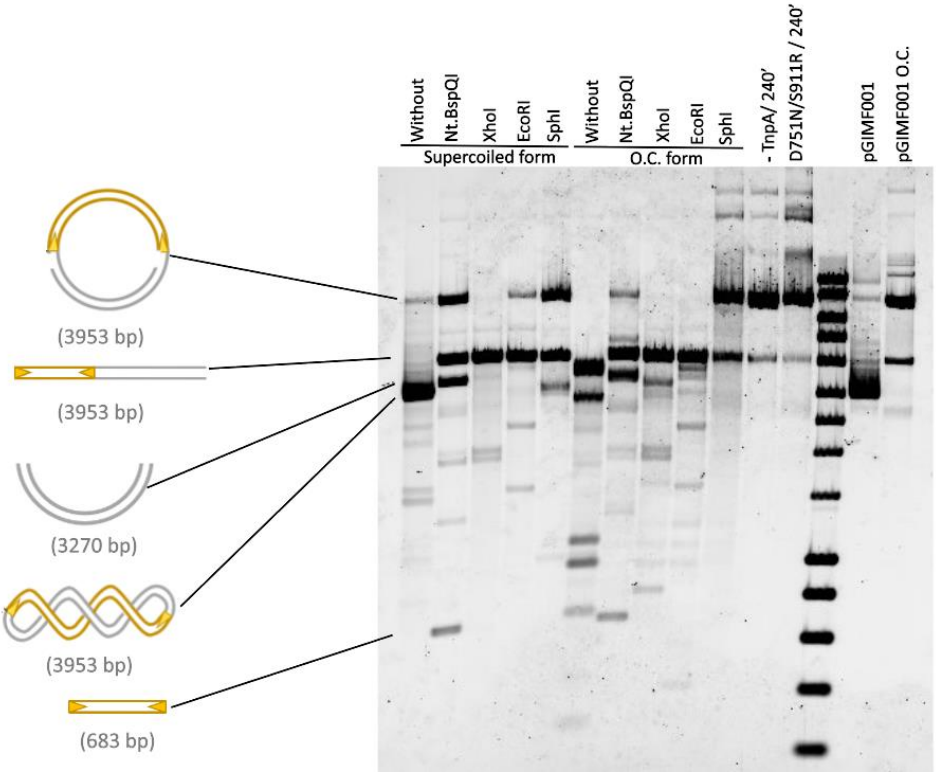


Figure III.7.2: ***Cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM) in presence and in absence (Without) of Nt.BspQI and restricted with different restriction enzymes.*** *4 of incubation. 1.2% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. The eleventh lane is a control without TnpA. The twelfth lane is a control with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*

In figure III.4.3, additional bands are also observed. The band that can be observed with pGILC001 is the same one that has been observed on other gels (Appendix 5, Fig. III.3.2A, Fig. III.7.1, Fig. III.7.2). For pGILC002, it is the same as observed on other gels (Appendix 5, Fig. III.3.2A, Fig. III.7.1, Fig. III.7.2). If these bands are found to be strand transfer products, the additional band with pGILC001 would be due to strand transfer from the right end of the transposon because the other end has been removed. For pGILC002, the additional bands observed would therefore be strand transfer products from the left end of the transposon

because the right end of the transposon has been removed. Nevertheless, this remains a hypothesis and must be proven.

In addition, when magnesium is used as a cofactor instead of manganese, there is a clear decrease in the appearance of additional bands and therefore perhaps a decrease in strand transfer. This may be due to the decrease in transposase activity due to the change of cofactor (Fig. III.5).

In summary, the different experiments in this section provide some clues that these additional bands could be strand transfer products because these bands are reproducible and can vary when using different enzymes. In addition, the experiment with a plasmid with a TIR indicates that some of the bands could be from strand transfer on either side of the transposon. However, at this stage, this remains a hypothesis. The results still need to be repeated and further experiments performed to determine whether these bands correspond to specific strand transfer products.

IV. Discussion and perspectives

During this master thesis, a new biochemical assay was set up and improved to study the activity of the Tn4430 transposase, TnpA, on circular DNA molecules containing a mini-derivative of the transposon that prefigures its natural organization. The analysis of the reaction has been improved to detect TnpA-mediated DNA cleavage at either or both ends of the element. A further improvement to the analysis would be to cleave the reaction products with an additional restriction enzyme that cuts asymmetrically to Mini-Tn4430 to discriminate cleavage at either end.

This assay was used to compare the activity of the hyper-active TnpA^{3X} and TnpA^{S911R} mutant on different substrates, and could potentially be used to characterise other variants of the transposase, including TnpA^{WT}.

The data indicates that TnpA activity is influenced by DNA topology, its activity with relaxed substrates appears to be higher than with supercoiled substrates. This result is interesting because it is in contrast to the topological filter model which explains that DNA supercoiling facilitates the assembly of the transposition complex (see I.4.1). In Tn4430, it seems that this topological filter does not apply. This could be explained by the fact that TnpA traps a positive DNA cross (see I.6.2) while the plasmid supercoil is negative. During the formation of the transposition complex with a supercoiled plasmid there is therefore a change from a negative to a positive cross, which is energetically unfavorable. Whereas, when our plasmid is in relaxed form the creation of a positive cross does not require energy. This is why we could observe a higher activity of hyperactive mutants with a relaxed plasmid. Thus, the system observed here limits transposition with a supercoiled substrate and this system could be a mechanism for regulating transposition at the level of the donor DNA molecule.

On both supercoiled and open circular substrates, TnpA^{3X} showed higher activity than TnpA^{S911R} with pGIMF001, its O.C. form and pGIAR045 (see V.1.2 ,Appendix 3).

The activity of TnpA can also be influenced by the plasmid itself. The activity of TnpA^{3X} and TnpA^{S911R} is higher with pGIAR045 (Appendix 3) than with pGIMF001 (Fig. III.3.2 and Fig. III.3.3). Thus, the size of the plasmid, the sequence of the plasmid, the sequence of the Mini-Tn4430, the size of the Mini-Tn4430 or several of these factors could impact on the transposase activity. The flanking sequence of the TIR could also have an impact on transposase activity as they also interact with transposase during the formation of the

transposition complex (see I.6.2). This point deserves further investigation to determine which factors actually have an impact on transposase activity.

The results indicate that TnpA^{3X} can be active on a plasmid with only one TIR. This result is particularly interesting as it is known that the cleavage reaction is in *trans* but that even with one end of the transposon cleavage on one side can take place. However, as it seems that the topological filter is not applicable in Tn4430, it could be that transpososome formation also occurs between 2 different plasmids each containing one TIR. Furthermore, when no TIR is present the transposase cannot recognise the transposon and therefore there is no cleavage activity.

The experiments also showed that the conditions under which the experiment is conducted can impact on transposase activity. In particular, the cofactor used will have an impact on the activity of TnpA. Mn^{2+} seems to make transposase more active and/or promiscuous. The presence of glycerol will also have an impact on the reaction when the starting substrate is supercoiled. With this substrate and without glycerol, the transposase is less active, probably due to the fact that the transition from a negative to a positive DNA cross during transpsosome formation is energetically unfavourable (see above). Glycerol has a "solvophobic" effect (Moelbert et al. 2004) and will stabilise the formation of the transpososome (Aryanpour, unpublished results) which will have an impact on the activity of the transposase with the supercoiled plasmid. However, the reaction with the relaxed form of pGIMF001 appears to be more favourable and therefore glycerol does not visibly impact transposition.

In this thesis, a very interesting activity was observed. This activity corresponds to the appearance of additional bands on the gels and these bands do not correspond to bands we would normally observe. The experimental data suggest that there may be specific intramolecular strand transfer *in vitro* in the cleavage assay and the additional bands observed may be the products of these strand transfers. These products could have many different topologies. It could result from single or double strand transfer. Furthermore, this transfer could also be *cis* or *trans*. Finally, the transfer products may also undergo disintegration (Nicolas et al. in prep). All these topologies will be different again if the starting substrate is a supercoiled or open circular plasmid. There is therefore a large number of possible topological products and for the moment it is difficult to determine the topology of our obtained

substrates. This phenomenon has not been observed before and deserves further study to ensure that specific strand transfer is indeed occurring.

In the future, to confirm our hypothesis that due to the DNA positive cross trapped by the transposase, TnpA^{3X} and TnpA^{S911R} are less active with a negatively supercoiled plasmid than with its relaxed form. The same experiment should be performed with a positively supercoiled form of pGIMF001 and if the transposase activity is the same or higher than with our relaxed plasmid, this will confirm our hypothesis.

It would also be interesting to perform a cleavage assay with the plasmid pGIAR045 and its O.C. form in order to compare the activity of TnpA^{3X} and TnpA^{S911R} on these 2 substrates and then perform quantifications from these gels. This experiment would be interesting to determine whether the topology of pGIAR045 also influences the activity of hyperactive TnpA mutants and thus confirm or refute the previous results.

Next, the experiments already performed can be repeated but this time with misaligned ends, i.e. in direct orientation, to see if the presence of two ends in reverse orientation on the same DNA molecule facilitates the assembly of the active complex and the cleavage activity.

Another experiment that can be performed is a cleavage assay, but this time using Nb.BtsI or Nb.BsrDI digested pGIMF001 to mimic TnpA mediated terminal cleavages. This will be done to determine whether TnpA-mediated cleavage at one end promotes PEC formation and cleavage at the other end, particularly with TnpA^{WT}. These 2 enzymes are two ssDNA endonucleases whose cleavage sites have been positioned at the 3' ends of Mini-Tn4430 to mimic TnpA-mediated end cleavages (see V.1.1).

The cleavage assays could be repeated with the O.C. form of pGIMF001 and the plasmid pGIAR045 and its O.C. form with magnesium as cofactor this time. This will be done in order to determine whether this change of cofactor also has an influence on the activity of the transposase with these substrates.

In addition, it would be interesting to generate a new plasmid with either a larger or smaller Mini-Tn4430 or a larger or smaller backbone plasmid than those of pGIMF001 and pGIAR045 or also with a different flanking sequence. With this new plasmid, a cleavage assay should be performed to compare the activity of TnpA on this plasmid to the activity on the two other plasmids. This experiment will allow us to determine whether the flanking sequence

or the size of the plasmid backbone or the size of Mini-Tn4430 has an influence on TnpA activity.

Finally, following the appearance of these additional bands, it would be important to determine whether they are really due to strand transfer. Furthermore, if these bands are indeed specific strand transfer products, the topology of these products should also be determined.

To do this, an AFM (atomic force microscopy) visualisation of the products could be carried out. This visualisation would be performed from the bands of the gels. If it turns out that these products are indeed strand transfer products, the topology of these products could also be revealed thanks to the visualisation.

Moreover, other experiments could also confirm or refute our hypothesis. For example, the products of the reaction could be put on a 2D gel to better separate our products according to their topology. This will give us additional information on the topology of these products.

V.Materials and methods

Table 1: ***Solutions used***

Solution	Composition
Buffer A (TpA)	50 mM Tris pH8, 1 M NaCl, 10 % glycerol, 20 mM imidazole
Buffer B (TpB)	Tris 50 mM pH8, NaCl 1 M, glycerol 10 %, imidazole 500 mM
Coomassie Blue	Ethanol 50 %, acetic acid 7 %, Coomassie blue 0.2 % Crack
Cleavage buffer	0,75 M NaCl, HEPES buffer pH 8 12,5%, $MnCl_2$ 25 mM, DTT 5 mM
Crack 5X	0.25 M Tris HCl pH 6.5, 0.1% SDS, 0.2 % bromophenol blue, 50 % glycerol, 15 % β- mercaptoethanol by volume
Hepes Buffer pH8	HEPES 500 mM, NaCl 2 M, L-arginine 1 M
LB liquid culture medium	Tryptone 10 g/l, yeast extract 5 g/l, NaCl 5 g/l
LB agar culture medium	Tryptone 10 g/l, yeast extract 5 g/l, NaCl 5 g/l, agar 20 g/l
Lysis buffer	Tris 50 mM pH 8, NaCl 1 M, glycerol 10 %
Orange G	Bromophénol 0,25 %, xylène cyanol 0,25 %, glycerol 30 %
Stop buffer	SDS 0,2 %, EDTA 20 mM, Proteinase K 1,2 mg/ml
SYBR-Gold buffer	10 µl for 100 ml of TAE buffer
TAE buffer	88 mM Tris base, 88 mM ortho boric acid, 2 mM EDTA, pH8
TE buffer	10 mM Tris-HCl pH8, 1 mM EDTA
TN buffer	Tris 10 mM pH8, NaCl 100 mM

V.1 Description of the substrates

V.1.1 pGIMF001

At the heart of this project, pGIMF001 is a fully designed artificial 3,9-Kb plasmid carrying a 0,68-Kb Mini-Tn4430 element delineated by properly oriented left and right 38-bp terminal inverted repeats of Tn4430 in their genuine context (TIRL and TIRR, respectively) (Fig. V.1.1). The recognition sites for the single-strand DNA (ssDNA) endonuclease Nt.BspQI has been introduced opposite to the TnpA cleavage sites on both ends. Since TnpA produces ssDNA nicks at the 3' ends of Tn4430, subsequent digestion of the reactions with Nt.BspQI generates double-strand DNA breaks (dsDBs) that can be readily identified by standard gel electrophoresis. Cleavage sites for two other ssDNA endonuclease (Nb.BtsI and Nb.BsrDI, respectively) have been positioned at the 3' ends of Mini-Tn4430 so as to mimic TnpA-mediated end cleavages.

An open circular form of this plasmid has been generated. For this purpose, we used either the Nt.BsmAI ssDNA endonuclease (Fig. V.1.1).

In the experiments (see III.4 and III.7), different restriction enzymes were used.

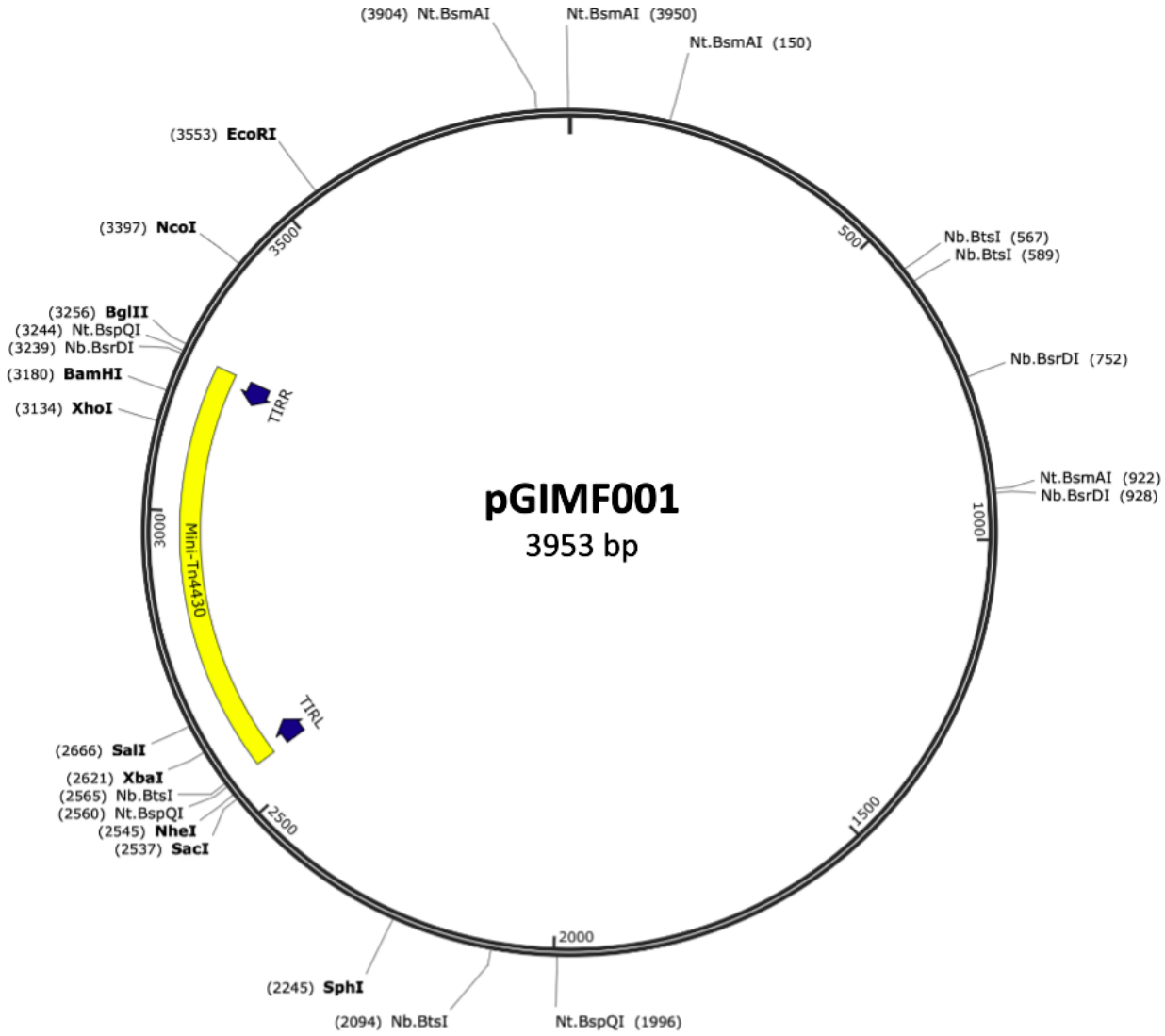


*Figure V.1.1: **Map of pGIMF001**. The Mini-Tn4430 is shown in yellow and the TIRR and TIRL are represented by the dark blue arrow. The different restriction sites of the single-stranded DNA endonucleases and the restriction enzymes are shown on the map.*

V.1.2 pGIAR045

pGIAR045 is also a fully engineered 4.629 kb artificial plasmid carrying a 1.151 kb Mini-Tn4430 element bounded by correctly oriented 38 bp terminal IR (Fig. V.1.2). The single-stranded DNA endonuclease (ssDNA) recognition sites Nt.BspQI were introduced opposite the TnpA cleavage sites at both ends. So, pGIAR045 has the same overall structure as pGIMF001 but the size and the sequence of Mini-Tn4430 are different.

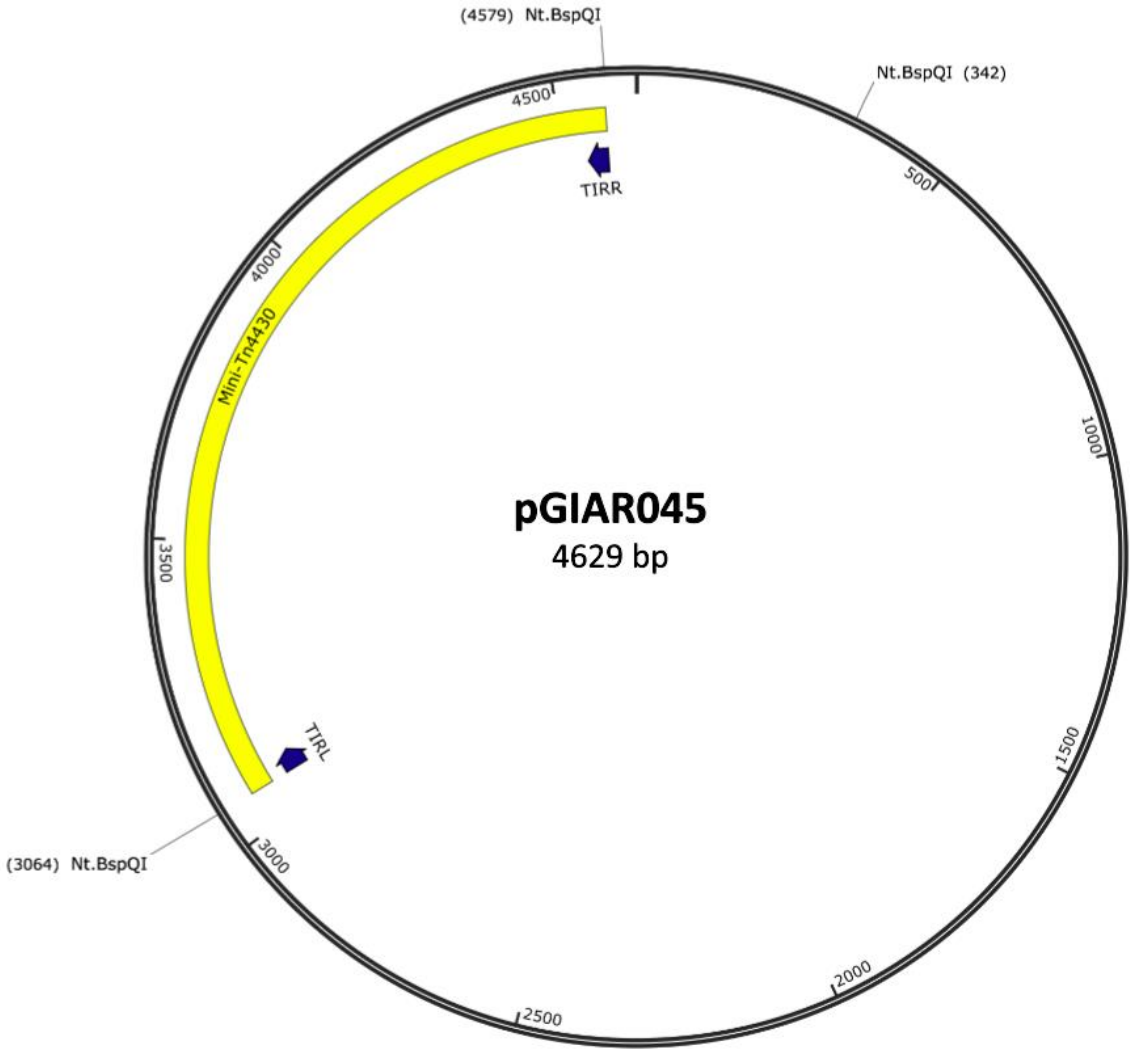


Figure V.1.2: ***Map of pGIAR045.*** *The Mini-Tn4430 is shown in yellow and the TIRR and TIRL are represented by the dark blue arrow. The restriction site of the single-stranded DNA endonuclease Nt.BspQI is indicated on the map.*

V.1.3 Plasmid purification

The plasmids are extracted from the *E. coli* Top10 strain in which our plasmid of interest is present. The plasmid contains an Ampicillin resistance gene (Amp^r). *E. coli* TOP10 strains containing the vector of interest are streaked onto LB-agar plates with Ampicillin. The plates are then incubated at 37°C overnight. An isolated colony is selected from the plate to inoculate a 15 ml LB liquid pre-culture at 37°C overnight.

After this, 1 ml of the LB liquid culture is placed in an eppendorf tube and centrifuged for 30 seconds at 10000 g. The supernatant is then removed. This step is repeated 3 times for each tube. The pellet is then re-suspended in 250 µl re-suspension buffer from the kit. 250 µl

of Lysis Buffer from the kit is added and each tube is mixed by 4-6 inversion. 350 µl of neutralisation buffer from the kit is then added and the solution is also mixed by 4-6 inversion. The sample is centrifuged for 10 min at 14000g and the supernatant is placed in one of the kit columns. After this, the column is centrifuged for 1 min at 6000 g and 650 µl of kit wash buffer is added. The whole is centrifuged again for 1 min at 12000 g and the flow-through is removed. This step is repeated twice. The column is then placed in a sterile eppendorf and 40 µl of distilled H_2O is added to the column. Finally, after a 1 min centrifugation at 12000 g, our purified plasmids of interest are found in the eppendorf tube.

V.2 Purification of transposase

V.2.1 Expression Vector

All TnpA proteins were produced from *E. coli* TOP10 F- strains. These strains contain an expression vector consisting of an L-arabinose inducible P-ARA promoter with the TnpA sequence inserted upstream in translational fusion with a MycHis6 tag. The expression vector also contains a Tetracycline resistance gene (Tc^r) (Nicolas et al. 2017).

*Table 2: **Vector used** (Nicolas et al. 2017)*

Name of the vector	**TnpA purified**
pGIAD003	WT
pGIML3X/M-H	3X
pGIML007	S911R
pGIAD002	D751N/S911R

V.2.2 Induction

First, *E. coli* TOP10 strains containing the vector of interest are stripped onto LB-agar plates Tetracycline. The plates are then incubated at 37°C overnight. For each strain used, an isolated colony is selected from the plate to inoculate a 30 ml LB liquid pre-culture containing Tetracycline at 37°C overnight. Afterwards, the pre-culture is used to inoculate a 750 ml LB liquid culture containing Tetracycline and Benzyl Alcohol 10 µM. Induction is done in the presence of benzyl alcohol as the transposase is quite unstable and not very soluble. After

inoculation, 1 ml of the liquid LB culture is harvested to determine its optical density (OD) which should be at least 0.04. The culture is incubated at 37°C with 150 rpm agitation. The OD is assessed regularly during this period. When the OD is between 0.5 and 0.7, the culture is placed on ice for 10 min to reduce the temperature of the culture to 20°C. 1 ml of the culture is also recovered and washed 2 times in 1 ml TN and stored at -20°C. This will serve as a pre-induction control. Subsequently, induction of the P-ARA promoter is initiated by adding 1.5 ml of L-arabinose (20%).The culture is then incubated for 2 h at 22.5°C with 100 rpm agitation. After 2 h of incubation, 0.75 ml of L-arabinose (10%) is added and the culture is incubated for a further 2 h under the same conditions. After the additional 2 h of induction, 1ml is collected and washed 2 times in 1 ml TN and stored at -20°C. This will serve as a post induction control. Next, the culture is centrifuged for 10 min at 4°C, 4000 rpm. Afterwards, the pellet is re-suspended in 100 ml lysis buffer and centrifuged again for 10 min at 4°C, 4000 rpm. Finally, the pellet is stored at -80°C in its centrifugation tube (Nicolas et al. 2017).

V.2.3 Lysis

The pellet is thawed on ice. It is then re-suspended in 36 ml of Buffer A to which cOmplete EDTA-free (protease inhibitor) is added. One pellet is added per 30 ml of Buffer A. Cell lysis is performed by centrifugation with glass beads. 1 ml of the solution is distributed in different 2 ml tubes containing 200 µl of glass beads and these tubes are centrifuged twice at 6.0 m/s for 25 sec. The different fractions of the sample are placed on ice between each step. After lysis, the samples are centrifuged for 15 min at 4°C 13000 rpm. The supernatant is collected in a tube and centrifuged again for 15 min at 4°C 13000 rpm. The supernatant is placed in a syringe and passed through a 0.45 µm mesh filter and placed at 4°C to purification (Nicolas et al. 2017).

V.2.4 Purification

Purification is based on the affinity of the N-terminal MycHis6 tag for a nickel column. The entire protocol is performed at 4°C using an Äkta Prime. The solutions used with the Äkta Prime are filtered and degassed. First, the machine is washed with water. Then the column is equilibrated with lysis buffer and a 0.1 M $NiSO_4$ solution. After this, the filtered lysate is loaded onto the column. The column is then washed with a mixture of buffer A and B with an

increasing proportion of buffer B. This generates a concentration gradient of imidazole which gradually increases in concentration from 20 mM to 500 mM. Nonspecific binding between contaminating proteins and the column is displaced by low concentrations of imidazol (20 mM) whereas elution of recombinant TnpA bound by its MycHis6 tag occurred at higher concentrations. The elution of proteins from the lysate is monitored spectrophotometrically allowing detection of the elution peak which corresponds to the elution of TnpA (Nicolas et al. 2017).

V.2.5 SDS-Page Gel

For reach purification, culture samples before and after 4 h of induction were compared to the elution fractions from the purification by SDS-PAGE electrophoresis. All samples were suspended in a mixture of "crack" buffer (Table 1) and H_2O and then heated for 10 min at 95°C. After this, the samples were loaded onto a pre-cast 4-10 % gel (Eurogentec). The gel was subjected to a current of 120 V for 1 h and 10 min. The gel was finally revealed by immersion in Instant Blue (Abcam) (Nicolas et al. 2017).

V.2.6 Protein reconcentration

Fractions containing the most TnpA and the least impurities are collected. They were then reconcentrated using an AMICON Ultra 4 – 10 kDa column. The AMICON column containing the fractions of interest is centrifuged at 4°C 4000g for 15 to 20 min. The column is then refilled with lysis buffer and centrifuged again at 4°C 4000 g for 15-20 min. The liquid remaining in the column is separated into different samples which are then stored at -80°C (Nicolas et al. 2017).

V.2.7 Bradford assay

The concentration of TnpA used was determined by Bradford assay. 40 µl of Bradford Dye and 150 µl of H_2O are added to 10 µl of sample and the concentration of this 10 µl of sample is assessed via its absorbance. From an absorbance calibration curve for known concentrations of BSA, the absorbance values of our samples are compared and the concentration of TnpA can be determined (Nicolas et al. 2017).

V.3 Cleavage assay

The cleavage assay is the central point of this master thesis. It consists of a biochemical assay to study the activities of TnpA on a DNA molecule carrying functional or mutants derivatives of Tn4430 (or "Mini-Tn4430").

V.3.1 Cleavage test

The cleavage test is performed in one eppendorf per condition. The final volume in the eppendorf is 30 µl. In this eppendorf 6 µl of cleavage buffer, 6 µl of BSA (10 mg/ml) and 2 µl of purified plasmid (250 ng/ml) as well as 2 µl distilled H_20 are placed. Next, the eppendorf is placed on ice and 8 µl of TnpA (300nM final concentration) and 6 µl of glycerol were added (22% final concentration) or not depending on the experiment. Finally, the resulting solution is mixed by pipetting and the eppendorf is placed for at 34 °C for different incubation times (from 0 to 240 min) depending on the experiment.

After incubation, 30 µl of stop buffer is added and the reaction is incubated again for 30 min at 37°C.

For the time course assay, all the reactions are performed in the same eppendorf (e.g. for 6 cleavage reactions the final volume was 180 µl). After each time point, a 30 µl sample is removed the reaction is stopped by adding 30 µl of stop buffer as described above.

V.3.2 Purification of the DNA products

Initially purification of the reaction products was carried out by phenol extraction but this was later replaced by purification with the NEB monarch kit.

V.3.2.1 Phenol extraction

The 60 µl of cleavage assay is collected and mixed with 140 µl of distilled H_2O, 40 µl of 5M $NaClO_4$ and with 100 µl of phenol at ph 7.5-8. The solution is then vortexed 30 secs and 100 µl of chloroform/isoamyl alcohol (24:1) is added. The solution is vortexed again for 30 secs and then centrifuged for 5 min at 15000 rpm. After this, the aqueous phase is collected and 100 µl of chloroform and 150 µl of TE buffer are added. The solution is vortexed again and then centrifuged for 5 min at 15000 rpm. The aqueous phase is again collected in the same

tube and 720 µl of 100 % ethanol and 40 µl of 5 M $NaClO_4$ are added. The sample is left for 1 h at -20°C and then centrifuged for 30 min at 4°C at 15000 rpm. The supernatant is removed and the tube is placed at 37°C for 45 min to dry. Finally, 20 µl of distilled H_2O is added to the tube and the solution is mixed by pipetting.

V.3.2.2 Purification with NEB monarch kit

The 60 µl of cleavage assay is collected and are mixed with 120 µl of binding buffer from the kit. Then the 180 µl obtained are placed in a column and then centrifuged for 1 min at 16000 g. The column is then washed by adding 200 µl of wash buffer of the kit and centrifuging for 1 min at 16000 g. This step is repeated twice. Finally, the DNA will be eluted by adding 10 µl of distilled H_20 to the column and then centrifuging for 1 min at 16000g. The 10 µl are collected in a new eppendorf.

V.3.3 Digestion with restriction enzymes and agarose gel

After DNA purification, the 10 µl containing DNA is collected and 8 µl of H_2O, 2 µl of restriction buffer and 1 µl of restriction enzyme (often Nt.BspQI) is added. The eppendorf is then incubated for 1 h at 50°C.

After incubation, 3µl of orange G is added to the sample and it is placed in a well of an agarose gel. First, the agarose gels contained 0.8% agarose and were subjected to a current of 140 V for 1 h and 15 min. In a second step, the gels contained 1.2% agarose and were subjected to a current of 140 V for 2 h and 30 min. The gel is then incubated in SYBR-Gold buffer for 30 min in the dark.

Finally, the gel is revealed by fluorescence using a fluorescence scanner (American Typhoon).

VI. Bibliography

Arthur, A., Nimmo, E., Hettle, S., & Sherratt, D. (1984). Transposition and transposition immunity of transposon Tn3 derivatives having different ends. *The EMBO journal*, 3(8), 1723-1729.

Beauregard, A., Curcio, M. J., & Belfort, M. (2008). The take and give between retrotransposable elements and their hosts. *Annual review of genetics*, 42, 587-617.

Blundell-Hunter, G., Tellier, M., & Chalmers, R. (2018). Transposase subunit architecture and its relationship to genome size and the rate of transposition in prokaryotes and eukaryotes. *Nucleic acids research*, 46(18), 9637-9646.

Cerqueira, G. C., Earl, A. M., Ernst, C. M., Grad, Y. H., Dekker, J. P., Feldgarden, M., ... & Hanage, W. P. (2017). Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks. *Proceedings of the National Academy of Sciences*, 114(5), 1135-1140.

Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics*, 18(2), 71-86.

Craig, N. L. (2015). A moveable feast: an introduction to mobile DNA. *Mobile DNA III*, 1-39.

Curcio, M. J., & Derbyshire, K. M. (2003). The outs and ins of transposition: from mu to kangaroo. *Nature Reviews Molecular Cell Biology*, 4(11), 865-877.

Eickbush, T. H., & Jamburuthugoda, V. K. (2008). The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus research*, 134(1-2), 221-234.

Fedoroff, N., Wessler, S., & Shure, M. (1983). Isolation of the transposable maize controlling elements Ac and Ds. *Cell*, 35(1), 235-242.

Finnegan, D. J. (1989). Eukaryotic transposable elements and genome evolution. *Trends in genetics*, 5, 103-107.

Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology*, 3(9), 722-732.

Goodwin, T. J., & Poulter, R. T. (2001). The DIRS1 group of retrotransposons. *Molecular biology and evolution*, 18(11), 2067-2082.

Guynet, C., Nicolas, E., Ton-Hoang, B., Bouet, J. Y., & Hallet, B. (2020). First biochemical steps on bacterial transposition pathways. *In Horizontal Gene Transfer*, 157-177.

Grindley, N. D., Whiteson, K. L., & Rice, P. A. (2006). Mechanisms of site-specific recombination. *Annual review of biochemistry*, 75, 567-605.

Hallet, B., Rezsöhazy, R., & Delcour, J. (1991). IS231A from Bacillus thuringiensis is functional in Escherichia coli: transposition and insertion specificity. *Journal of bacteriology*, 173(14), 4526-4529.

Hallet, B. (2021). Replicative transposition of Tn3-family transposons on target: from atomic structures to molecular mechanism. F.R.S.-FNRS CREDITS AND PROJECTS CALL. *Unpublished document.*

Harshey, R. M. (2012). The Mu story: how a maverick phage moved the field forward. *Mobile DNA*, 3(1), 1-10.

Heffron, F., McCarthy, B. J., Ohtsubo, H., & Ohtsubo, E. (1979). DNA sequence analysis of the transposon Tn3: three genes and three sites involved in transposition of Tn3. *Cell*, 18(4), 1153-1163.

Hickman, A. B., Chandler, M., & Dyda, F. (2010). Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Critical reviews in biochemistry and molecular biology*, 45(1), 50-69.

Hickman, A. B., & Dyda, F. (2015). Mechanisms of DNA transposition. Microbiology spectrum, 3(2), 3-2.

Hickman, A. B. & Dyda, F. (2016). DNA Transposition at Work. *Chem Rev*. 116:12758-12784.

Kazazian, H. H. (2004). Mobile elements: drivers of genome evolution. *Science*, 303(5664), 1626-1632.

Krupovic, M., Makarova, K. S., Forterre, P., Prangishvili, D., & Koonin, E. V. (2014). Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC biology*, 12(1), 1-12.

Lambin, M., Nicolas, E., Oger, C. A., Nguyen, N., Prozzi, D., & Hallet, B. (2012). Separate structural and functional domains of Tn4430 transposase contribute to target immunity. *Molecular microbiology*, 83(4), 805-820.

Mahillon, J., & Chandler, M. (1998). Insertion sequences. Microbiology and molecular biology reviews, 62(3), 725-774.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences*, 36(6), 344-355.

Mizuuchi, K. (1992). Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annual review of biochemistry*, 61(1), 1011-1051.

Moelbert, S., Normand, B., & De Los Rios, P. (2004). Kosmotropes and chaotropes: modelling preferential exclusion, binding and aggregate stability. *Biophysical chemistry*, 112(1), 45-57.

Nicolas, E., Lambin, M., Dandoy, D., Galloy, C., Nguyen, N., Oger, C. A., & Hallet, B. (2015). The Tn 3-family of replicative transposons. *Microbiology spectrum*, 3(4), 3-4.

Nicolas, E., Lambin, M., & Hallet, B. (2010). Target immunity of the Tn 3-family transposon Tn 4430 requires specific interactions between the transposase and the terminal inverted repeats of the transposon. *Journal of bacteriology*, 192(16), 4233-4238.

Nicolas, E., Oger, C. A., Nguyen, N., Lambin, M., Draime, A., Leterme, S. C., ... & Hallet, B. F. (2017). Unlocking Tn3-family transposase activity in vitro unveils an asymetric pathway for transposome assembly. *Proceedings of the National Academy of Sciences*, 114(5), E669-E678.

Nicolas, E., Oger, C. A., Nguyen, N., Lambin, M., Chandler, B., & Hallet, B. A replication fork hijacking model for Tn3- family replicative transposition. *In prep*.

O'Neill, J. (2016). Tackling drug-resistant infections globally: final report and recommendations.

Partridge, S. R., Kwong, S. M., Firth, N., & Jensen, S. O. (2018). Mobile genetic elements associated with antimicrobial resistance. *Clinical microbiology reviews*, 31(4).

Peters, J. E. (2015). Tn7. Mobile DNA III, 647-667.

Piégu, B., Bire, S., Arensburger, P., & Bigot, Y. (2015). A survey of transposable element classification systems–a call for a fundamental update to meet the challenge of their diversity and complexity. *Molecular phylogenetics and evolution*, 86, 90-109.

Ronning, D. R., Guynet, C., Ton-Hoang, B., Perez, Z. N., Ghirlando, R., Chandler, M., & Dyda, F. (2005). Active site sharing and subterminal hairpin recognition in a new class of DNA transposases. *Molecular cell*, 20(1), 143-154.

Sakai, J., & Kleckner, N. (1997). The Tn10 synaptic complex can capture a target DNA only after transposon excision. *Cell*, 89(2), 205-214.

SanMiguel, P., Tikhonov, A., Jin, Y. K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., ... & Bennetzen, J. L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science*, 274(5288), 765-768.

Shkumatov, A. V., Aryanpour, A., Hallet, B., & Efremov, R. G. Structures of Tn3 family transposase reveal novel mechanism of transposition regulation. *In prep*.

Smit, A. F. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Current opinion in genetics & development*, 9(6), 657-663.

Snesrud, E., Maybank, R., Kwak, Y. I., Jones, A. R., Hinkle, M. K., & McGann, P. (2018). Chromosomally encoded mcr-5 in colistin-nonsusceptible Pseudomonas aeruginosa. *Antimicrobial agents and chemotherapy*, 62(8).

Thomson, P. (2016). Political Declaration of the High-Level Meeting of the General Assembly on Antimicrobial Resistance : draft resolution / submitted by the President of the General Assembly. *United Nation Digital Library.*

Touchon, M., & Rocha, E. P. (2007). Causes of insertion sequences abundance in prokaryotic genomes. *Molecular biology and evolution*, 24(4), 969-981.

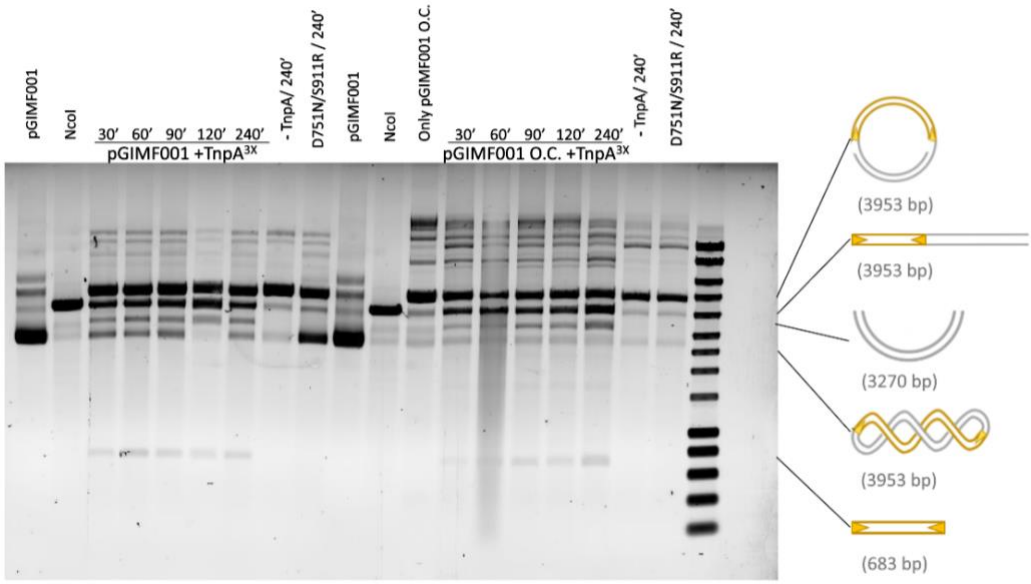
Toussaint, A., & Merlin, C. (2002). Mobile elements as a combination of functional modules. *Plasmid*, 47(1), 26-35.

Vanhooff, V., Galloy, C., Agaisse, H., Lereclus, D., Révet, B., & Hallet, B. (2006). Self‐control in DNA site‐specific recombination mediated by the tyrosine recombinase TnpI. *Molecular microbiology*, 60(3), 617-629.

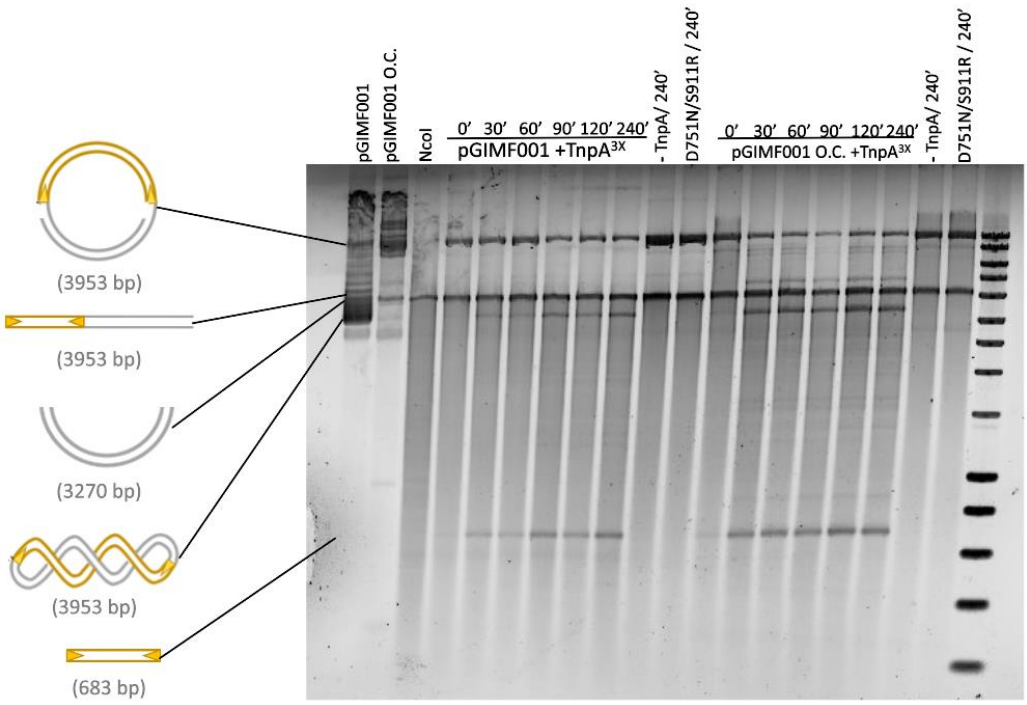
Yoder, J. A., Walsh, C. P., & Bestor, T. H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in genetics*, 13(8), 335-340.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., ... & Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, 8(12), 973-982.

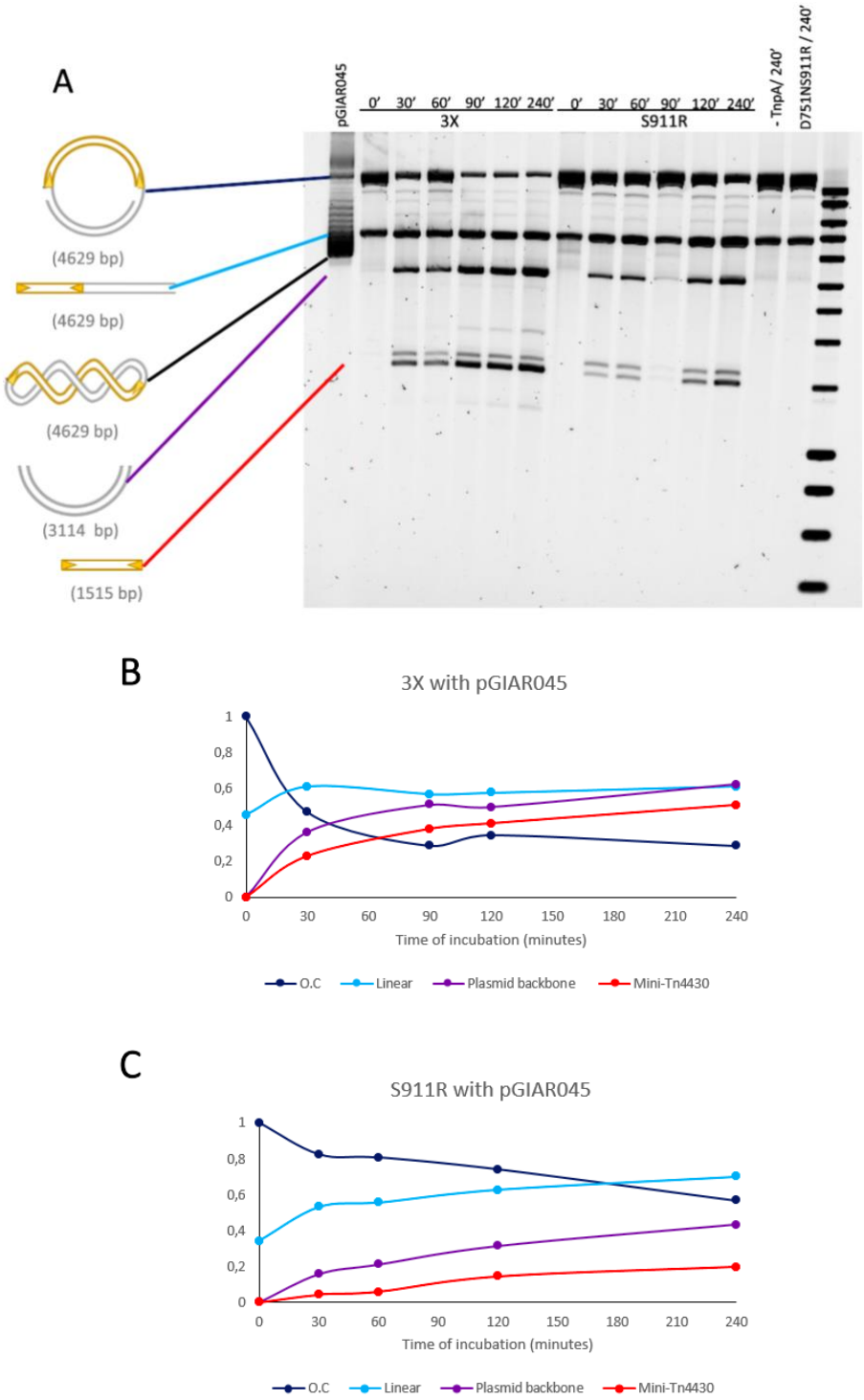
VII. Appendices



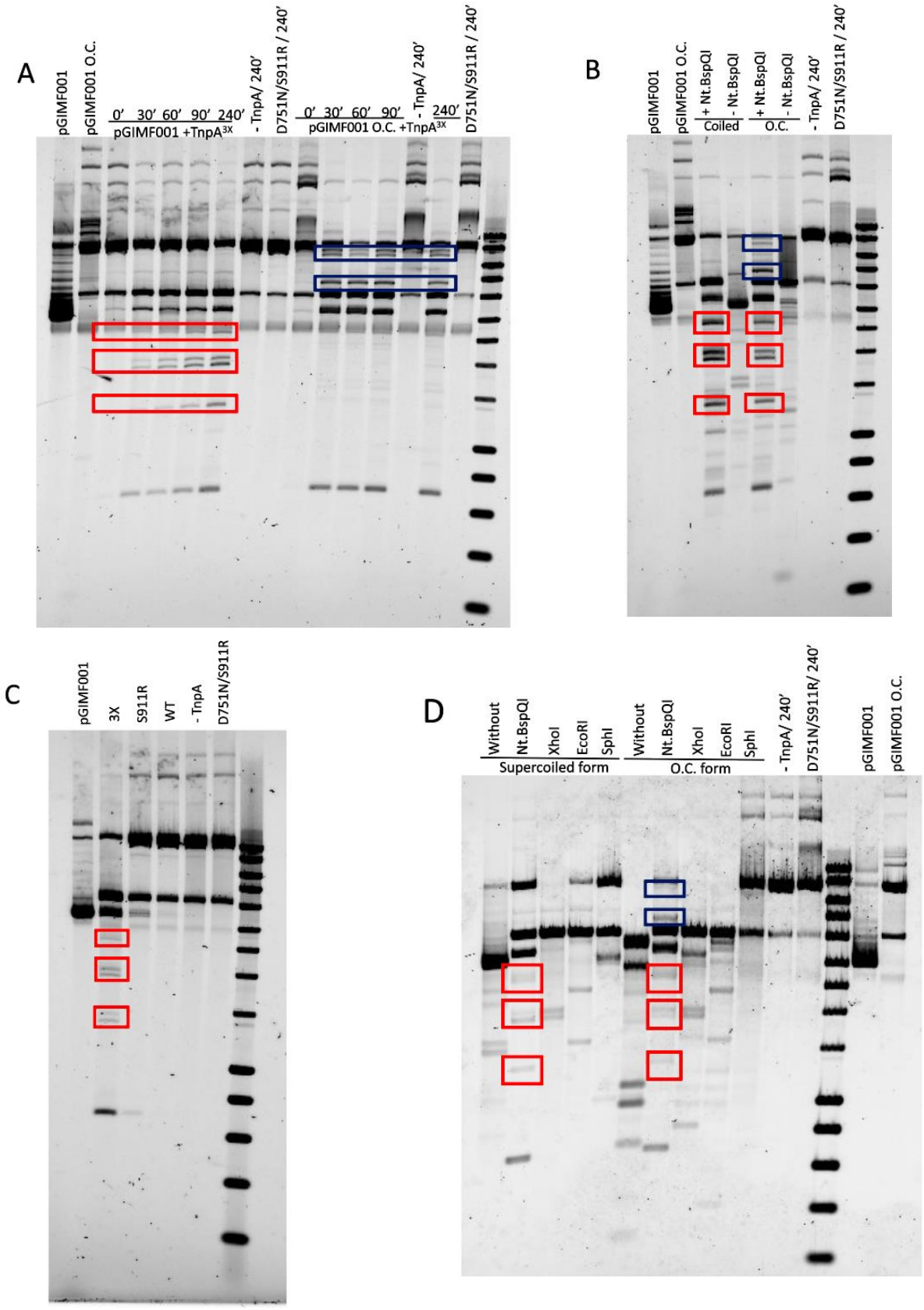
Appendix 1: ***Kinetic cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM).*** *The incubation time is noted on the legend. 0.8% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. NcoI is an enzyme that cuts pGIMF001 at a single site, generating a linear form of the plasmid. The eighth and penultimate lanes are controls without TnpA. The ninth and last lane are controls with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*



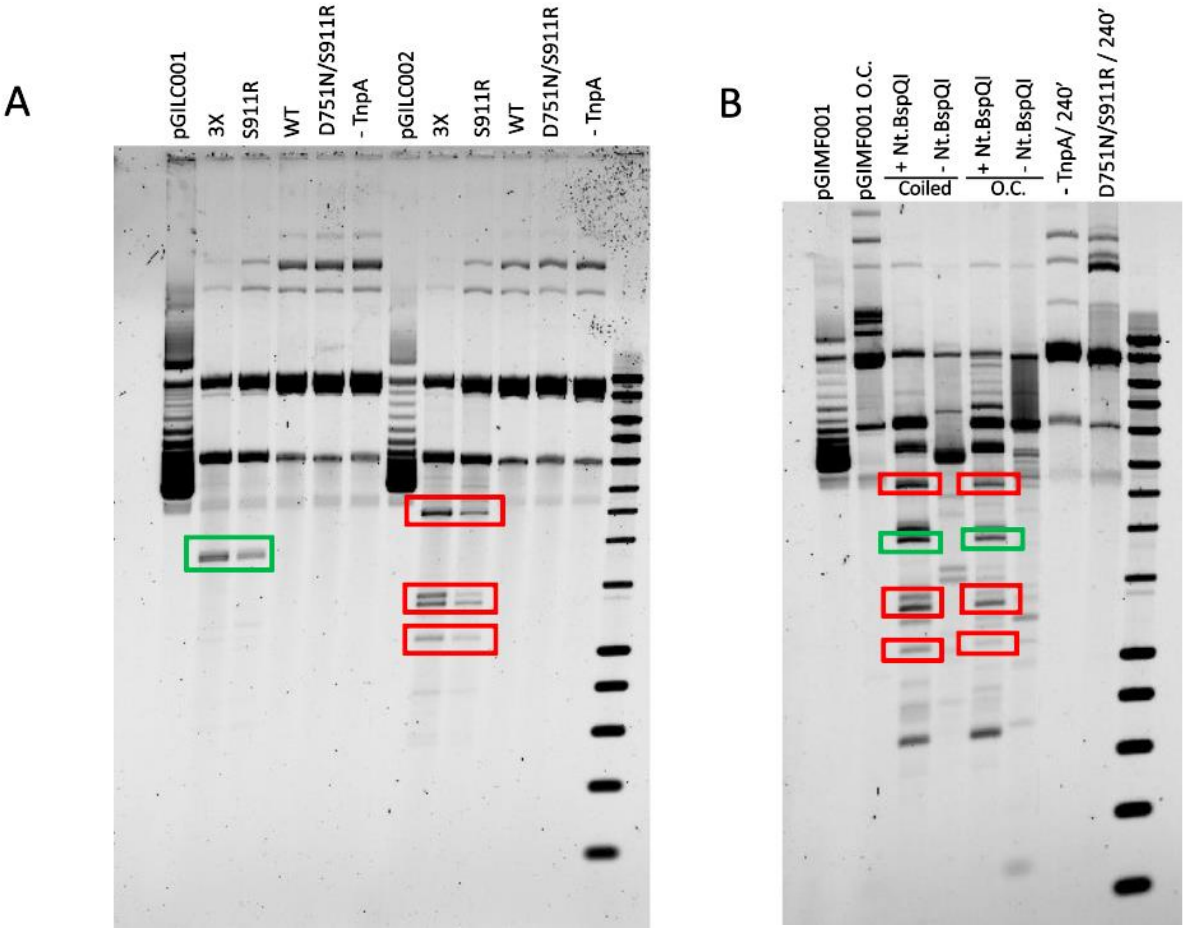
Appendix 2: ***Kinetic cleavage test between pGIMF001 (250 ng/µl) or its open circular form (O.C.) (250 ng/µl) and TnpA^{3X} (300nM).*** *The incubation time is noted on the legend. 0.8% agarose gel. pGIMF001 and pGIMF001 O.C. are only the plasmid without treatment. NcoI is an enzyme that cuts pGIMF001 at a single site, generating a linear form of the plasmid. The tenth and penultimate lanes are controls without TnpA. The eleventh and last lane are controls with the catalytic mutant TnpA$^{D751N/S911R}$ (300 nM).*



Appendix 3: ***(A) Kinetic cleavage test between pGIAR045 (250 ng/µl) and TnpA^{3X} (300nM) and TnpA^{S911R} (300nM).*** *The incubation time is noted on the legend. 1.2% agarose gel. pGIAR045 is only the plasmid without treatment. The penultimate lane is controls without TnpA. The last lane is controls with the catalytic mutant TnpA$^{D751/S911RN}$ (300 nM).* ***(B and C)*** *Quantification of TnpA^{3X}and of TnpA^{S911R} activity from the kinetic cleavage. Graph of the evolution of the quantity of the different substrates as a function of the incubation time. The graphs show the relative amount of the different cleavage products with respect to the initial quantity of the form OC at time 0 reported on 1. The transposase used for graph B is TnpA^{3X} and the transposase used for graph C is TnpA^{S911R}.The dark blue curve represents the relative amount of the circular open form of the plasmid, the light blue curve represents the linear plasmid, the purple curve represents the backbone plasmid and the red curve represents the Mini-Tn4430. Quantification was performed with ImageJ software.*



Appendix 4: ***Comparison of the additional bands between different figures****. The red rectangles indicate the identical additional bands for supercoiled pGIMF001 and the O.C. form. The blue rectangles indicate the additional bands for the O.C. form of pGIMF001. (A) Figure III.3.2A (B) Figure III.7.1 (C) Figure III.1.2 (D) Figure III.7.2.*



Appendix 5: ***Comparison of the additional bands between different figures****. The green rectangles indicate the identical additional bands for the pGILC001 and the pGIMF001 supercoiled and its O.C. form with Nt.BspQI. The red rectangles indicate the additional bands for pGILC002 and the pGIMF001 supercoiled and its O.C. form with Nt.BspQI. (A) Figure III.4.3 (B) Figure III.7.1.*

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des sciences
Place des sciences, 2 bte L6.06.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/sc**