UCLouvain

FACULTE DES SCIENCES

Faculté des sciences

In vitro characterization of the Tn4430 transposase activity on DNA molecules containing Mini–Tn4430 element

Author: **Veronica Alejandra GONZALEZ GONZALEZ**
Supervisors: **Pr. Bernard HALLET, Nicolas ARYANPOUR, Dr. Maricruz FERNANDEZ**
Readers: **Dr. Corentin CLAEYS BOUUAERT, Pr. David ALSTEENS**

Master [120] en Biochimie et Biologie Moléculaire et Cellulaire, finalité spécialisée : biotechnologie

Academic year 2022–2023

Acknowledgement

I would like first to thank my supervisor Bernard Hallet for accepting me into his group and trusting me. I want to thank him for the time he took to explain, help and guide me through my project.

As for my PhD students, Maricruz and Nicolas, I would like to thank them for all the time, guidance, and advice they gave me during this project. Thanks to them, I was able to feel comfortable, and learn a lot of different things about transposons and laboratory techniques.

I would like to thank all the BGM members, especially Brigitte and Damien whose advice and guidance during my experiments were helpful. I am grateful for all the work they do in the lab and I would like to thank them for helping me feel comfortable and welcome in this lab.

I would like to especially thank my BGM coworker (well, if we could say that) and friend, Deven, whose presence throughout the project, even in summer, was very helpful. Thanks for being there on joyful and not-so-joyful days, for sharing with me all the steps, and for taking coffee breaks. I would also like to thank my other BGM coworkers, Flore, Louise, Anaëlle, Mathilde, and Gergana, and a friend from another lab, Victoria. Thank you for sharing almost every day with me, your presence at the lab and during the breaks was very pleasant.

Last but not least, I would like to give special thanks to my family for always being there for me and supporting me in every step and decision I take. Thanks to them I am where I am and I feel grateful and very happy. My biggest thanks go to my brother, Marco, whose presence during this Master's and especially during my Master's thesis was incomparable. Even if you do not know anything about transposons, your kindness, attentiveness, and advice helped me understand many things and made me surpass my capabilities in many ways. I still would not know what to do without you.

I am very grateful and happy to have been able to contribute as much as I could to the study of the Tn3-family transposons.

Abstract

Transposons are defined DNA sequences capable of moving from one location to another within and between genomes. This process is known as transposition and it happens in a nucleoproteic complex, the transpososome. This complex contains the protein allowing the transposition: the transposase, the target DNA molecule, and the transposon ends.

The Tn4430 transposon is a member of a widespread family of transposons, the Tn3 family, and is used as a representative model for the study of this family. The Tn3 family significantly impacts human health because it is implicated in the dissemination and persistence of antibiotic resistance among pathogens. Despite this biological significance, there are still many unknown molecular aspects of the transposition mechanism of this family.

In previous works, the initial steps of the Tn4430 transposition have been studied and they have shown that this pathway leads to the assembly and activation of the transposition complex. Thanks to cryo-EM, different conformations of the transposase (TnpA) have been obtained and have allowed us to characterize the transposition complex at different steps of the reaction.

Moreover, *in vivo* and *in vitro* data have converged into a new transposition mechanism called "replication hijacking". During this mechanism, Tn3-family transposons can integrate into intermediates of replication allowing them to recruit the host's replication machinery and then create a new copy of themselves. During transposition, TnpA mediates another mechanism called "immunity" where it prevents the insertion of a new transposon copy into a target DNA molecule that already presents a copy of the same transposon. Its mechanism is still not completely understood, but we have been able to isolate TnpA mutants that have lost this "immunity" mechanism which makes them more prone than the wild-type transposase to catalyze the DNA cleavage and strand transfer reactions.

In this work, we use the hyperactive mutants of the TnpA ($TnpA^{3X}$ and $TnpA^{S911R}$) to continue the *in vitro* study of their activity. To this end, a biochemical assay was used to determine how the DNA and transposon topology on the donor molecule impacts the TnpA activity. A supercoiled, open circular, or linear version of a plasmid containing a Mini-Tn4430 was used as a substrate. Furthermore, other important factors for transposition were studied. The temperature to further optimize the Tn4430 transposition reaction, and the donor and protein concentrations to test the hypothesis of assembly site occlusion (ASO) control of Tn4430 transposition.

This thesis shows that the DNA topology of the donor molecule (supercoiled, open circular, or linear) does not seem to impact the TnpA activity. Additionally, the transposon topology showed that the transposase needs at least one transposon end to be activated and the presence of both transposon end on the same molecule does not appear to have an impact. Finally, we started an analysis of unknown cleavage products that appeared in different reactions Although this analysis will have to be expanded and completed in further works, we proposed some hypotheses concerning the potential topologies and origins of these products which could be the result of strand transfer.

Resumé

Les transposons sont des séquences d'ADN particulières capables de se déplacer d'un endroit à un autre à l'intérieur et entre les génomes. Ce processus est connu sous le nom de transposition et a lieu dans un complexe nucléoprotéique, le transpososome. Ce complexe comprend la protéine permettant la transposition : la transposase, la molécule d'ADN cible et les extrémités du transposon.

Le transposon Tn4430 fait partie d'une famille répandue de transposons, la famille Tn3, et est utilisé comme modèle représentatif pour l'étude de cette famille. La famille Tn3 a un impact significatif sur la santé humaine car elle est impliquée dans la dissémination et la persistance de la résistance aux antibiotiques chez les pathogènes. Malgré cela, il existe de nombreux aspects moléculaires du mécanisme de transposition de cette famille qui restent encore inconnus.

Dans des travaux précédents, les étapes initiales de la transposition de Tn4430 ont été étudiées et elles ont montré que cette voie conduit à l'assemblage et à l'activation du complexe de transposition. Grâce à la cryo-EM, différentes conformations de la transposase (TnpA) ont été obtenues et nous ont permis de caractériser le complexe de transposition à différentes étapes de la réaction.

De plus, les données *in vivo* and *in vitro* ont convergé vers un nouveau mécanisme de transposition appelé "replication hijacking". Au cours de ce mécanisme, les transposons de la famille Tn3 peuvent s'intégrer dans des intermédiaires de réplication leur permettant de recruter la machinerie de réplication de l'hôte puis de créer une nouvelle copie d'eux-mêmes. Lors de la transposition, TnpA assure un autre mécanisme appelé "immunité" où elle empêche l'insertion d'une nouvelle copie de transposon dans une molécule d'ADN cible qui présente déjà une copie du même transposon. Il n'y a pas d'analyses complètes sur ce mécanisme mais nous avons pu isoler des mutants TnpA qui n'ont pas ce mécanisme "d'immunité", ce qui les rende plus enclins que la transposase de type sauvage à catalyser les réactions de clivage de l'ADN et de transfert de brin.

Dans ce travail, nous utilisons les mutants hyperactifs de la TnpA ($TnpA^{3X}$ and $TnpA^{S911R}$) pour poursuivre l'étude *in vitro* de leur activité. À cette fin, un test biochimique a été utilisé pour déterminer comment la topologie de l'ADN donneur et celle du transposon incluencent l'activité de la transposase grâce à un substrat plasmide sous forme superenroulée, ouverte circulaire ou linéarisée, contenant un Mini-Tn4430. En outre, d'autres facteurs importants pour la transposition, par ex. la température et la concentration de la molécule donneuse et de la protéine, ont été testés pour optimiser davantage la réaction de clivage et transfert de brin *in vitro* Tn4430.

Ce mémoire indique que la topologie de l'ADN donneur (superenroulé, ouvert circulaire ou linéaire) ne semble pas influencer l'activité de la TnpA. La topologie du transposon aurait également peu d'influence car la présence d'une seule extrémité de Tn4430 est suffisante pour atteindre une activité similaire à celle observée en présence des deux extrémités sur une même molécule. Enfin, nous avons démarré une analyse de certains produits de clivage inattendus. Bien que cette analyse doive être poursuivie, nous proposons des hypothèses sur leur topologies et origines potentielles qui pourraient être le résultat de transfert de brin.

Table of Contents

1 Introduction **1**
- 1.1 Transposable elements: engine of evolution . . . 1
- 1.2 Classification of transposons . . . 2
  - 1.2.1 Classification based on the transposition intermediate . . . 2
  - 1.2.2 Classification based on the transposase and the transposition mechanism 3
- 1.3 DDE/D transposases . . . 3
  - 1.3.1 Transposition mechanisms in DDE/D transposases . . . 4
  - 1.3.2 Structure of DDE/D transposases . . . 5
    - 1.3.2.1 Catalytic domain: RNase H-like fold . . . 6
    - 1.3.2.2 DNA-binding domain . . . 6
- 1.4 Regulation of transposition . . . 7
  - 1.4.1 Transpososome assembly . . . 7
  - 1.4.2 Autoregulation: OPI and ASO . . . 8
  - 1.4.3 Target molecule . . . 9
    - 1.4.3.1 Target recruitment . . . 9
    - 1.4.3.2 Target immunity . . . 9
  - 1.4.4 DNA topology . . . 10
    - 1.4.4.1 Bacteriophage Mu . . . 11
    - 1.4.4.2 Tn10 . . . 11
    - 1.4.4.3 Hsmar1 . . . 12
- 1.5 Tn3 family . . . 13
- 1.6 Tn4430: a representative model of the Tn3 family . . . 13
  - 1.6.1 Replication hijacking mechanism . . . 14
  - 1.6.2 Assembly and activation of Tn4430 transposition complex . . . 17
    - 1.6.2.1 TnpA binding to the transposon end and the target . . . 17
    - 1.6.2.2 Cryo-EM structures of the transposase and the DNA-bound complexes . . . 18
    - 1.6.2.3 Models for transposition complex assembly . . . 18

2 Objectives **20**

3 Results **22**
- 3.1 Optimization of the conditions for plasmid cleavage by the transposase . . . 22
  - 3.1.1 Optimal temperature . . . 23
  - 3.1.2 Is TnpA reaction subjected to assembly site occlusion (ASO)? . . . 24

3.2 Impact of the topology of the transposon . . . 27
3.3 Impact of the topology of donor DNA molecules . . . 29
3.3.1 Circular substrates . . . 29
3.3.2 Linear substrate . . . 32
3.4 Strand transfer products . . . 34

4 Discussion **39**

5 Conclusion and Perspectives **43**

6 Materials and Methods **44**
6.1 Substrates . . . 45
6.1.1 pGIMF001 . . . 45
6.1.2 Plasmid purification . . . 45
6.2 Transposase purification . . . 46
6.2.1 Expression vectors . . . 46
6.2.2 Induction . . . 46
6.2.3 Lysis . . . 47
6.2.4 Purification . . . 48
6.2.5 SDS-Page gel . . . 48
6.2.6 Protein reconcentration . . . 48
6.2.7 Bradford assay . . . 48
6.3 Cleavage assay . . . 49
6.3.1 Cleavage test . . . 49
6.3.2 DNA product purification . . . 49
6.3.3 Digestion with restriction enzymes and agarose gel . . . 50
6.3.4 Quantification of the cleavage reactions . . . 50

A Appendices **51**
A.1 Supplementary figures . . . 51
A.2 Creation of pGIVGG001 vector . . . 53

1 Introduction

1.1 Transposable elements: engine of evolution

During the 1940's and 1950's, an American scientist, Barbara McClintock, was studying the irregular coloring appearing on maize grains of *Zea mays*. This phenomenon was due to unstable mutations that were still unknown. McClintock observed some severe structural modifications happening on chromosome 9. She discovered a dissociation (Ds) locus, corresponding to a chromosome break, that could change its position within the chromosome [1, 2]. Besides this, she also discovered that its activity is dependent on another locus, the activator (Ac), which can also move from one position to another in the genome [1]. Even though at the beginning her work was not entirely recognized, after the discovery of other transposable loci in other species, she obtained the Nobel Prize in Physiology or Medicine in 1983 for the discovery of the first two transposable elements: Ac and Ds [2, 3, 4].

Nowadays, transposable elements (TE) or transposons are defined as discrete DNA fragments that can move from one position to another within and between genomes. Transposons are part of the mobile genetic elements (MGE), like the plasmids and bacteriophages, and share a modular organization based on genetic functions [5, 6, 7, 8]. The minimum requirement for an autonomous TE is a mobility gene coding for a recombination enzyme genetically referred to as "transposase", contained within delimiting sequences. During transposition, the transposase generally recognizes the TEs ends and mediates its displacement. In non-autonomous TEs, the transposase gene is truncated or absent but their mobility can be achieved if another copy of a related autonomous transposon is present [3, 9, 10].

The transposase-encoding genes are recognized as the most ubiquitous and most abundant genes in nature [11, 12]. For instance, in eukaryotic organisms, TEs correspond to up to 80% of the genome in plants and 70% in vertebrates. Because of this abundance, they play an essential role in chromosomes' functional and structural organization [6, 13]. In addition to that, the movement of transposons can have an impact on the expression of the host genes. If they insert themselves into or near a gene, it can result in its inactivation or a change in its expression level [5]. They are thus an important driving force of evolution [14, 15].

In prokaryotic organisms, transposons also have an important impact. They allow bacterial

communities to face environmental changes by exchanging specific genes of interest. This is done through a process known as horizontal gene transfer (HGT) [13, 16]. One clear example of this phenomenon is the implication of these MGEs in the dissemination of antibiotic resistance genes. They have helped bacteria to capture and accumulate these genes and thus allowed the emergence of multiresistant strains. Antibiotic resistance is a major healthcare-associated issue around the world because any infection of a multiresistant pathogen increases the risk of complications and mortality [7].

1.2 Classification of transposons

Transposons constitute a significant portion of almost all genomes that have been sequenced and are very diverse [17]. Since their discovery, many different classification methods have been used to classify them, among which the most informative are based on the way they move within and between genomes [18]. The two main criteria are the nature of the transposition intermediate, which can be either RNA or DNA, and the mechanism of action of the transposase protein [17, 18].

1.2.1 Classification based on the transposition intermediate

In 1989, David Finnegan presented the first comprehensive transposable element classification system. He differentiated two different classes based on their transposition intermediate. The class I elements transpose via an RNA molecule and the class II elements directly DNA from one locus to another [19, 20].

Class I elements or retrotransposons comprise around 40% of the human genome [21], and 75% of the maize genome [22]. They move by reverse transcription, meaning that during transposition, the transposon is first transcribed into RNA and then a reverse transcriptase generates the complementary DNA (cDNA) molecule. Finally, this cDNA is the molecule that will be integrated at a new position within the genome [21, 22]. There are two types of retrotransposons based on the DNA sequence topology [23]. The first ones are flanked by long terminal repeats (LTR), they are thus called LTR retrotransposons and their cDNA is produced in viral-like particles before integration. The other ones are non-LTR retrotransposons and their cDNA molecule is synthesized at the insertion site concomitant to its integration [17, 21, 24, 25].

Finally, class II elements have a DNA intermediate and are known as DNA transposons. The transposition mechanism of these TEs is done by integrating a DNA molecule into a new site in the genome, without synthesizing an RNA intermediate [26]. These transposons are composed of a transposase-encoding gene and delimiting sequences. These sequences are often constituted by a pair of terminal inverted repeats (TIR) [17, 26, 27]. In this class of TEs, we can find two different types of transposition depending on whether the transposon is duplicated or not. In the replicative transposition, the transposon can be duplicated either

before or after insertion into the genome by the host replication machinery [28]. On the other hand, in the non-replicative mechanism, the transposon is cut out from its position and then inserted elsewhere, without necessarily being replicated in the process [29].

In addition to this general classification system based on the transposition intermediates, there is another one based on the mechanism of transposition and the transposase proteins, which is more informative [18, 30].

1.2.2 Classification based on the transposase and the transposition mechanism

The transposition mechanism of class II elements requires the breakage and formation of phosphodiester bonds to allow the movement and insertion of DNA fragments within genomes. In order to catalyze these reactions, an enzyme – the transposase – is required [31]. In 2003, M. Joan Curcio and Keith M. Derbyshire proposed a classification system for transposons based on the transposases and their mechanism of action [18]. They described five families that differ in the amino acids involved in the catalytic reaction during transposition, and the formation of a covalent transposase-DNA intermediate or not [18, 30].

The first protein family is the HUH-motif transposase whose catalytic domain comprises two histidine residues separated by a hydrophobic one. These transposases need one or two additional tyrosine residues at the active site to complete the strand cleavage [30, 32]. Another family is the tyrosine transposases whose active site comprises a tyrosine residue acting as a nucleophile in the DNA cleavage reaction [30, 33]. Then, there are serine transposases that use a serine residue as their active site nucleophile [30, 33]. A fourth family is formed by casposases in which case the catalytic site uses a triad of two aspartate residues separated by a histidine residue (EHE) [34]. Finally, the last and largest family is the DDE/D transposases, comprising a broad diversity of enzymes whose catalytic domain has in common to contain a triad of two aspartate residues followed by either a glutamate or an aspartate as the last residue [30, 35]. This last family will be further discussed in the next section.

1.3 DDE/D transposases

Transposases of the DDE/D family present a common structural motif that allows the three essential catalytic residues (Asp, Asp, Glu or Asp) to be positioned for catalysis. This motif is known as RNase H-like fold because it was initially found in RNase H and subsequently identified in a variety of polynucletidyl transferases [36]. Importantly, some retroviral integrases are members of the DDE/D polynucletidyl transferase superfamily. Since retroviruses are known to be a medical threat to humans (e.g. human immunodeficiency viruses (HIV)), studying a protein containing the same type of catalytic domain increases the chances of finding important information about these viruses [37]. The DDE/D transposase family catalyzes two types of

reactions represented in Figure 1.1, a cleavage step and a strand transfer step. First, during the cleavage step, the three important residues (DDE/D) are responsible for the coordination of two divalent metal ions with the scissile phosphate of the substrate. After this, a hydroxyl ion does a nucleophilic attack on the scissile phosphate forming the transition intermediate of the reaction. Then, the phosphodiester bond is broken which releases a free 3'-OH leaving group on the substrate. During the strand transfer reaction, the free 3'-OH does a nucleophilic attack on the phosphate of the target DNA. This reaction then allows the breakage of the phosphodiester bond and the formation of a new bond between the target and the donor DNA molecules [35, 36, 38]. The absence of any covalently bound nucleoproteic intermediate during the reaction is a specific feature of the DDE/D family [35].

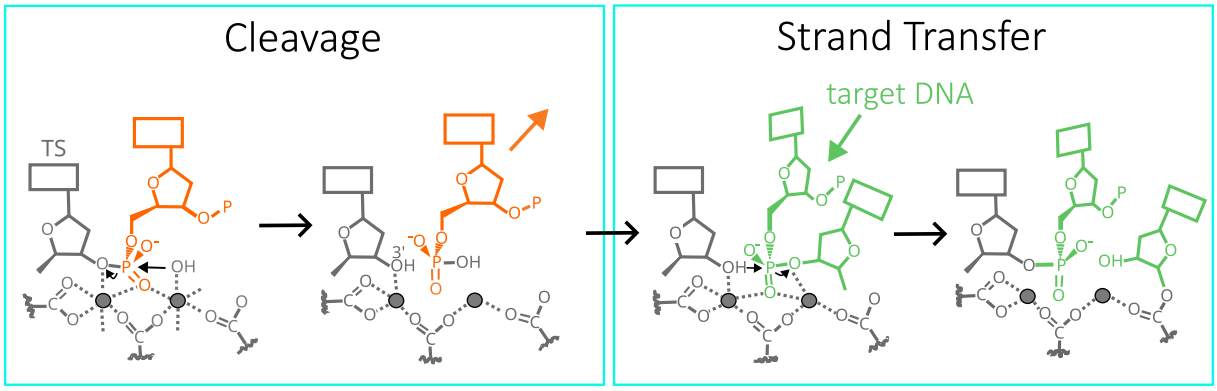

Figure 1.1: Chemical mechanisms catalyzed by DDE/D transposases. Gray DNA indicates the donor molecule. Orange DNA represents the cleaved dinucleotide. Green DNA shows the target DNA strand. Spheres indicate the divalent metal ions. Figure taken from [38].

1.3.1 Transposition mechanisms in DDE/D transposases

DDE/D transposases all share a similar catalytic site but their transposition mechanisms can be divided into two groups: replicative and non-replicative transposition (Figure 1.2) [9].

First, in the replicative transposition, a new copy of the transposon is created by replication. The final product consists of two copies of the same transposon, one that remains in the donor molecule and the other that gets integrated into the target DNA. This mechanism allows the TEs to proliferate within genomes. Replicative transposition has a key feature, it does not create double-strand breaks on the DNA at the transposon end [9]. There are two different mechanisms of replicative transposition. The "copy-in" or "paste-and-copy" mechanism is known to be used by different kinds of transposons, for example, the bacteriophage Mu [39] and the Tn3 family [40]. The other type is called "copy-and-paste" and it is used, for example, by the insertion sequence IS911, a member of the IS3 family [9, 41].

Then, in the non-replicative transposition or "cut-and-paste", the transposon is not duplicated. In this case, double-strand breaks are formed to allow the physical transfer of the whole transposon molecule from the donor DNA to the target [9]. Well-studied models from different

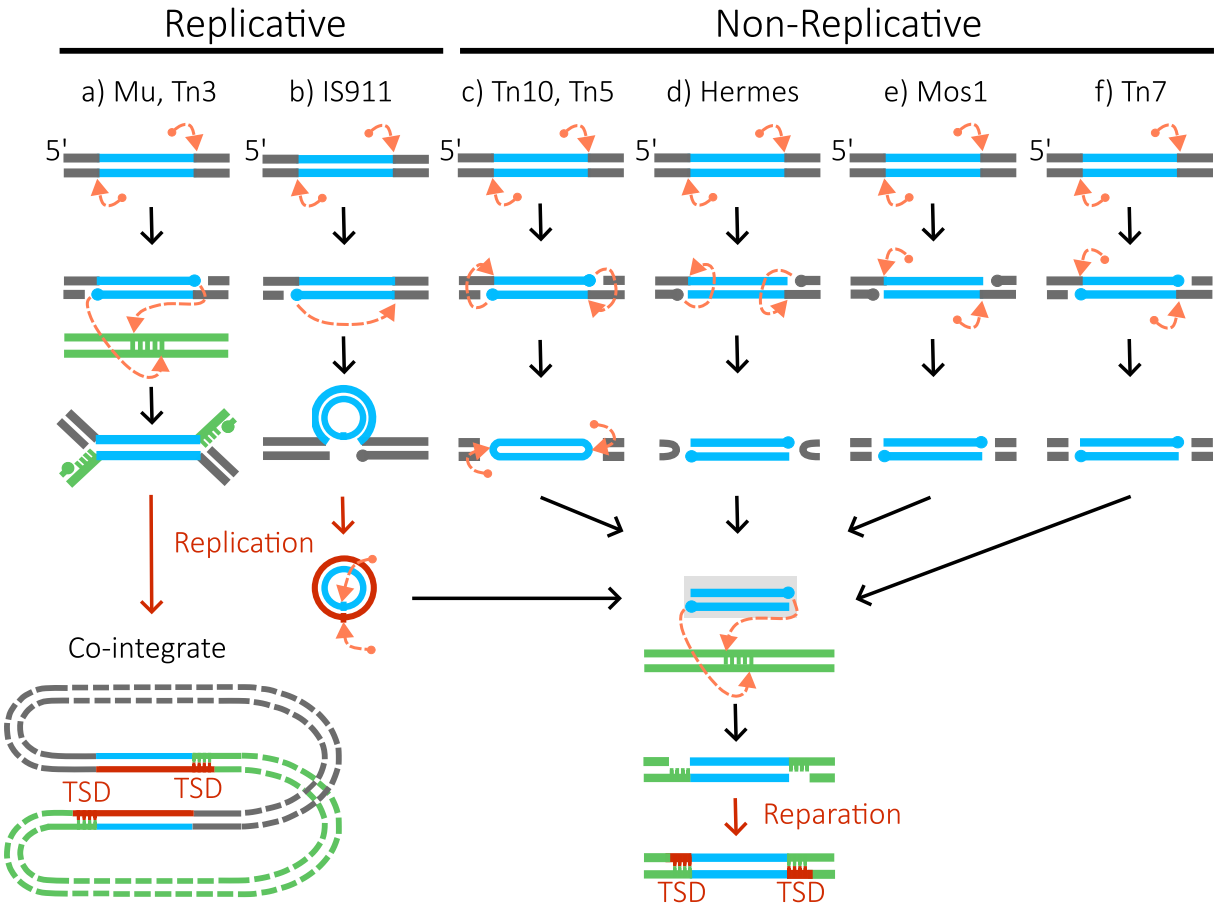

Figure 1.2: Transposition mechanisms of DDE/D transposases family. Blue represents the transposon. Green indicates the target DNA. Gray represents the donor DNA. Blue, green, or grey dots indicate the 3'-OH groups. Red shows the DNA obtained after replication. Orange dots represent water molecules. Orange arrows show nucleophilic attacks. The transposition reaction leads to the duplication of the sequence between the insertion sites of the two transposon ends, which is called Target Site Duplication (TSD). a) Intermolecular transposition by the "copy-in" mechanism, used by phage Mu and Tn3 family, leads to the formation of a co-integrate where the donor and target molecules are linked by two copies of the same transposon. b – f) "Cut-and-paste mechanisms. Pathways leading to the same intermediate of an excised transposon (in a grey rectangle) which is then integrated into the target DNA. Figure adapted from [38].

families of transposon that use this mechanism include Tn5 (IS4 family) [9, 42], Hermes (*hAT* family) [9, 43], Mos1 (Tc1/mariner family) [9, 44], and Tn7 (Tn7 family) [45, 46]. All of them use a different type of the "cut-and-paste" mechanism with various intermediates [9].

1.3.2 Structure of DDE/D transposases

As illustrated in the previous point, DDE/D transposases are very diverse and they can catalyze different kinds of transposition reactions. Although they vary in their architecture and structural domains, they all share the same basic functions: to specifically recognize the transposon ends thanks to the DNA binding domains and to mediate the transesterification reactions thanks to the RNase H-like domain [35, 36, 47].

1.3.2.1 Catalytic domain: RNase H-like fold

The catalytic domain of DDE/D transposases does not share significant similarities in their sequences or size. However, all of them contain a structurally similar region, the RNase H-like fold (Figure 1.3). It consists in a β-sheet constituted of five strands surrounded by three α-helices [36, 44, 47]. In some cases, in addition to this similar fold, the catalytic domain can be interrupted by other sequences. In particular, an "insert" of variable size, structure, sequence, and function, is present in many transposases, interrupting the RNase H-like fold between the last β-strand and the last α-helix. For instance, in Tn5 transposase, the insert is 90 amino acids long, while the same region of Hermes is almost 300 residues long with a different fold. In contrast, in other families, the catalytic domain stays uninterrupted but with an addition of an extra surrounding amino acid sequence. For example, the RNase H-like fold of Mos1 does not have any interruptions but there are two different α-helices flanking it [36].

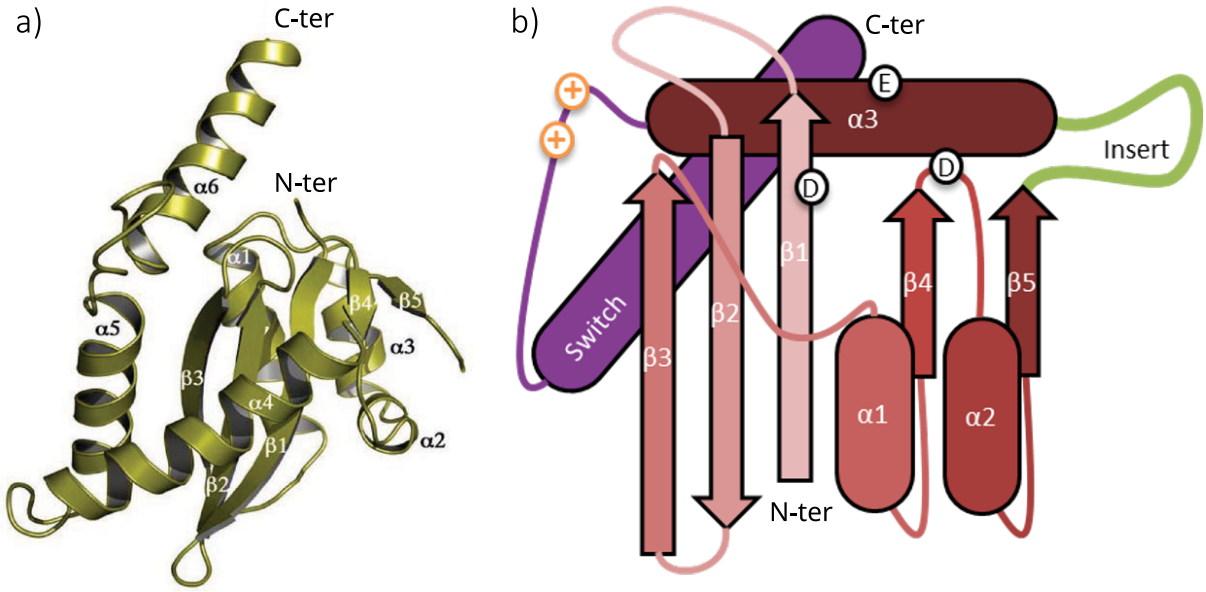

Figure 1.3: RNase H-like fold. The secondary structure elements (β-sheets and α-helices) are numbered according to their position along the protein sequence (from N- to C-terminal (ter)). a) Ribbon diagram of the RNase H domain of HIV-1 integrase. Figure taken from [35]. b) Typical topology of the RNase H-like fold domain showing the conserved succession of secondary structure elements and the position of the insert when present. The catalytic triad (DDE) is shown.

1.3.2.2 DNA-binding domain

The catalytic domain positions and processes the DNA but it is the DNA-binding domain (DBD) that will specifically recognize the transposon ends. This domain is generally located near the N-terminus of the transposase and, contrary to the catalytic domain, DBD is variable in terms of shape, size, and sequence between different transposons [36]. Some transposases have several DBD, while others only have one, and in some cases, they share structural motifs with other DNA binding proteins, like the three helix-turn-helix (HTH), motifs of the bacteriophage Mu transposase [36, 48]. Even though this domain is quite variable between transposases, all

of them share the same function: to specifically bind the transposon ends so that the catalytic domain can process them.

1.4 Regulation of transposition

DNA transposition is done in several successive steps and requires different actors. Successful transposition requires the presence of all the actors, as well as the correct recognition of the transposon ends, their cleavage, and their insertion into a specific DNA target site. In order to avoid non-productive or incomplete transposition activity that could lead to a fatal outcome, for example introducing breaks in the chromosome, all the transposition steps require spatial and temporal coordination. In addition to that, the frequency of transposition events must also be regulated, it has to be sufficiently low to prevent excessive rearrangements in the genome that could cause severe changes in its organization [49].

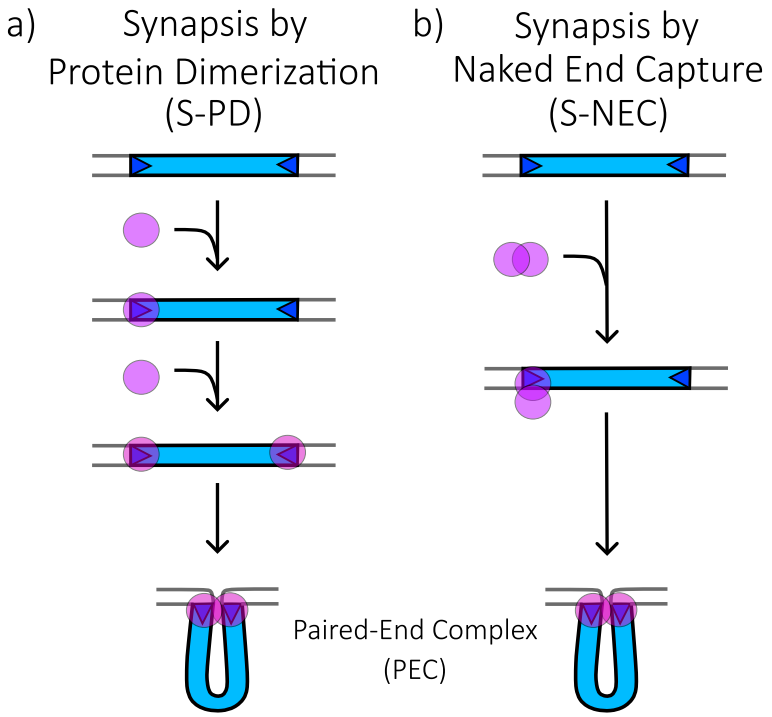

Figure 1.4: Transpososome assembly pathways. Purple spheres indicate the transposase monomers. Light blue represents the transposon. Dark blue triangles show the transposon ends. a) Synapsis by Protein Dimerization (S-PD): a transposase monomer binds to each transposon end and then, the paired-end complex (PEC) is formed by dimerization of the transposase. b) Synapsis by Naked End Capture (S-NEC): a transposase oligomer, in this case, represented as a dimer, contacts one transposon end first and then recruits the other one to form the PEC. Figure adapted from [50, 51].

1.4.1 Transpososome assembly

The formation of a transpososome, or synaptic complex, is needed to have a successful transposition event. It is a nucleoproteic complex composed of a transposase oligomer and

the two transposon ends. This complex is responsible for coordinating and carrying out the reaction and can adopt different conformations as transposition progresses. Transpososome assembly is triggered, in part, by the recognition of the transposon ends by the transposase [49]. This recognition can happen in different ways depending on the oligomeric state of the transposase in solution [51].

Most prokaryotic transposases are monomeric in solution. This means that they can interact independently with the transposon ends, forming the single-end complex (SEC), and then they multimerize and form the paired-end complex (PEC) where both ends of the transposon are brought together. This assembly mechanism is known as synapsis-by-protein-dimerization (S-DP) (Figure 1.4). On the other hand, eukaryotic transposases often have a multimeric state in solution, meaning that a transposase multimer binds to one transposon end and then captures the other end. This mechanism is called synapsis-by-naked-end-capture (S-NEC) (Figure 1.4). In the first case, the transposon ends are simultaneously recruited while in the second, one end is recruited before the other [50, 51]. In the final complex of the transpososome, all transposases are multimeric. This allows to hold the transposon ends together and to perform the DNA cleavage and strand transfer reactions in a concerted way at both ends. In all cases studied, the active core of the transpososome exhibits a "*cis*/*trans*" organization where a specific transposase monomer specifically bound to one end – *cis*-interaction – directs it to the catalytic site of another monomer – *trans*-interaction [12].

1.4.2 Autoregulation: OPI and ASO

Transposons are tolerated by long-lived multicellular organisms. This means that there must be control mechanisms in the host to manage their presence. Indeed, in living organisms there are different types of autoregulation mechanisms: i) the host-mediated regulation, in which the host organism regulates the activity of the transposon, for example, by using the methylation to reduce the expression of transposable elements [52]; and ii) the endogenous regulation which comes from the transposon itself. A type of endogenous regulation is the "overproduction inhibition" (OPI) mechanism (Figure 1.5)[53].

OPI is a self-regulating process that steps in when there is a higher concentration of transposase compared to the substrate, in this case, the transposon ends. This type of autoregulation can happen in the case of an S-NEC mechanism (cf. Section 1.4.1). OPI could lead to the assembly-site-occlusion (ASO) model. Indeed, if the concentration in transposase rises, there will be a decrease in available transposon ends for the reaction to occur because there will be two transposase dimers bound to both TIR of the same transposon. Consequently, the transposition rate will decrease [53].

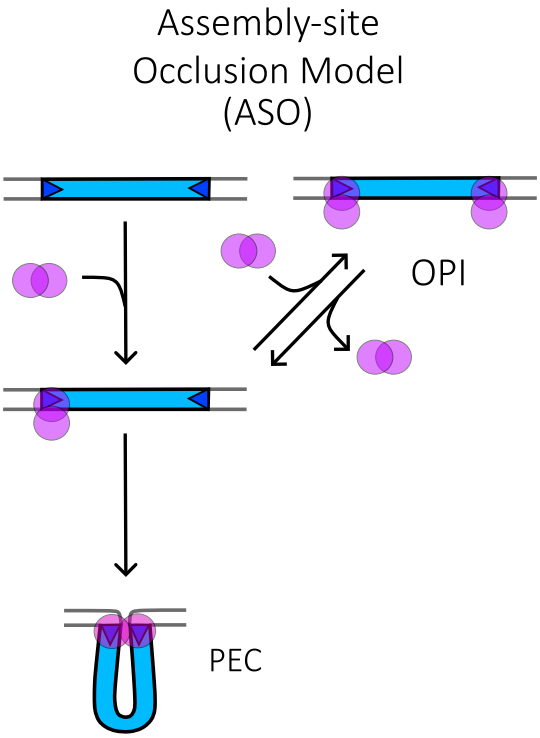

Figure 1.5: Autoregulation mechanism. Purple spheres represent the transposase monomers. Light blue represents the transposon. Dark blue triangles show the transposon ends. Bottom left: the transposase concentration is low, and most of the transposon ends are occupied by only one transposase dimer which leads to a productive synapsis. Top right: the transposase concentration is high, and the transposon ends are occupied by two transposase dimers which cause the overproduction inhibition (OPI) mechanism. Figure adapted from [50].

1.4.3 Target molecule

The target molecule choice is an important level of control in transposition. It has strong implications for the transposon and also for the host because it will determine if the outcome after the transposition is beneficial or not [40].

1.4.3.1 Target recruitment

One level of control on the choice of the target molecule is the moment when it is recruited during transposition. This depends on the type of transposase involved [54]. The target molecule can be recruited before the formation of the transpososome, before the excision of the transposon, for example in Tn7 [54, 55]. On the contrary, in Tn5 and Tn10, the target is captured after excision, when the transposase forms a synaptic complex with the cleaved ends [46, 54]. Finally, we have other transposons in which the target can be recruited before or after excision (Himar1) [54, 56] or at different points of the transposition pathway (phage Mu) [54, 57].

1.4.3.2 Target immunity

Another way of controlling the transposition target is a mechanism known as target immunity. This is a particularly selective process in which a transposon cannot insert itself into a target

DNA molecule that already contains a copy of the same transposon. This mechanism has been described in a few transposon families, such as in members of the Tn3 family, as well as in phage Mu and Tn7. The feature making this regulation extraordinary is the range of activity. It can act over extended regions in the genome, for example from 20 kilobases to several dozen kilobases [40].

Target immunity has different roles in different transposon families. For example, it can avoid self-destruction that would be caused by the insertion of the new transposon into another copy of itself or the insertion of the "moving" transposon into itself. Besides this, it is also an important strategy used to favor intermolecular and intercellular dispersal of transposons. For instance, it prevents a transposable element to accumulate into the chromosome and allows it to spread by horizontal transfer by transposing into other replicons such as conjugative plasmids. Finally, for the Tn3 family, immunity has a broader role of avoiding damaging the donor molecule due to intramolecular replicative transposition which causes adjacent deletion or inversion [40].

1.4.4 DNA topology

Another level of regulation in the transposition mechanism is the topology of DNA. It refers to how the two complementary DNA strands are intertwined and includes supercoiling, catenates, and knots [58, 59]. Linked to this important concept, there is also the topology of the transposon itself which depends on the relative orientation of the transposon ends or their presence on separate DNA molecules [50, 60, 61, 62, 63, 64].

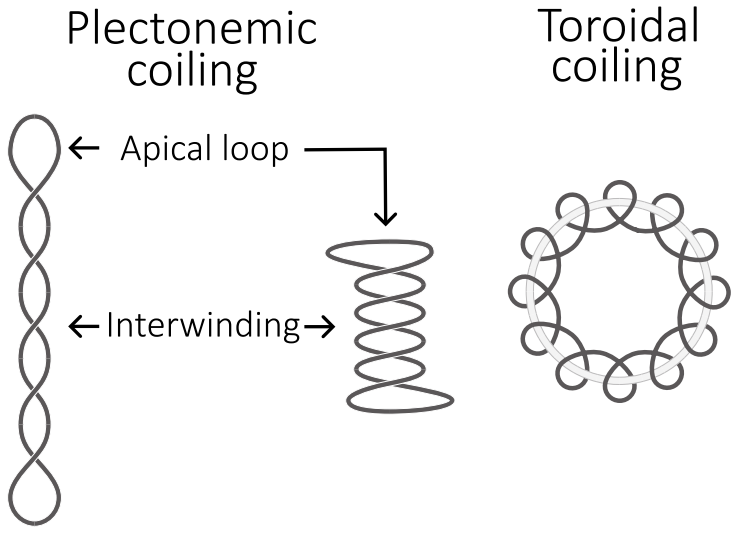

Figure 1.6: Supercoiled DNA packaging. Left: two geometries of plectonemes with the same superhelical density, found in bacteria. Right: a single form of toroidal DNA, found in eukaryotes. Figure adapted from [65].

For some transposons, the topologies of either the donor or target molecule, or both, have an impact on the transposition reaction. It can affect the cleavage step, the transpososome assembly, and the rate of the reaction [61, 64]. This is biologically significant because although DNA is generally negatively supercoiled (with some exceptions), bacteria and eukaryotes package

supercoiled DNA in different forms (Figure 1.6). For example, in bacteria, the dominant form is plectonemic, while in eukaryotes, DNA is wrapped in a toroid around the nucleosome. However, in both cases, there exist proteins that are able to stabilize the other packing mode [64, 65].

Finally, several processes produce high transient levels of free supercoiling in both bacteria and eukaryotes. These processes are DNA replication and transcription [64, 66]. In eukaryotes, there is an additional process: the displacement of nucleosomes associated with episodes of chromatin remodeling [64]. This suggests that free supercoiling is accessible to use as a mechanism of regulation in any system found in eukaryotes and bacteria [64].

1.4.4.1 Bacteriophage Mu

In bacteriophage Mu, DNA topology has been shown to be an essential actor in the transposition reaction. It is specifically necessary for the assembly of the synaptic complex, the transpososome [60, 61, 62, 67]. There are two main features that the donor DNA molecule needs to have. First, the two Mu ends must be in the correct orientation (inverted) on a single DNA molecule [60, 68], and secondly, the donor must be supercoiled [60, 61, 62, 67, 68, 69, 70], and more specifically, negatively supercoiled [70].

These requirements have their significance. First, DNA supercoiling allows a proper binding of Mu transposase (MuA) to the DNA [61, 70]. It also allows the proper positioning of both Mu ends as they are brought together, in the proper orientation, for them to cooperatively participate in the assembly of the transpososome [60, 70]. In addition to that, it has also been described that the transposition reaction has a very high activation enthalpy. This implies that the free energy of the supercoiling is used to lower the activation barrier of the rate-limiting step [61]. Finally, negatively supercoiled DNA donors allow stabilizing the synaptic complex and the end-end interaction [60].

The importance of supercoiling in Mu is illustrated by the fact that the reactions with a relaxed DNA substrate are slow and cannot generate a significant amount of product before the transposase loses its activity. The effect is even greater on a linear substrate which is unable to promote the assembly of a MuA tetramer. Finally, although certain factors, such as the presence of the Integration Host Factor (IHF) protein can modulate the need for supercoiling, they cannot replace it entirely [60, 61, 62].

1.4.4.2 Tn10

Interestingly, the topology requirements differ greatly for other transposons such as Tn10.

In contrast to Mu, Tn10 transposition is very flexible in terms of the transposon ends configuration. The synapsis of both ends, the cleavage, and the strand transfer reactions can happen efficiently regardless of whether the ends are present as direct or inverted repeats or even if they are on different DNA molecules. There are no geometry constraints playing a major role in this case [63].

However, as for bacteriophage Mu, IHF and its related protein, HU, play an important role in the transpososome assembly of Tn10 [71]. In the absence of this factor, there appears to be a dependency on a negatively supercoiled donor DNA [71, 72]. In fact, reactions on a linear substrate, therefore in the absence of negative supercoiling, require the presence of IHF whereas reactions on a supercoiled plasmid do not require it. The fact that IHF can rescue the transposition activity in absence of negative supercoiling is due to its ability to bend the DNA. Upon DNA-binding, IHF introduces a 180° bend into DNA, this kind of bending is essential for the transposition to occur and can also be induced by supercoiling [71]. In addition to that, bending seems to be an important factor in the insertion of Tn10 into its target site. The target DNA needs to be bent because it will allow the formation of a larger interface between the transpososome and the target site and, it is important for the activation of the scissile phosphate during the strand transfer reaction [71].

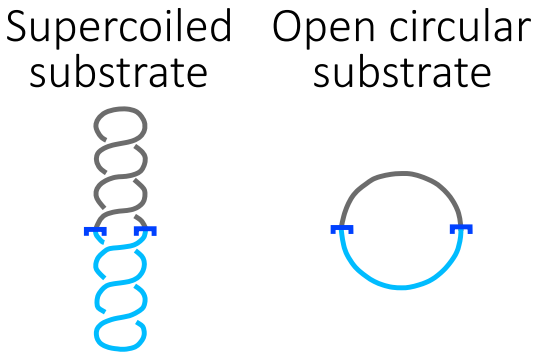

Figure 1.7: DNA substrates. Dark blue brackets represent the transposon ends. Light blue indicates the transposon. Gray represents the DNA donor backbone molecule. Left: DNA supercoiled substrate. Right: open circular (nicked) substrate. Figure adapted from [64].

1.4.4.3 Hsmar1

In the case of the eukaryotic DNA transposon Hsmar1, the transposition reaction appears to be sensitive to DNA topology, at the level of transposon excision and integration [64].

Again, the main feature that the DNA donor requires is to be negatively supercoiled. This is because this configuration accelerates the synapsis of the transposon ends while the positively supercoiled inhibits it [50, 64]. The orientation of both transposon ends is also an important feature. When they are in an inverted orientation, the synapsis is favored [64]. Interestingly, in the case of Hsmar1, the topology of the target is also a factor that affects the transposition with a preference for negatively supercoiled targets [50].

These features have their importance. First, supercoiled DNA increases the concentration between two sites at a local level, increasing the probability of having collision events between two ends in the correct orientation. The juxtaposition of two sites in plectonemic negative supercoils has a specific angular configuration (Figure 1.7). These two factors are responsible for the acceleration of transposon ends synapsis. Secondly, negative supercoiling can increase

the bendability of the DNA by underwinding it. This is the factor that favors the integration of a transposon in a negatively supercoiled target [50, 64].

Finally, for Hsmar1 transposition end orientation as well as supercoiling of the substrate molecules are very important. The DNA donor molecule with inverted repeats and negative supercoiling is the best substrate because it enhances the transposition reaction [50, 64].

In conclusion, there are general conditions shared between different transposons but the exact conditions changes within each family.

1.5 Tn3 family

The Tn3 family is a widespread transposon family, discovered in the 1970s [12]. It is present in almost all bacterial phyla. They are known to be involved in the dissemination of antimicrobial resistance genes, allowing the emergence of what is known as multi-resistant pathogens [40]. Recently, they were involved in the emergence of the carbapenem-resistant *Enterobacteriaceae* (CRE), which are considered to form part of the most severe threats to human health because they are resistant to the most powerful β-lactam antibiotics (carbapenem), considered as the "last chance" antibiotics [73, 74].

One of the characteristics that allows them to be so successful is their replicative transposition mechanism, known as "copy-in" or "paste-and-copy" (Figure 1.8), where the transposon creates a new copy of itself when integrating into the DNA target [40]. During this process, the transposon-encoded transposase (TnpA) catalyzes the cleavage of the DNA and the joining reactions allowing the transposon ends to be joined to the target molecule. Following replication, an intermediate called cointegrate is formed, which carries two copies of the transposon. After this, the cointegrate is resolved by another transposon-encoded protein, the resolvase, producing a donor and a target DNA molecule, both carrying a copy of the same transposon [12, 40, 75].

Tn3-family transposons are characterized by the presence of a transposition module, which contains the DDE/D transposase, and the terminal inverted repeats. They also have a resolvase which can be either a member of the serine recombinase family or the tyrosine recombinase family. Finally, many members of this transposon family contain passenger genes such as genes encoding for antimicrobials or catabolic functions [40].

1.6 Tn4430: a representative model of the Tn3 family

Tn4430 is a representative model of the Tn3 family. It was first isolated from *Bacillus thuringiensis* but it is completely functional in *Escherichia coli* [12, 16, 76]. The use of this transposon model has already given insight into a better understanding of the transposition pathway of the Tn3 family [75]. Tn4430 is delimitated by two TIR of 38 base pairs (bp) and contains one sequence coding for the transposase, denoted as TnpA, and a second one for

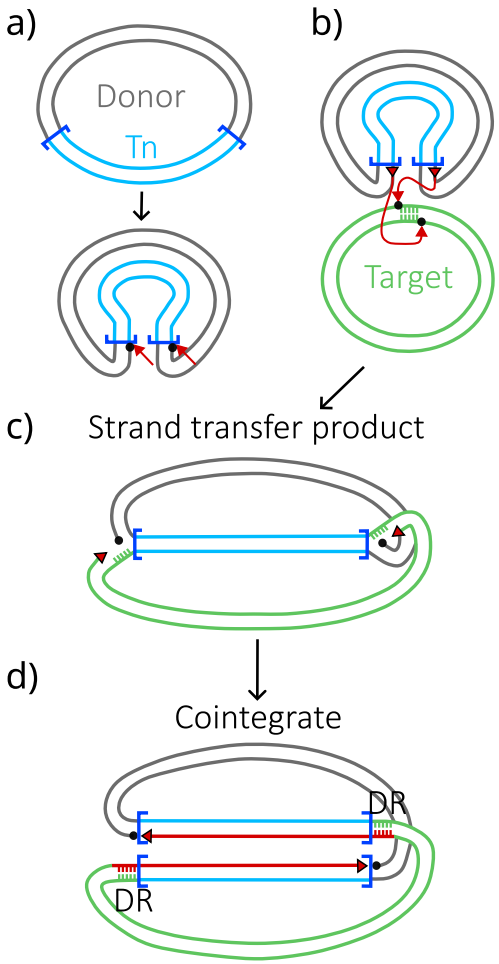

Figure 1.8: Copy-in transposition model of the Tn3 family. Gray and green represent the DNA donor and target molecules, respectively. Light and dark blue indicate the transposon and the transposon ends, respectively. Black dots show the target phosphates. Red arrows represent the free 3'-OH groups. a) Beginning of the transposition by the generation of single-strand breaks by the transposase at both 3' ends of the transposon, creating 3'-OH groups in the DNA donor molecule. b) The 3'-OH ends of the transposon do a nucleophilic attack on the phosphodiester bond on the DNA target strands. c) This reaction generates a strand transfer product that comprises the transposon bound to the DNA donor and target molecules at its 5' ends. d) Finally, the replication starts at the free 3'-OH group, released by the target cleavage, synthesizing the complementary strands of the transposons (red lines). The DNA synthesis will repair the single-strand gaps at the end of the transposable element. There is then a generation of 5 base pairs repeat (DR) duplications that flank the two replicates of the element in the final cointegrate. Figure adapted from [40].

the tyrosine recombinase, TnpI. Besides this, there is also an internal recombination site (IRS) sequence, used by the TnpI to catalyze the DNA site-specific recombination that resolves the cointegrate (Figure 1.9) [77, 78].

1.6.1 Replication hijacking mechanism

The success of Tn3-family transposons relies on their transposition mechanism which creates a new copy of the transposon while integrating into the target DNA. As a result of the "copy-in" mechanism, the after strand transfer reaction produces an intermediate, in which

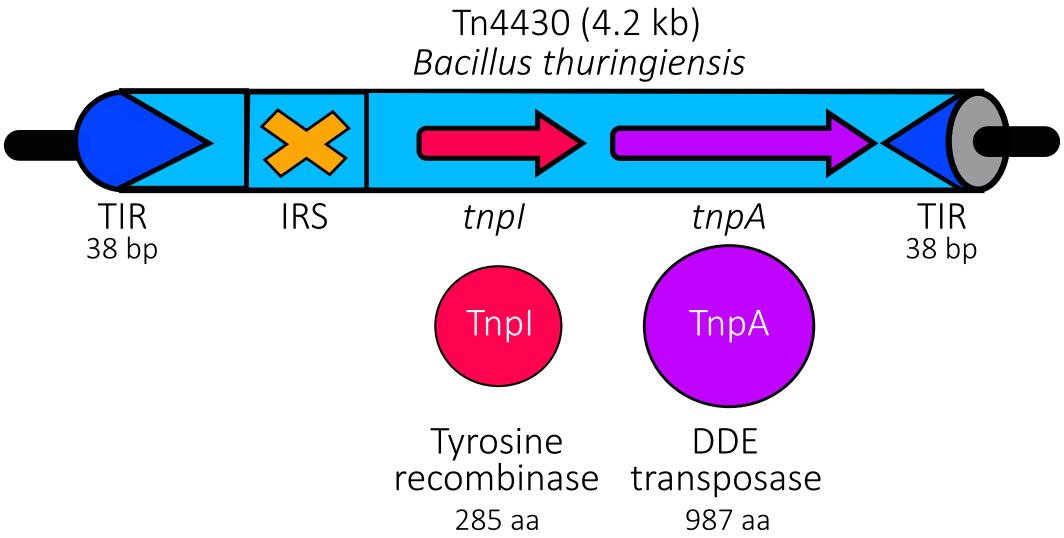

Figure 1.9: Tn4430 structure. Light blue represents the transposon. Dark blue arrows indicate the transposon ends (TIR). The orange cross shows the position of the internal recombination site (IRS), necessary for the resolution of the cointegrate by site-specific recombination. Pink and purple arrows represent the genes coding for the tyrosine recombinase (TnpI) and the transposase (TnpA), respectively.

each transposon end is bound by one DNA strand to the donor and by the other strand to the target creating forked junctions (Figure 1.8). This intermediate is resolved by the passage of the replication machinery. In the "replication hiring" model for "copy-in" transposition proposed in textbooks based on bacteriophage Mu (Figure 1.10(a)), the replication is recruited after the formation of the strand transfer intermediate by the transposase. For Mu, this requires the successive action of different host factors to first disassemble the transposition complex, and then promote the assembly of the replisome at one end of the Mu genome in a primosome-dependent mechanism. Our work on Tn4430 supports a conceptually distinct mechanism called "replication hijacking" (Figure 1.10(b)) during which replication is recruited concomitantly to the reactions catalyzed by the transposase TnpA. In fact, the transposon cannot transpose successfully unless there is replication happening on the target molecule. We believe that is because it needs to hijack an already existing replication machinery to integrate into the target. Consistently, Tn4430 was found to preferentially insert in DNA regions where replication forks slow down or stop and restart *in vivo*. In addition, DNA molecules mimicking replication intermediates such as DNA forks were shown to be the preferred substrate for TnpA-catalyzed strand transfer *in vitro* [13, 79]. This type of mechanism is thought to allow the DNA replication and transposition mechanism to be synchronized, which increases the chance of a proper resolution of the Shapiro intermediate [13].

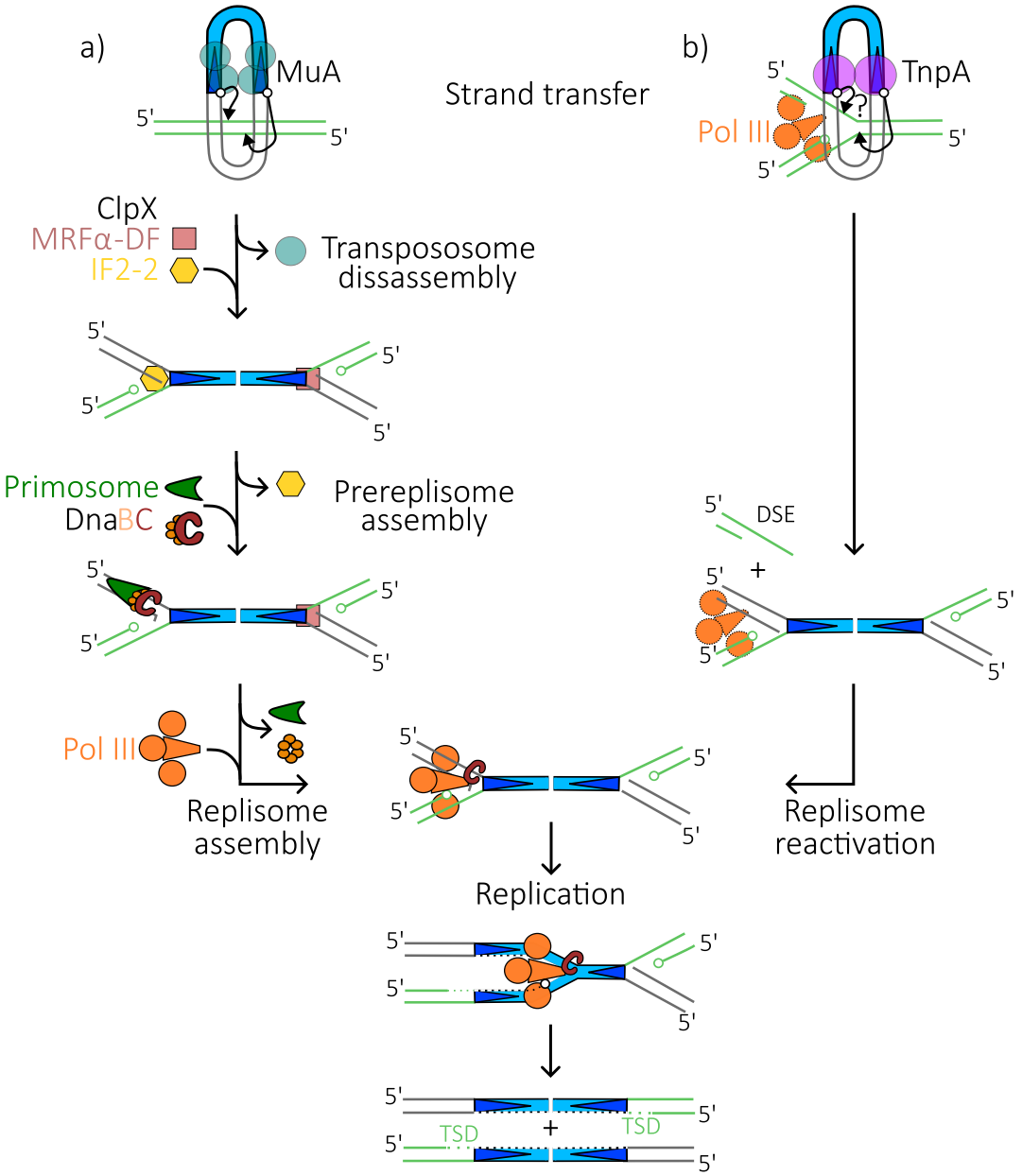

Figure 1.10: Copy-in replicative transposition models. a) "Replication hiring" mechanism found in bacteriophage Mu. During this mechanism, a tetramer of MuA transposase (turquoise circles) joins both Mu ends to the target DNA duplex, in a concerted way (curved arrows). The transpososome is disassembled in an ATP-dependent reaction that involves host factors: the ClpX chaperone, MRFα-DF (a replication-disassembly factor), and a truncated version of IF2-2 (translation initiation factor). IF2-2 is bound to one end and allows the recruiting of the primosome proteins which give access to the helicase loading complex DnaB-DnaC. The loading of DnaB helicase onto the lagging strand template promotes the assembly of DNA polymerase III holoenzyme and the replication of the Mu genome by the replisome. b) "Replication hijacking" model proposed for Tn4430 and Tn3 transposons family. The transpososome is formed when the TnpA dimer (purple circles) contacts the replication machinery on the target (green) and joins both transposon ends to a DNA replication intermediate which is represented here as a replication fork. One end will be transferred to one end of the fork. Replisome is reactivated by an unknown mechanism that promotes replication of the transposon. The second transposon end is suggested to be transferred somewhere in the lagging strand template (arrow with an interrogation point). This is based on the fact that Tn4430 generates 5-bp target site duplications (TSD). The reaction generates a double-strand end (DSE) which is going to be either repaired by homologous recombination or degraded. Dark blue arrows represent Mu or Tn4430 ends. Light blue indicates Mu or Tn4430. Green and grey show the target and the donor DNA, respectively. Open circles represent the position of free 3'-OH ends. Dashed lines indicate the newly synthesized DNA strands. Figure adapted from [13].

1.6.2 Assembly and activation of Tn4430 transposition complex

1.6.2.1 TnpA binding to the transposon end and the target

The TnpA transposase of Tn4430 transposon is a member of the DDE/D transposase superfamily and is big in comparison to other DDE/D transposases members (987 amino acids) [12, 75]. TnpA is responsible for the target immunity (cf. Section 1.4.3.2). The mechanisms regulating this phenomenon are still poorly understood but it has been shown that transposition and immunity can be uncoupled. In fact, we have identified and characterized TnpA mutants that can still transpose but have lost their ability to mediate immunity (T+/I–) [75]. Many different mutations can give this phenotype but two mutants, in particular, have shown higher activity than the $TnpA^{WT}$: $TnpA^{S911R}$ and $TnpA^{3X}$ which combines three different mutations: W24R, A174V, and E740G (Figure 1.11(b)). In correlation with this lack of immunity, these hyperactive mutants are more prone to adopt the active configuration than the wild-type (WT) transposase in the absence of an appropriate target. Moreover, these two mutants have allowed us to obtain more *in vitro* insights into the different steps of the transposition mechanism of the Tn3 family [12]. During this pathway, the TnpA, present as a dimer in solution, binds the DNA of one of the transposon ends and it creates an inactive single-end complex (SEC). Then, TnpA changes to an activated conformation, the paired-end complex (PEC), in which it is bound to both transposon ends. In this state, the TnpA is able to make single-strand DNA breaks and catalyze strand transfer [80].

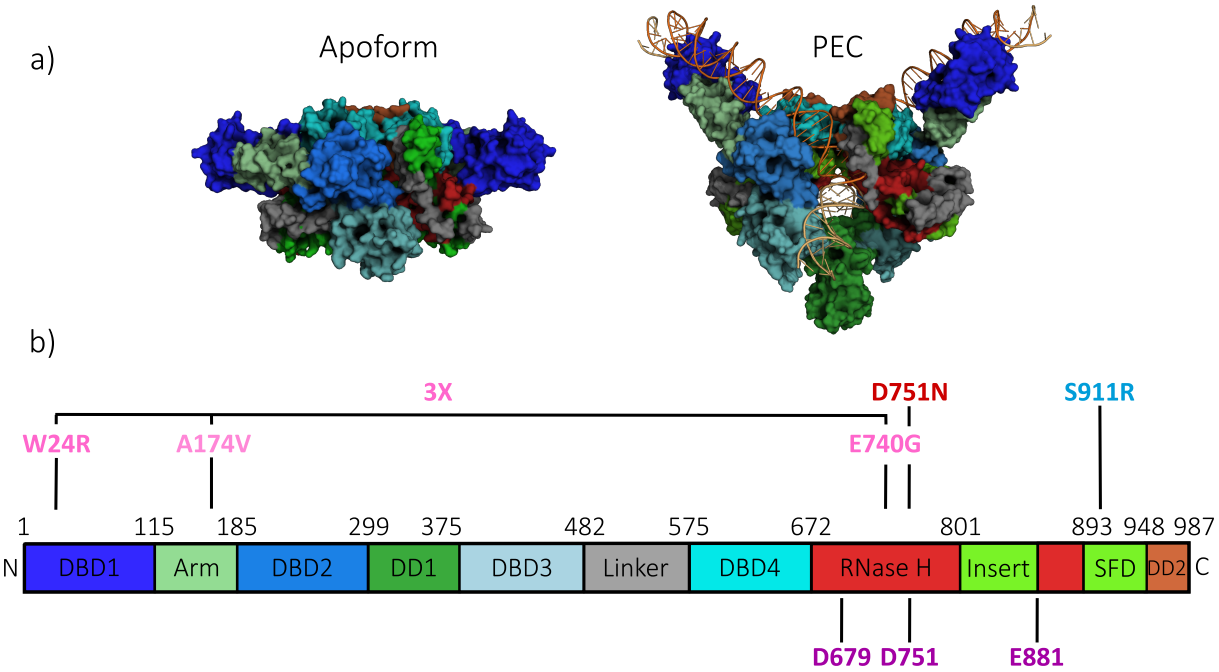

Figure 1.11: TnpA architecture. a) Cryo-EM reconstitution of the two conformations of TnpA: $TnpA^{WT}$ apoform and $TnpA^{S911R}$ PEC state. b) Linear diagram of the structural domains found in TnpA. The position of the T+/I– mutations in $TnpA^{3X}$ (pink) and $TnpA^{S911R}$ (blue), the catalytic mutant ($TnpA^{D751N}$, red), and the DDE residues of the active site (purple) are shown. DBD: DNA-binding domain, DD: dimerization domain, SFD: scaffold domain. Figure adapted from [81].

1.6.2.2 Cryo-EM structures of the transposase and the DNA-bound complexes

Thanks to a collaboration with Ruslan Efremov and Alexander Shkumatov from the Vlaams Instituut voor Biotechnologie (VIB) at the Vrije Universiteit Brussel (VUB), our team was able to obtain the structures of the TnpA by Cryogenic Electron Microscopy (Cryo-EM) (Figure 1.11(a)). The structure of the apoform of the TnpA was solved with the $TnpA^{WT}$. However, due to very low TnpA-DNA complexes found with the $TnpA^{WT}$, the PEC state was solved with the hyperactive mutant, $TnpA^{S911R}$. This could be mainly because of the absence of the appropriate target [81]. Additionally, a structure for the post-strand transfer complex was solved with the $TnpA^{S911R}$ mutant. It takes an almost identical conformation to the PEC, but for the DNA, the TIR is linked to the target molecule. However, the resolution of the target DNA molecule in the cryo-EM images was not sufficient to accurately model its position. Finally, although the TIR takes the same position as it does in the PEC, the nicked donor DNA molecule, at the external part of the transposon, is subjected to a 120° rotation around the double helix axis which drives it away from the active site [81].

Each monomer of the TnpA is constituted by ten structural domains (Figure 1.11(b)). It contains four DNA-binding domains (DBD) and two dimerization domains (DD). It also has a protrusion of a long linker (LN) that brings two of the DBD and the catalytic RNase H-like domain together forming the scaffold domain (SFD). In the substrate-bound structure, at the center of the dimer, there is the scissile phosphate facing the RNase H-like domain of a TnpA monomer [81]. During the apo-PEC transition, each TnpA monomer is bound to one transposon end allowing them to converge to the center of the TnpA dimer and thus have the DNA cleavage sites within the active site of the transposase. The cleavage reaction happens in a cross-reactivity mechanism. One subunit recognizes a specific DNA sequence by *cis*-interactions and the other subunit catalyzes, in a *trans*-interaction, the cleavage of the DNA strand and the strand transfer. This is a striking characteristic that has converged in most DDE/D transposases despite the differences in their structure [81].

1.6.2.3 Models for transposition complex assembly

By combining microscopy with biochemical analyses, we are now able to describe the different complexes formed by the TnpA before strand transfer. In solution, in the absence of any substrate, the transposase dimer is present in a closed conformation (apoprotein). In the presence of forked DNA oligonucleotides of random sequences, it makes a low-stability Target Complex (TC). In the presence of its specific substrate, Tn4430 ends, the $TnpA^{WT}$ forms a catalytically inactive single-end complex (SEC). Finally, in specific artificially optimized conditions (e.g. cleaved IR substrates, pre-nicked IR substrates, high glycerol concentrations, etc.), the WT transposase is able to form an active paired-end complex where the dimer opens up to allow access to the catalytic site and takes a V shape. The formation of PEC is greatly facilitated in deregulated and hyperactive mutants ($TnpA^{3X}$ and $TnpA^{S911R}$). The final complex containing the ends of the transposon and the target is still unresolved but has been modeled.

Overall it is currently considered that the transition from apoform to PEC is the regulatory step controlling transposition activity and that the recruitment of the target prior to the PEC formation is necessary (Figure 1.12) [81].

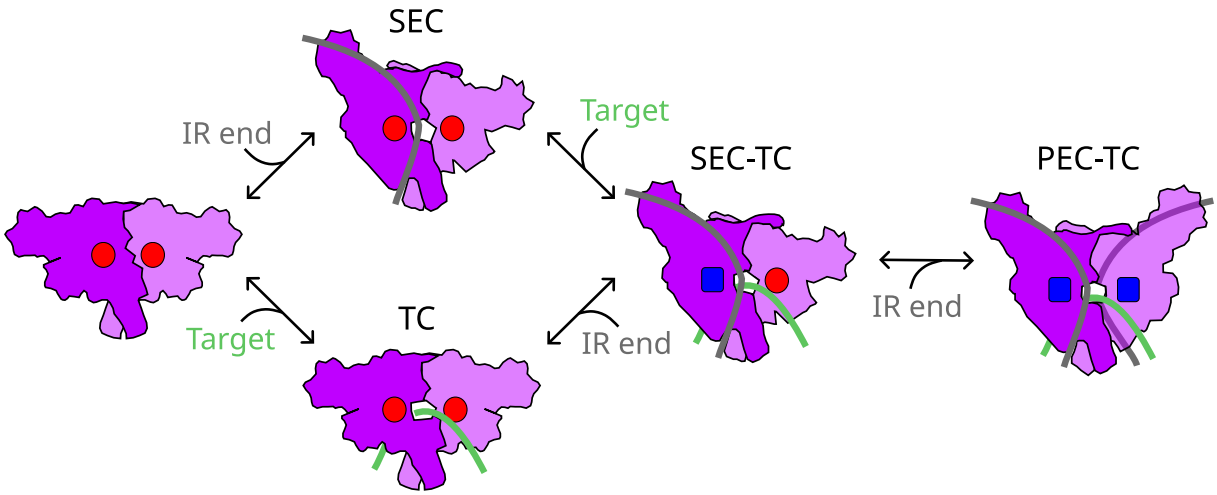

Figure 1.12: Model of transpososome assembly proposed for Tn4430 and Tn3-family transposons. Purple represents the transposase. Gray indicates the DNA donor. Green indicates the DNA target. The red circle shows the inactive RNase H. The blue square represents the active RNase H. The pathway starts on the left with the TnpA, in solution without any substrate, as a closed dimer. Then there can be two possibilities: either one monomer binds one IR end (top) or it binds the target DNA (bottom). In both ways, we have an inactive RNase H. After that, the transposase monomer will bind again to either the other IR end or the target molecule, depending on what was bound first, and the left RNase H domain is going to be activated but the reaction still cannot happen. Finally, the second IR end is captured by the second TnpA monomer which is going to activate the other RNase H domain and the reaction can occur.

2 | Objectives

Nowadays, the Tn3 transposons family has an important impact on the living world because of its implication in the dissemination and persistence of antibiotic resistance genes. Despite this biological importance and in comparison with other transposons families, there are still many unknown molecular aspects of the transposition mechanism of this family [12].

In our lab, we work with the transposon Tn4430, and its transposase, as a representative model for the Tn3 family. To date, *in vitro* studies of the transposition mechanism of the Tn4430 were done with the use of simplified DNA segments mimicking the donor (typically just one end of the transposon surrounded by flanking DNA) and the target substrate. These substrates were assembled by specific annealing of oligonucleotides. With them, we were able to visualize cleavage activity by the transposase and strand transfer products. However, the activity of the transposase in those assays often remains weak and requires particular conditions such as pre-processing of the donor DNA substrates, high glycerol concentrations, or hyperactive mutants of the transposase, specifically for strand transfer assays [12, 82].

To go further in the understanding of the Tn4430 transposition mechanism (e.g. the role of the target, the possible contribution of hosts factors and the order of events during transposition), a new experimental system has been developed based on DNA substrates closer to the *in vivo* conditions. For this, circular DNA molecules containing a functional derivative of Tn4430 ("Mini-Tn4430") with properly oriented ends were created. These substrates allow us to assess the importance of flanking DNA, transposon ends orientation, and supercoiling during the cleavage and then strand transfer reactions. In a previous project ([83]; M.C. Fernandez, unpublished data), the cleavage activity of the TnpA hyperactive mutants with these new substrates was examined and proved to be highly efficient with deregulated (i.e. immunity deficient) TnpA mutants. Interestingly, the results indicated that donor DNA topology might have an impact on TnpA activity. In particular, negative supercoiling of the donor DNA seemed to reduce TnpA cleavage activity *in vitro* [83].

Furthermore, in some experiments, unexpected products appeared that could not be explained solely by the cleavage activity of TnpA [83]. These products have not been characterized yet but Atomic Force Microscopy (AFM) imaging of these reactions identified multiple loop structures that likely correspond to auto-insertion complexes potentially giving rise to

these unidentified products (M. C. Fernandez, unpublished data).

This project aims first to further optimize the transposition reaction with circular plasmids and linear substrates by establishing the optimal temperature and changing the concentration of TnpA to see whether the transposition complex assembly is controlled by assembly site occlusion.

Then, we will continue to investigate whether DNA topology has an impact on the TnpA activity. To assess this, different configurations of the donor DNA molecule (either supercoiled, open circular, or linearized) will be incubated along with the TnpA mutants.

Besides that, it is important to determine if the Tn4430 transposition reaction requires correctly oriented transposon ends on the same DNA molecule. To this end, reactions will be performed with plasmids containing a single transposon end or two ends in direct repeats instead of inverted.

Finally, we will also be investigating and checking if the hypothetic "auto-insertion products" are reproducible and start their identification. This will be done by recreating the same gels where the additional bands appeared, followed by tests with different restriction enzymes.

These experiments will be performed and further optimized with "hyperactive" variants of TnpA ($TnpA^{S911R}$ and $TnpA^{3X}$) with the prospect of exploring conditions required to activate transposition under wild-type situation.

3 Results

3.1 Optimization of the conditions for plasmid cleavage by the transposase

During a recent project [83], a new system was developed to get closer to the *in vivo* conditions for transposition by using a plasmid substrate containing a Mini-Tn4430 and flanked by the TIR ends of Tn4430 in their natural inverted orientation. TnpA-mediated cleavage of the substrate generates a nick at the 3' end of either one or the other TIR sequence or both. In order to detect this activity on an agarose gel, we then use a restriction enzyme (Nt.BspQI) that nicks the ends at their 5' external extremity transforming any eventual nick caused by the TnpA into a double-strand break (Figure 3.1). If only one end was processed by the transposase, this creates a linearized plasmid, whereas if both ends were processed, the whole Mini-Tn4430 segment is excised from the plasmid backbone (Figure 3.1). Using this assay, previous experiments allowed us to measure relatively high levels of DNA cleavage with the hyperactive mutants (TnpA^{3X} and

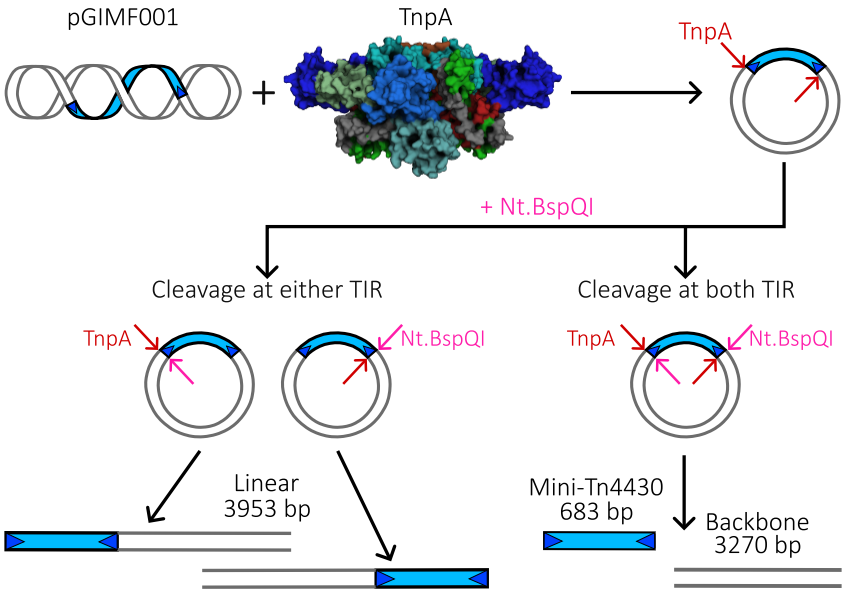

Figure 3.1: Scheme of the reaction between the substrates (in this case the negative supercoiling form of pGMIF001) and the transposase and the expected single- and double-cleavage products.

TnpA^{S911R}) but not with the TnpA^{WT}. The presence of a proper cofactor (Mn_2 instead of Mg_2) and a certain concentration of glycerol were found to greatly improve the activity of the mutants TnpAs, but the reaction with the TnpA^{WT} remained weak, requiring further optimization.

This is why, during this project, the optimal temperature as well as the hypothesis of OPI/ASO control of transposition were tested.

3.1.1 Optimal temperature

In other transposons families, the optimal reaction temperature varies, possibly in correlation with the host organism's preferred temperature. For instance, for Tn10, a transposon from the enteric bacterium *Shigella flexneri* [84], transposition is most effective between 28°C and 32°C [85], and for bacteriophage Mu that infects *Escherichia coli*, the optimal temperature is 30°C [60, 70, 86]. On the other hand, for Hsmar1, a human transposon, the best temperature is 37°C [87].

To determine the optimal reaction temperature for Tn4430, the TnpA^{3X} mutant was incubated with the supercoiled pGIMF001 plasmid at different temperatures (Figure 3.2(a)). As detailed above, we obtained the three expected cleavage products (linear fragment, backbone, and Mini-Tn4430) at the different temperatures. High amounts of products were obtained between 25 and 34°C. Indeed, this can be confirmed thanks to the quantification of the gel (Figure 3.2(b)), where we can observe that between these temperatures more than half of the initial substrate was converted into product. Unfortunately, the samples incubated at 27 and

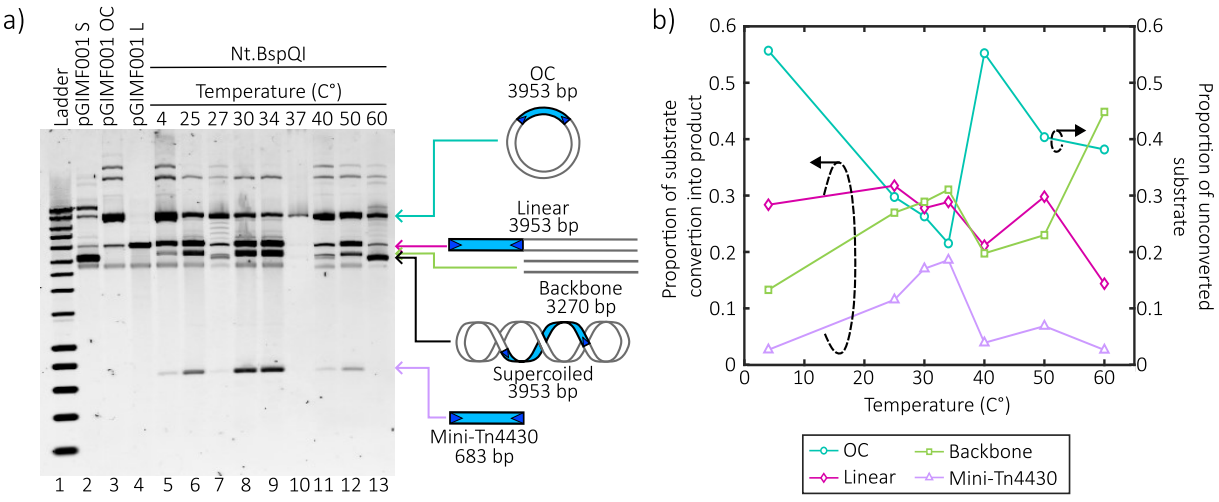

Figure 3.2: Cleavage reaction between the TnpA^{3X} mutant (300 nM) and the supercoiled pGIMF001 plasmid (6.49 nM) under different temperatures (°C). Lanes 2 – 4 correspond to the three different forms of the plasmid: supercoiled (S), open circular (OC), and linear (L). The last two were obtained by digesting the plasmid with Nt.BsmAI and NcoI, respectively. Lanes 5 – 13 show the cleavage reaction at different temperatures and each reaction was incubated for 2 hours. 1.2% agarose gel and SYBR Gold staining. b) Quantification of the evolution of the proportion of the substrate unconverted and converted into products, performed using Image J. The cleavage products (linear, backbone, and Mini-Tn4430) use the left Y-axis, and the OC substrate uses the right Y-axis (cf. Section 6.3.4).

37°C have been corrupted by a technical mishap, preventing us from determining a precise optimal temperature. Low to no activity was detected at 4°C, and between 40 and 60°C, reflecting the specific catalytic properties of TnpA^{3X} or its thermal stability. In addition to that, in general, proteins start to denature in ranges between 0 and 10°C, and 40 to 80°C [88, 89].

These results are similar to the optimal temperature determined for other bacterial transposases activity (Tn10 and phage Mu) and the optimal growth temperature for Tn4430's original host organism, *Bacillus thuringiensis* (30°C, under laboratory conditions) [90]. A finer estimate of the optimal temperature could be found if we reiterate the experiment with a more precise gradient between 27 and 37°C. Nevertheless, the results are consistent with former estimations obtained with oligonucleotide substrates (E. Nicolas, unpublished data), and the temperature of 34°C was chosen for further experiments (cf. Section 6.3.1).

3.1.2 Is TnpA reaction subjected to assembly site occlusion (ASO)?

Overproduction inhibition (OPI) is an autoregulation mechanism that takes place *in vivo* when there is an excess of protein compared to the substrate (cf. Section 1.4.2). For transposons that use a naked end capture pathway for synaptic complex formation (S-NEC), OPI can lead to transposition inhibition through assembly site occlusion (ASO), a process in which binding of a multimeric form of the transposase occupies both transposon ends, thereby preventing assembly of a productive synapse. Our models for Tn4430 transposition complex assembly during which a TnpA dimer first binds to one transposon end to form the single-end complex (SEC) and then captures the other end to form the active paired-end complex (PEC) suggests that transposition could be subjected to ASO (cf. Section 1.6.2). If OPI and ASO occur, the enzymatic activity would typically increase proportionally to enzyme concentration until a threshold is reached from where it would start to decrease. Interestingly, in the presence of a short oligomeric substrate representing a single transposon end, at a certain threshold of TnpA concentration, the inactive SEC complex is formed in priority at the expense of the higher order, active, PEC [12]. This observation might indicate that an ASO-like phenomenon can happen in our *in vitro* experiments.

To test this possibility, different transposase-to-substrate ratios were tested in our *in vitro* experiments (Figure 3.3).

First, we measured the activity of a gradient of the TnpA^{3X} mutant (from 0 to 300 nM) on a constant concentration of the plasmid substrate (16.67 ng/μl = 6.49 nM, cf. Section 6.3.1) during a two-hour incubation. In Figure 3.3(a), we can see that when the protein-DNA ratio increases, there is more consumption of the OC substrate which yields increasing amounts of the single- and double-cleavage products, indicating a higher TnpA activity. However, we can also observe that at a higher protein-DNA ratio (46.22, lane 16) the proportion of Mini-Tn4430 seems to slightly decrease. Indeed, in Figure 3.3(c), at the highest ratio, more than half of the OC form was consumed while the proportion of Mini-Tn4430 decreased. Although it is a weak effect, this final decrease of Mini-Tn4430 could indicate that we have reached the protein

concentration threshold at which ASO starts to occur.

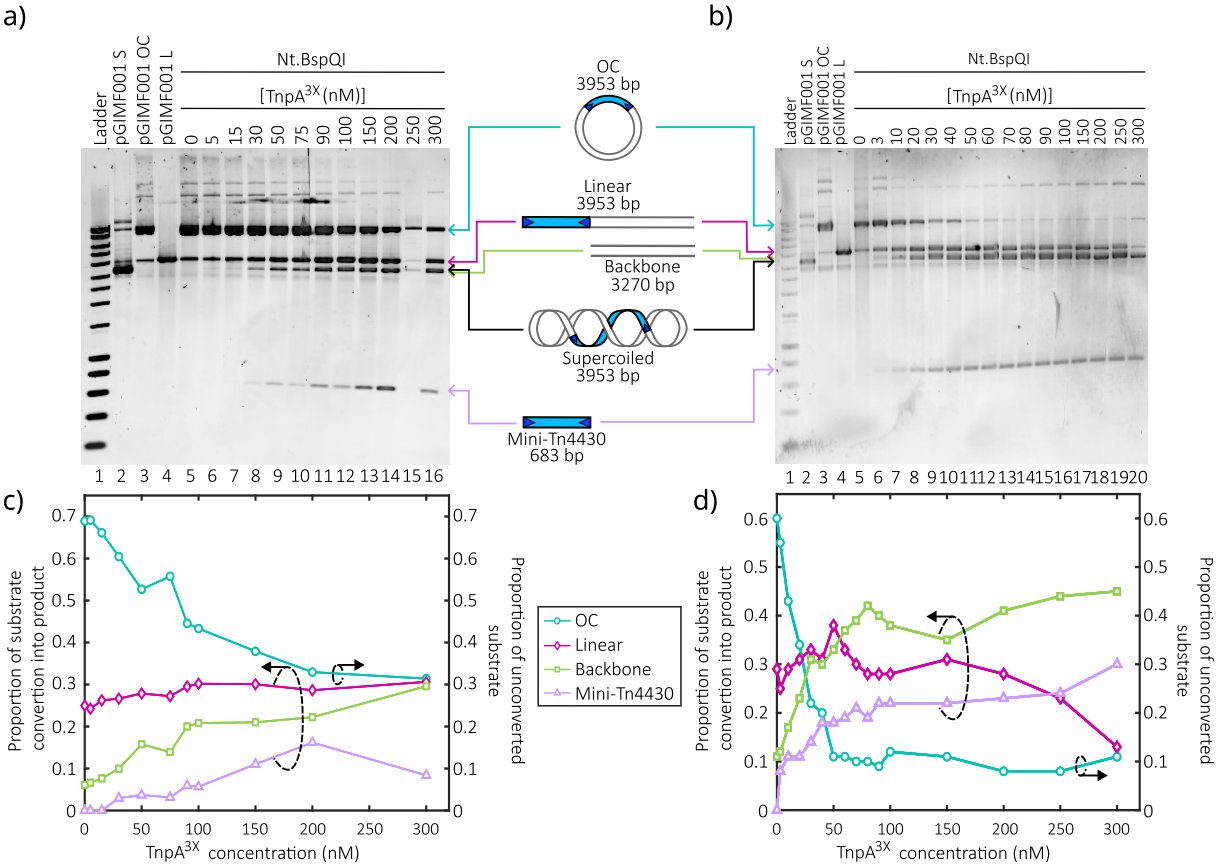

Figure 3.3: Comparison of the cleavage reaction between the supercoiled pGMIF001 plasmid at a constant concentration and the TnpA^{3X} at a variable concentration. Reactions were incubated for 2 hours at 34°C. 1.2% agarose gel and SYBR Gold staining. Lanes 2 – 4 correspond to the three different forms of our plasmid: supercoiled (S), open circular (OC), and linear (L). The last two were obtained by digesting S with Nt.BsmAI and NcoI, respectively. a) Cleavage reaction between 6.49 nM of supercoiled substrate with variable concentrations of TnpA from 0 to 300 nM (lanes 5 – 16), corresponding to protein-to-substrate ratios of: 0, 0.77, 2.31, 4.62, 7.7, 11.56, 13.87, 15.41, 23.11, 30.82, 38.52, and 46.22. b) Cleavage reaction between 1.56 nM of supercoiled substrate with variable concentrations of TnpA from 0 to 300 nM (lanes 5 – 20), corresponding to protein-to-substrate ratios of: : 0, 1.92; 6.41, 12.82, 19.23, 25.64, 32.05, 38.46, 44.87, 51.29, 57.69, 64.10, 96.15, 128.21, 160.26, and 192.31. c) and d) Quantification of the gels a) and b), respectively, performed using Image J (cf. Section 6.3.4). The cleavage products (linear, backbone, and Mini-Tn4430) use the left Y-axis, and the OC substrate uses the right Y-axis.

Secondly, we, therefore, decided to repeat the experiment with a higher protein-to-DNA ratio. Because the maximum concentration of TnpA that we could reach in our assays is 300 nM, instead of increasing the protein concentration, we reduced the DNA concentration to 1.56 nM (or 4 ng/μl). This concentration was chosen to have an equable concentration of transposase and substrate at the highest TnpA concentration (300 nM). Usually, the TnpA concentration used is 300 nM, but this corresponds to a TnpA monomer. In terms of TnpA dimer, we would have a concentration of 150 nM. In the end, at the highest TnpA concentration, there are thus $\sim$ 50 times more TnpA dimers than Tn4430 ends. The results, in Figure 3.3(b), show this time a quick conversion of the OC form into the different products after reaching

a plateau where almost all the OC form is converted, the accumulation of Mini-Tn4430 has slowed down but continues until the highest TnpA concentrations.

Then, to further test the hypothesis of an ASO phenomenon intervening in Tn4430 transposition, we wanted to increase the protein-DNA ratio to have even more transposase than substrate. To do this we decreased the DNA concentration instead of the amount of protein, especially because the highest TnpA concentration that can be used is limited by our TnpA purification yields. Therefore, we tested the activity of a constant amount of TnpA on a gradient of plasmid DNA substrate (from 0.52 to 5.19 nM (being 1.33 to 13.33 ng/μl)). Allowing us to reach protein-substrate ratios of $\sim$ 600. In Figure 3.4, we can see that at a higher protein-DNA ratio (578.13, lane 5) – at the lowest substrate concentration – the OC substrate is almost completely consumed, while at the lowest protein-DNA ratio (46.25, lane 13), there is still a huge proportion of the OC form and it even seems to have a slight decrease in the proportion of Mini-Tn4430 (lane 13).

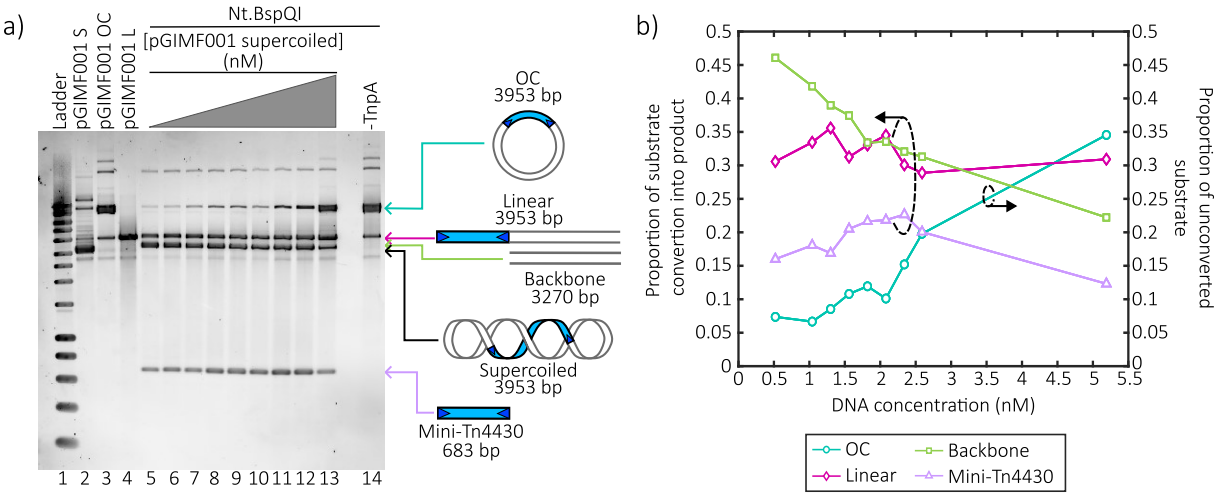

Figure 3.4: a) Cleavage reaction between the TnpA^{3X} mutant at a constant concentration of 300 nM and the supercoiled pGIMF001 plasmid at variable concentrations. Lanes 2 – 4 correspond to the three different forms of our plasmid: supercoiled (S), open circular (OC), and linear (L). The last two were obtained by digesting S with Nt.BsmAI and NcoI, respectively. Lanes 5 – 13 show the cleavage reaction with different concentrations in nM of the supercoiled pGIMF001 plasmid (resp. protein-to-substrate ratios): 0.52 (578.13), 1.04 (289.06), 1.30 (231.25), 1.56 (192.71), 1.82 (165.18), 2.08 (144.53), 2.34 (128.47), 2.59 (115.63), 5.19 (46.25). Lane 14 represents the negative control without TnpA. Reactions were incubated for 2 hours at 34°C. 1.2% agarose gel and SYBR Gold staining. b) quantification of the gel a), performed using Image J (cf. Section 6.3.4). The cleavage products (linear, backbone, and Mini-Tn4430) use the left Y-axis, and the OC substrate uses the right Y-axis.

Taken together, these results show that with a gradient of TnpA, the activity seems to increase with the protein-DNA ratio with a typical saturation curve, up to reach a plateau when the substrate is totally consumed. Therefore, we cannot confirm that there is any ASO phenomenon occurring in these assays even if there seems to be a slight decrease in the accumulation of the products. Moreover, these results cannot be used to extrapolate to what can happen *in vivo* with the TnpA^{WT} because i) we did not test it, and ii) the behavior of the hyperactive mutants compared to the wild-type is different, with the TnpA^{3X} mutant, the

PEC state is formed faster and thus could prevent the possible ASO mechanism. The PEC formation with the TnpA^{WT} is slower so it cannot prevent the ASO and we hypothesized that is indeed the case but it has to be tested.

Finally, the ASO phenomenon should be further studied because it seems that it is not taking place with the TnpA^{3X} and it would be necessary to see if it is the case in the *in vivo* conditions. In order to improve these assays, it would be interesting to do a more rigorous kinetic analysis when instead of measuring the final formation of cleavage products, we measure the initial slopes of the reaction at different ratios.

3.2 Impact of the topology of the transposon

In nature, Tn4430 is found with both 38-bp ends in an inverted orientation. As mentioned in Section 1.4.4, there are some transposon families in which transposition can occur regardless of the topology of the ends, and others for which proper orientation of the transposon ends is preferred or required. Because of this, we were interested in investigating if having only one transposon end or having them in a direct orientation had an impact on the transposase activity.

Plasmids carrying a single copy of one or the other Mini-Tn4430 TIR (pGILC001 and pGILC002) or none of them (pGILC003) were constructed from the pGIMF001 plasmid by Loïc Codemo during his Master's student [83]. Each single-TIR plasmid was incubated with the TnpA^{3X}. As a negative control to rule out any effect due to non-specific TnpA activity, we also incubated the TIR-less plasmid pGILC003 with the transposase (Figures 3.5 and 3.6, lanes 14 – 16).

Figures 3.5(a, c) and 3.6(a, c) show the cleavage reaction performed with the supercoiled form of the single-TIR plasmids pGILC001 and pGLC002, respectively. In both cases, the kinetic analysis shows that the reaction with TnpA^{3X} leads to the accumulation of the linear fragment at the expense of the OC form of the substrate. This is not observed in the reaction where the transposase is incubated with the TIR-less plasmid (Figure 3.5(a) and 3.6(a), lane 14) indicating that accumulation of linear products observed with both single-TIR plasmid results from TnpA-mediated processing of Tn4430 TIR. Similar results can be observed in the cleavage reaction of the OC form of pGILC001 and pGILC002 plasmid with the TnpA^{3X} (Figures 3.5(b) and 3.6(b), respectively). Taken together these results suggest that TnpA^{3X} can specifically cleave at either one or the other TIR of the substrate in the absence of a second transposon end on the same plasmid. Interestingly, the rate of cleavage at the left TIR (pGILC002) seems to be higher than the one at the right TIR (pGILC001), in both supercoiled and OC forms. This is somewhat unexpected because the ends are identical

In addition to the expected cleavage products, interesting secondary bands appeared during the reactions performed with both conformations of pGILC001 and only with the OC form of pGILC002 (indicated by red bars in Figures 3.5(a – b), and 3.6(b), respectively). These

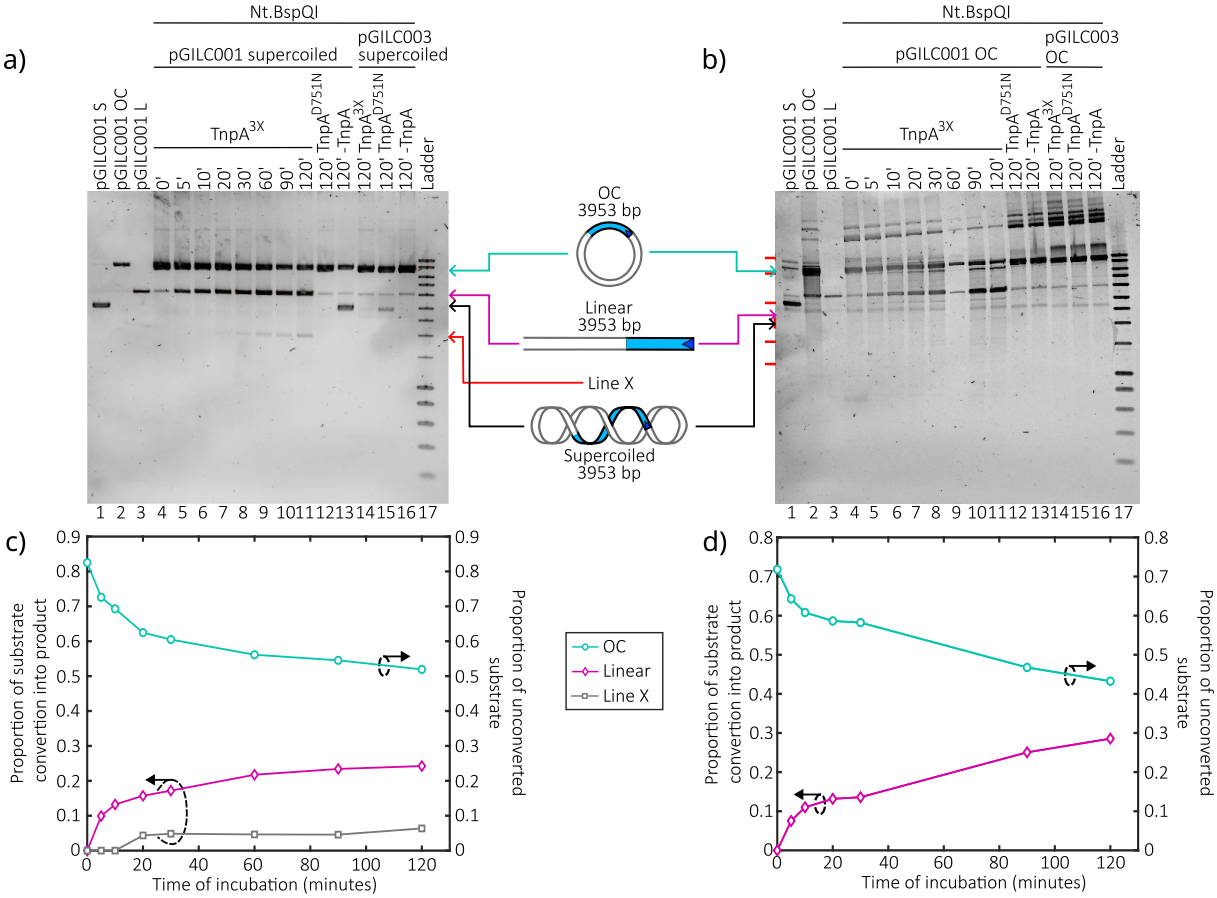

Figure 3.5: Comparison of the cleavage reaction between the TnpA^{3X} mutant (300 nM) and the supercoiled pGILC001 plasmid (6.49 nM) (a) or the relaxed open circular (OC) pGILC001 plasmid (6.49 nM) (b). Lanes ▪ 1 – 3 correspond to the three different forms of our plasmid: supercoiled (S), OC, and linear (L). The last two were obtained by digesting S with Nt.BsmAI and NcoI, respectively. Lanes 4 – 11 show the products of the cleavage reaction at different incubation times. Lanes 12 and 16 are the reactions performed with the TnpA^{D751N} catalytic mutant as a negative control. Reactions were done at 34°C. 1.2% agarose gel and SYBR Gold staining. c) and d) Quantification of the gels a) and b), respectively, performed using Image J (cf. Section6.3.4). The cleavage products (linear) use the left Y-axis, and the OC substrate uses the right Y-axis.

bands cannot be observed in the control experiments performed in the absence of active TnpA (lanes 12 and 13) or with the TIR-less plasmid pGILC003 (lanes 14 – 16), indicating that they require the specific activity of the TnpA. Since they tent to accumulate during the reaction (possibly with some delay with respect to the primary cleavage products), these bands likely correspond to strand transfer products arising from further processing reactions occurring at the TIR end. However, the pattern of the bands is different for both plasmids, indicating that the specificity is not the same. Finally, these results suggest that the presence of only one TIR on the substrate is enough to trigger specific cleavage by TnpA, and the presence of both ends in the same DNA molecule does not appear to have an impact. All of these different bands need to be further and carefully analyzed and characterized.

To broaden this question, a cleavage assay with a plasmid carrying directly repeated TIR would be necessary. A cloning strategy was set up to construct such substrate, but because of

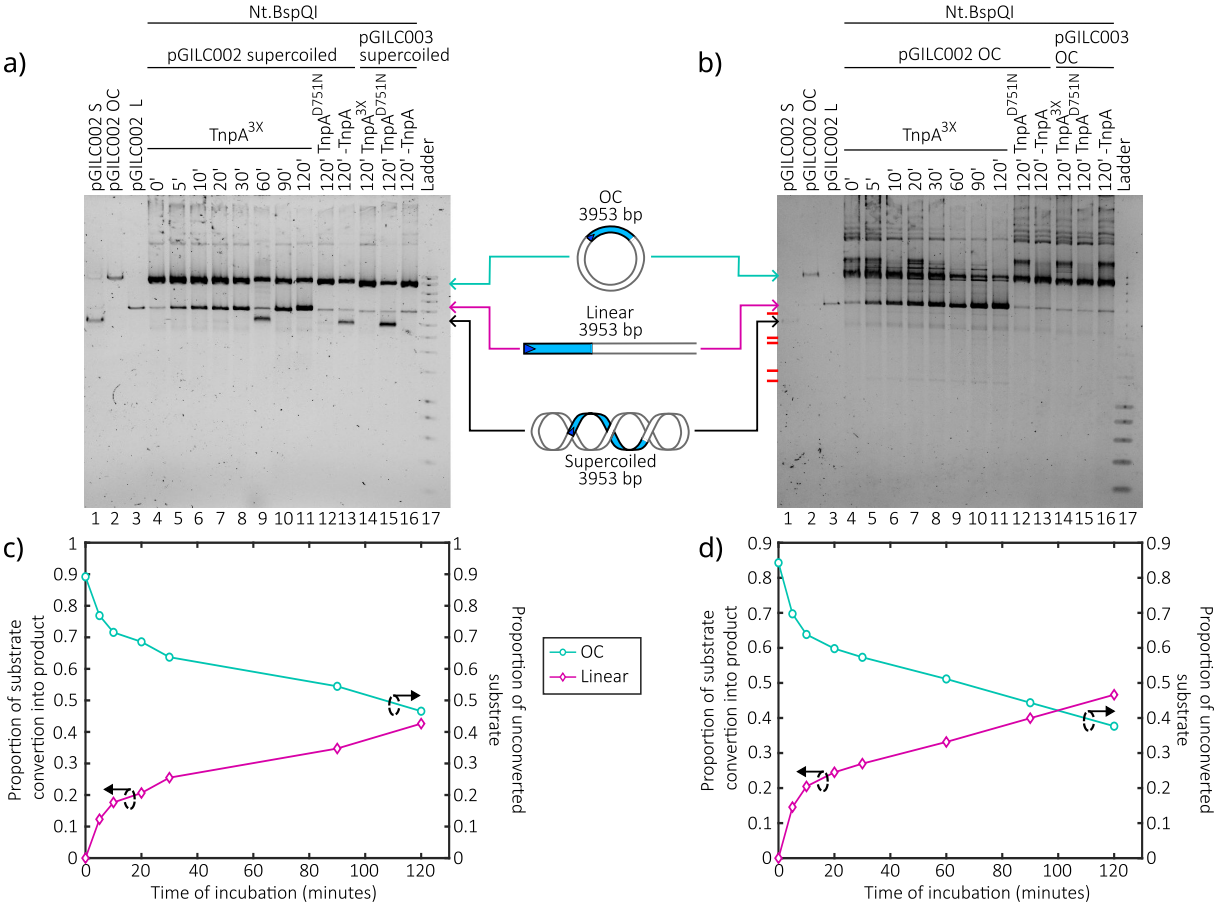

Figure 3.6: Comparison of the cleavage reaction between the TnpA^{3X} mutant (300 nM) and the supercoiled pGILC002 plasmid (6.49 nM) (a) or the relaxed open circular (OC) pGILC002 plasmid (6.49 nM) (b). Lanes 1 – 3 correspond to the three different forms of our plasmid: supercoiled (S), OC, and linear (L). The last two were obtained by digesting S with Nt.BsmAI and NcoI, respectively. Lanes 4 – 11 show the products of the cleavage reaction at different incubation times. Lanes 12 and 13 are the negative controls performed with the catalytic mutant TnpA^{D751N} and without transposase, respectively. Reactions were done at 34°C. 1.2% agarose gel and SYBR Gold staining. c) and d) Quantification of the gels a) and b), respectively, performed using Image J (cf. Section 6.3.4). The cleavage product (linear) uses the left Y-axis, and the OC substrate uses the right Y-axis.

the lack of time, it was not possible to bring it to completion (cf. Section A.2). The use of this substrate makes this part of the perspectives of this project.

3.3 Impact of the topology of donor DNA molecules

3.3.1 Circular substrates

As discussed in the introduction (cf. Section 1.4.4), in some transposons families, DNA supercoiling appears to be an important factor in the transposition mechanism because it favors the synapsis of properly oriented ends. The activity of the transposase depending on supercoiling was already started to be studied in a previous Master's thesis project [83]. The results indicated a preference for relaxed DNA, and negative supercoiling was shown to lower

TnpA activity. To further study the DNA cleavage activity catalyzed by the TnpA and determine if the results were reproducible, the experiments were performed with more accurate kinetics by increasing the time points and using shorter time laps between them, as well as the use of lower TnpA concentrations in comparison to previous experiments (80 nM) [83]. TnpA-mediated DNA cleavage was compared for two different plasmids: the supercoiled and relaxed forms of pGIMF001 plasmid. The relaxed open-circular (OC) form of pGIMF001 was generated by cleaving the plasmid with the single-strand DNA (ssDNA) endonuclease Nt.BsmAI. As illustrated in Figure A.1, this enzyme introduces four different ssDNA breaks on the DNA which relax the plasmid by removing the negative supercoils.

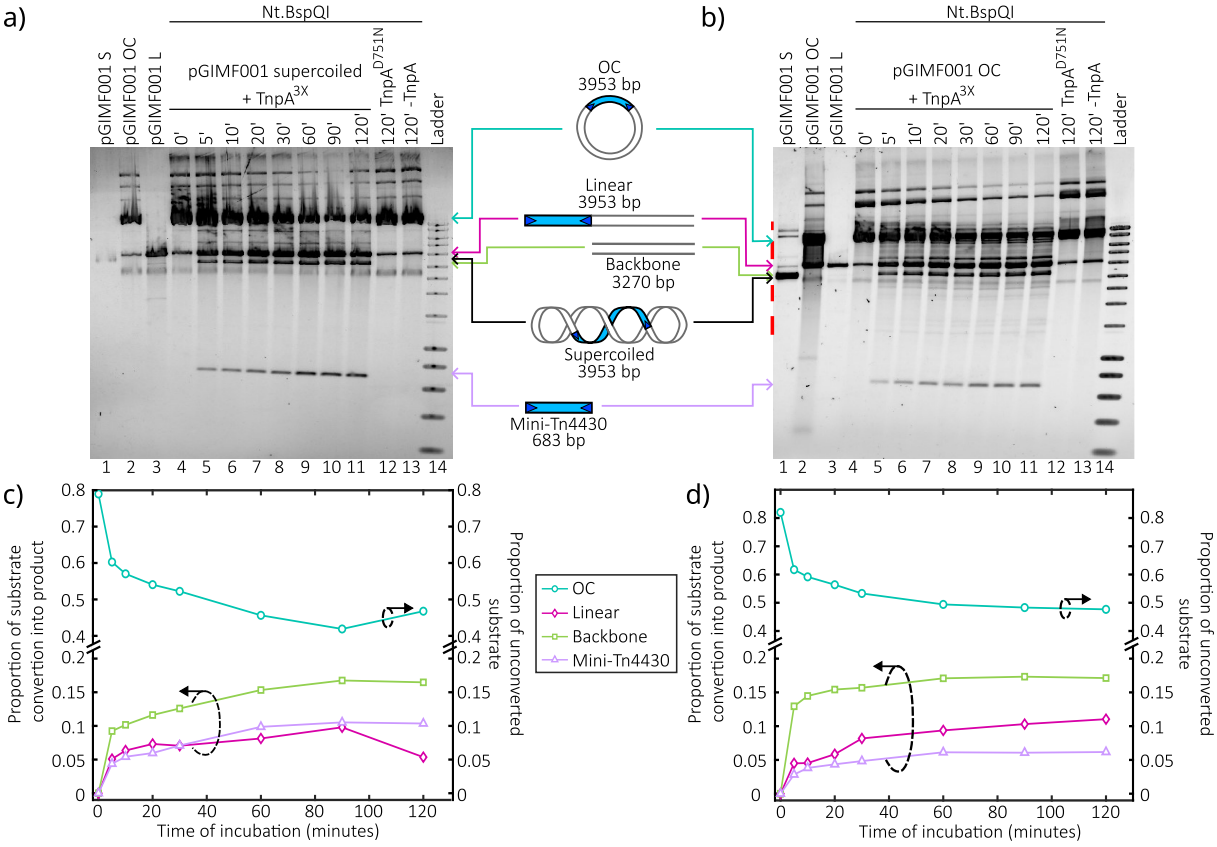

Figure 3.7: Comparison of the cleavage reaction between the TnpA^{3X} mutant (80 nM) and the supercoiled pGIMF001 plasmid (6.49 nM) (a) or the relaxed open circular (OC) pGIMF001 plasmid (6.49 nM) (b). Lanes 1 – 3 correspond to the three different forms of our plasmid: supercoiled (S), OC, and linear (L). The last two were obtained by digesting S with Nt.BsmAI and NcoI, respectively. Lanes 4 – 14 show the products of the cleavage reaction at different incubation times. Lanes 12 and 13 are the negative controls. The reaction was done at 34°C. 1.2% agarose gel and SYBR Gold staining. c) and d) Quantification of the gels a) and b), respectively, performed using Image J (cf. Section 6.3.4). The cleavage products (linear, backbone, and Mini-Tn4430) use the left Y-axis, and the OC substrate uses the right Y-axis.

TnpA^{3X} was first incubated with the supercoiled (Figure 3.7(a – c)) and relaxed forms of pGIMF001 (Figure 3.7 (b – d)). This was done by analyzing short-time kinetics from zero to two hours of the accumulation of the cleavage products. As detailed above (cf. Section 3.1), products were analyzed by digesting the reaction samples with Nt.BspQI, which introduces ssDNA breaks at the opposite side of the TnpA cleavage site, converting singly nicked products

into double-strand DNA breaks that are easily recognizable by gel electrophoresis (Figure 3.7). The results show the accumulation of the three expected TnpA cleavage products: the linear plasmid (3953 bp) arising from the cleavage of either the right or left Tn4430 end, and the backbone (3270 bp) and the Mini-Tn4430 (683 bp), both resulting from the cleavage of both ends.

By comparing the levels of cleavage obtained with the supercoiled and relaxed OC forms of the plasmid (Figure 3.7, compare panels a & c with panels b & d), we cannot confirm that DNA supercoiling has an impact on Tn4430 TnpA cleavage activity because the reaction occurred with similar kinetics with both forms of the plasmid. At each time point, a similar proportion of the initial substrate was converted into single- and double-end cleavage products. These findings contradict previous observations obtained with longer incubation times and lower resolution of the kinetics analysis suggesting that negative supercoiling of DNA substrate reduced TnpA activity [83].

Further and careful analysis of the gels revealed that the reaction with TnpA^{3X} and the OC form of pGIMF001 gave additional bands that do not appear with the supercoiled substrate. Although these results need to be further analyzed to determine the exact nature of these bands, at least some of them likely correspond to post-cleavage processing of the transposon ends such as strand transfer products.

Similar results were obtained for the incubation of the supercoiled and OC forms of pGIMF001 with another hyperactive mutant of TnpA, TnpA^{S911R} (cf. Section A.1, Figure A.3). As for TnpA^{3X}, we could not observe an impact of the negatively supercoiled DNA on TnpA^{S911R} activity. But, a comparison of the reactions performed with both transposases reveals that TnpA^{3X} has a higher activity than TnpA^{S911R} (compare panels a & c in Figures 3.7 and A.3, and panels b & d in Figures 3.7 and A.3).

Interestingly, by comparing the topology of the DNA in the single-TIR substrates (Figures 3.5 and 3.6), we can see that both, for the left and for the right end, a marginally higher cleavage activity could be observed with the OC substrate. However, as this preference is not observed in the substrate containing both ends, we cannot state that there is a clear general preference.

In addition, by comparing the single-end substrates with the substrate containing both ends (Figures 3.5, 3.6, and 3.7), in supercoiled and OC form, it seems that the presence of one or both ends does not drastically affect the activity of TnpA. Indeed for the double-TIR substrate, the total amount of products (linear, backbone, and Mini-Tn4430) and the remaining of the substrate are actually between the amount of linear product and remaining substrate generated with the right and left single-TIR substrates.

Finally, further and careful analysis of the gels revealed that the reaction with the TnpA^{3X} mutant and the OC form gave additional bands that do not appear with the supercoiled substrate. Although these results need to be further analyzed to determine the exact nature of these bands, at least some of them likely correspond to post-cleavage processing of the

transposon ends such as strand transfer products.

3.3.2 Linear substrate

To further study the impact of the topology of donor DNA molecules, the TnpA^{3X} was incubated with the linear form of the pGIMF001 plasmid. This was obtained by a polymerase chain reaction (PCR) with the specific primers (cf. Table 6.2) allowing us to amplify a 2339-bp fragment containing the Mini-Tn4430 (Figure A.2).

In this experiment, it is important to note that the outcome and the readout of the cleavage reaction are different compared to the circular substrates (Figure 3.8). In this case, the production of the Mini-Tn4430 with the left flanking region is diagnostic for cleavage at the right TIR, while the production of the Mini-Tn4430 with the right flanking region results from TnpA cleavage at the left TIR. Cleavage at both TIRs releases the fully excised Mini-Tn4430. Thus, the experiment can be used to quantify differential cleavage at the left or right TIR, and compare the kinetics of double-end cleavages compared to single-end cleavages.

The results, in Figure 3.9(a), show two important things: i) the kinetics of the cleavage reaction between the linear PCR substrate and the TnpA^{3X} mutant (lanes 3 – 10); and ii) the reaction of the linear PCR substrate with two restriction enzymes, BtsI-v2 and BsrDI, to simulate the cleavage at the left or the right TIR, respectively, and at both (lanes 12 – 14; cf. Figure 3.8 and A.2). With this result, we can observe the decrease of the linear substrate and the usual double-cleavage product, the Mini-Tn4430 (683 bp, Line 8), and it allows us to confirm which bands arise from the single- and double-cleavages from others that could be

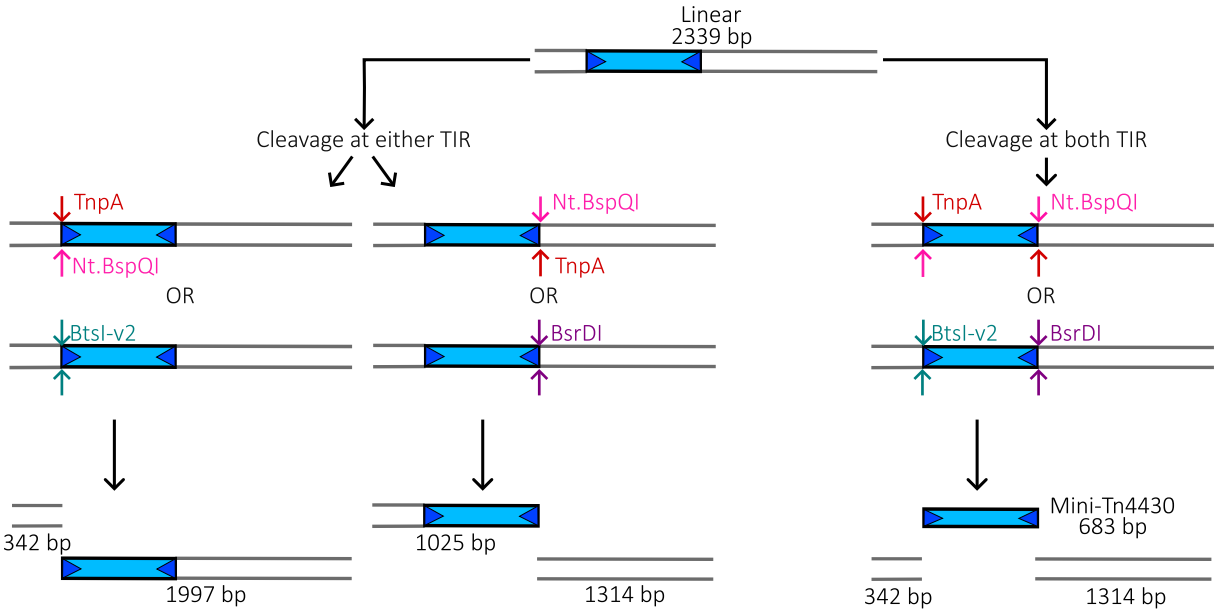

Figure 3.8: Scheme of the reaction between the linear PCR pGIMF001 substrate and the transposase and the expected single- and double-cleavage products. TnpA and Nt.BspQI do single-strand DNA breaks while BsrDI and BtsI-v2 do double-strand DNA breaks.

secondary products. Indeed, in the reaction with BsrDI, we obtained the Mini-Tn4430::small flank fragment (1025 bp, Line 6) and the large flanking fragment (1314 bp, Line 4). On the other hand, when BtsI-v2 is used, we obtained the small flanking fragment (342 bp, Line 9) and the Mini-Tn4430::large flank fragment (1997 bp, Line 2). Finally, with both enzymes, with the Mini-Tn4430, we obtained also the small and the large flank fragments (Figure 3.8).

By comparing the kinetics between the single- and double-cleavage (Figure 3.9(b)), there seems to be a preference for the cleavage site at the right TIR because the proportion of the Mini-Tn4430::small flank fragment (Line 6) increases at a higher rate than the Mini-Tn4430::long

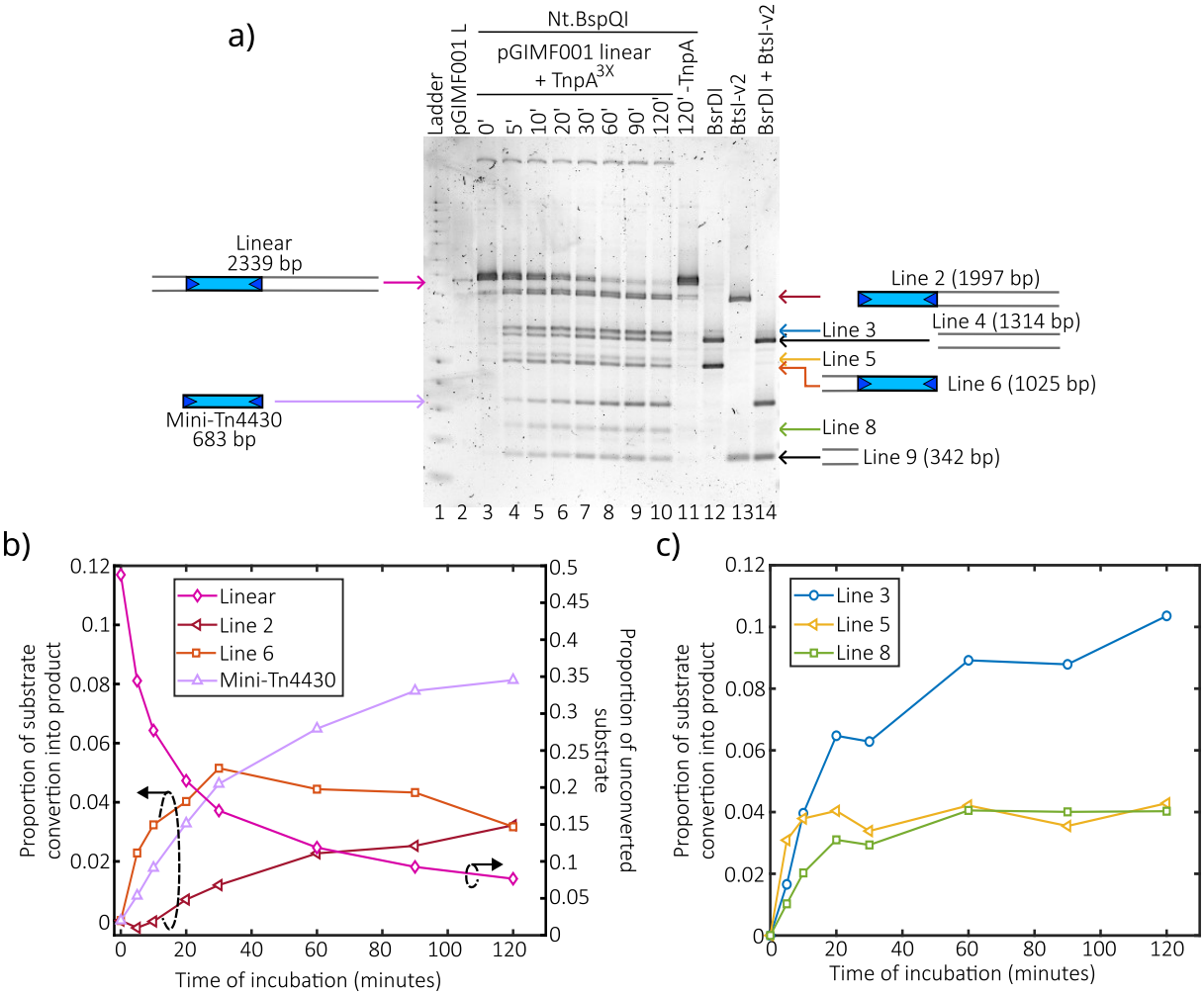

Figure 3.9: Cleavage reaction between the TnpA^{3X} mutant (300 nM) and the PCR fragment (L) of the pGIMF001 plasmid (6.49 nM). Lane 2 corresponds to the linear plasmid without any treatment. Lanes 3 – 10 show the products of the cleavage reaction at different incubation times. Lane 11 is the negative control. Lanes 12 – 14 are reactions with restriction enzymes that do double-strand DNA cuts either at the left (BtsI-v2) or the right (BsrDI) transposon end. The reaction was done at 34°C. 1.2% agarose gel and SYBR Gold staining. b) Quantification of the evolution of the proportion of the linear substrate and the single-cleavage (Line 2 and Line 6) and double-cleavage (Mini-Tn4430) products. The cleavage products (Line 2, Line 6, and Mini-Tn4430) use the left Y-axis, and the linear substrate uses the right Y-axis. c) Quantification of the evolution of the proportion of the additional secondary products (Lines 3, 5, and 8). Quantification was performed using Image J (cf. Section 6.3.4).

flank fragment (Line 2). Interestingly, after 30 minutes of fast accumulation, the right-end cleavage product starts decreasing whereas the left cleavage product keeps steadily increasing until the end. This means that the cleavage at the right TIR is faster and more efficient than at the left end. A possible explanation is that the transposase has a slight preference for this end, causing substrates that are first cut at this end to accumulate more at the beginning because there still is plenty of initial substrate available. As this availability decreases, the accumulation slows down. Concerning the slower accumulation of the left-end cleavage products, there could be two possible explanations: i) the cleavage on the left end is slow and thus the accumulation is slower, or ii) a cut on the left end is quickly followed by a cut on the right.

Although this assay does not allow us to have more insight into the topology and origin of the other products observed in Figure 3.9(a) (Lines 3, 5, and 8), we can say that they are secondary products of TnpA activity that tend to accumulate during the cleavage reactions. Indeed, in Figure 3.9(c) we can observe that their proportion increases with time with a big increase, especially at the beginning of the reaction. However, after 20 minutes or even less, the product of Line 3 accumulates more than the other two whose plateau is almost at the same height. It is also important to note that Line 5 and 8 products increase rapidly and then plateau in a similar manner to the right-end cleavage product which could mean that they are by-products of the right-end processing reaction, for example, strand transfer of the right end.

Finally, if we compare the TnpA^{3X} activity between the circular and the linear substrates (Figure 3.7 and 3.9, respectively), we can say that the linear substrate seems to increase the TnpA activity because the proportion of the substrate converted into products is higher, since, after two hours, almost all the substrate was consumed (Figure 3.9(b)).

3.4 Strand transfer products

In some experiments shown above, bands that do not correspond to the expected cleavage products were present (Figures 3.5(a – b), 3.6(b), 3.7(b), and 3.9(a)). These bands appeared in previous works [83] but not with the same profile. Although they do not appear with every substrate or under every conditions, they are secondary products of TnpA activity which tend to accumulate during the cleavage reactions. Because of this, it was interesting to investigate the nature and origin of these products in further detail. We decided to do this analysis on the reactions performed with the linear Mini-Tn4430 substrate because they gave reproducible, quantitative, and discrete extra bands resulting from secondary processing reactions (Figure 3.9).

To this end, we repeated a two-hour incubation of the linear substrate with the TnpA^{3X} mutant and analyzed the products by treating the reaction with Nt.BspQI (Figure 3.10, lane 3). As above, the unreacted substrate was cut with BtsI-v2 and BsrDI to identify the cleavage products at one or the other Mini-Tn4430 end, or both (Figure 3.10, lanes 4 – 6). The result confirmed that the substrate was efficiently cleaved by TnpA^{3X} giving the expected single- and double-end cleavage products as previously observed (Figure 3.9(a)). However, the reaction

generated proportionally lower TnpA-dependent secondary products corresponding to very faint bands at the expected position within the gel (compare Figure 3.10, lane 3 with Figure 3.9(a), lane 10). Intriguingly, additional slow-migrating bands were observed at an upper position of the gel, which were not found in previous reactions.

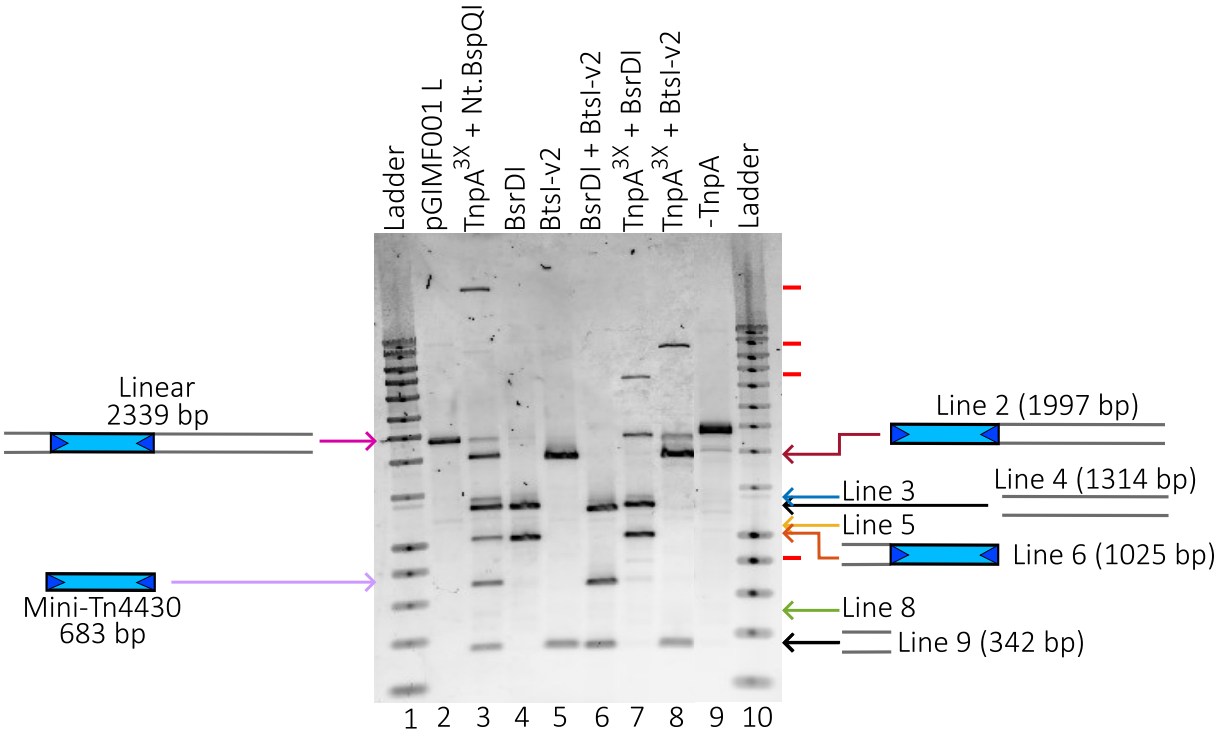

Figure 3.10: Reactions with the Mini-Tn4430 linear substrate (6.49 nM). Lane 2 corresponds to the PCR fragment used as the substrate. Lanes 3 is the 2-hour cleavage reaction with the TnpA^{3X} mutant (300 nM), followed by Nt.BspQI digestion and incubated at 34°C for 2 hours. Lanes 4 – 6 are reactions with restriction enzymes that do double-strand DNA cuts at the left (BtsI-v2), at the right (BsrDI), or at both transposon ends. Lanes 7 – 10 correspond to cleavage reaction with TnpA^{3X} (300 nM) followed by reactions with different restriction enzymes (BtsI-v2, BsrDI, Nt.BspQI). Lane 11 is the cleavage reaction without transposase, the negative control. 1.2% agarose gel and SYBR Gold staining.

In an attempt to gain some insight into the identity of these extra bands, the reaction was digested with BsrDI or BtsI-v2 alone (Figure 3.10, lanes 7 and 8). Since the BsrDI and BtsI-v2 restriction sites were engineered in pGIMF001 so as to cleave at the same position as TnpA (Figure A.2), we reasoned that strand transfer reactions involving the 3' end of the Mini-Tn4430 would perturb cleavage by these restriction enzymes due to the formation of branched Y-shaped DNA structures at the junction between the TIR sequence, the donor flanking DNA and the new target locus (Figure 3.11). As a consequence, we expected that specific bands arising from strand transfer reactions would be refractory to BsrDI, to BtsI-v2, or to both cleavages depending on which Mini-Tn4430 ends have been used in the transfer (i.e. single-end transfer of the right or left TIR would be resistant to BsrDI or BtsI-v2, respectively, while the concerted transfer of both ends would be resistant to digestion by both enzymes (Figure 3.11).

Treatment of the reaction by either BsrDI or BtsI-v2 showed a pattern of bands similar to the one obtained for the digestion of the unreacted substrate with the same enzymes (Figure 3.10,

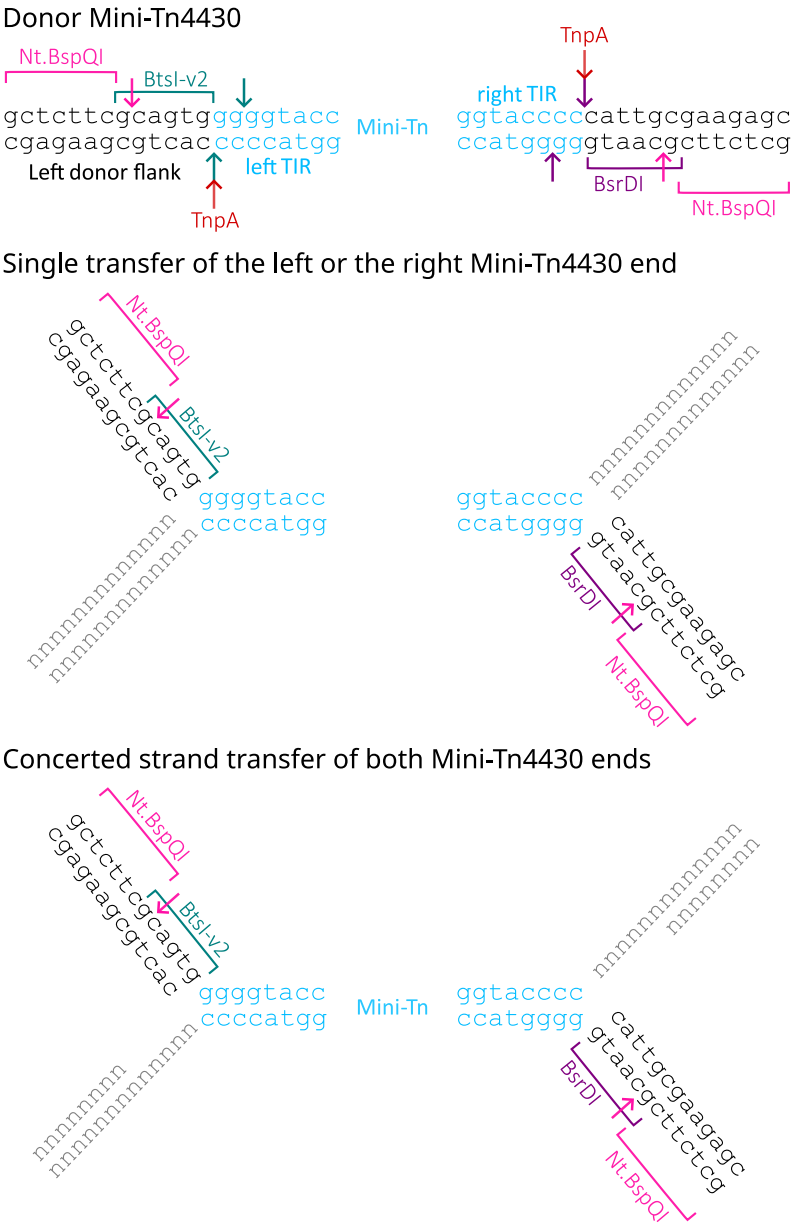

Figure 3.11: External flanking DNA sequences of Mini-Tn4430 and position of the restriction sites of Nt.BspQI, BsrDI, and BtsI-v2. Left and right end insertions create a branched substrate at the restriction site of BtsI-v2 and BsrDI respectively, possibly protecting the ends of the transposon from restriction. On the other end, Nt.BspQI restriction should be unperturbed as it takes place further away from the branching.

compare lane 4 with lane 7, and lane 5 with lane 8). This indicates that simple 3' end ssDNA cleavage by TnpA does not affect the activity of BsrDI and BtsI-v2. However, both enzymes gave an additional band migrating at the position of the full-length linear substrate (around 2.3 kb), as could be expected for intramolecular strand transfer products involving the right TIR (resistant to BsrDI) or the left TIR (resistant to BtsI-v2) or both (resistant to both BtsI-v2 and BsrDI) (Figure 3.10, lanes 7 and 8). Despite different structures, the restriction-resistant products are expected to have the same molecular weight explaining their colocalization within the gel (Figure 3.12). However, transfer products involving the right end alone would be cut by BsrDI,

while those involving the left end would be cut by BtsI. Depending on the position of the target site with respect to the Mini-Tn4430 ends, some of these products would lose the large (1314 bp) or small (342 bp) flanking DNA segments of the substrate, which may explain the presence of bands (or a smear) migrating between the Mini-Tn4430 (683 bp) and Mini-Tn4430::left flank (1025 bp) fragment (Figure 3.10, lane 7), and between the substrate (2339 bp) and the Mini-Tn4430::right flank (1997 bp) fragments (Figure 3.10, lane 8), respectively.

The nature of the low-mobility DNA species migrating above the initial substrate remains unclear. One possibility is that they correspond to high molecular weight products arising from intermolecular strand transfer reactions between separate substrates. Consistent with this interpretation, digestion of the reaction with BsrDI and BtsI-v2 converted the putative intermolecular products into smaller DNA molecules (Figure 3.10, compare lane 3 with lanes 7 and 8). Although the mobility of the resulting bands does not strictly correlate with the expected size difference, the pattern is consistent with a larger fragment being deleted after BsrDI digestion than after BtsI-v2 treatment. Again, the discrepancy between the observed relative mobility and the expected size difference might result from the structural features of the DNA products (Figure 3.12). Although the results proved to be difficult to interpret due to the poor yield of secondary products and the complexity of the different outcomes of the reaction, this analysis supports the possibility that TnpA^{3X} can catalyze strand transfer reactions on Mini-Tn4430 containing linear DNA substrates.

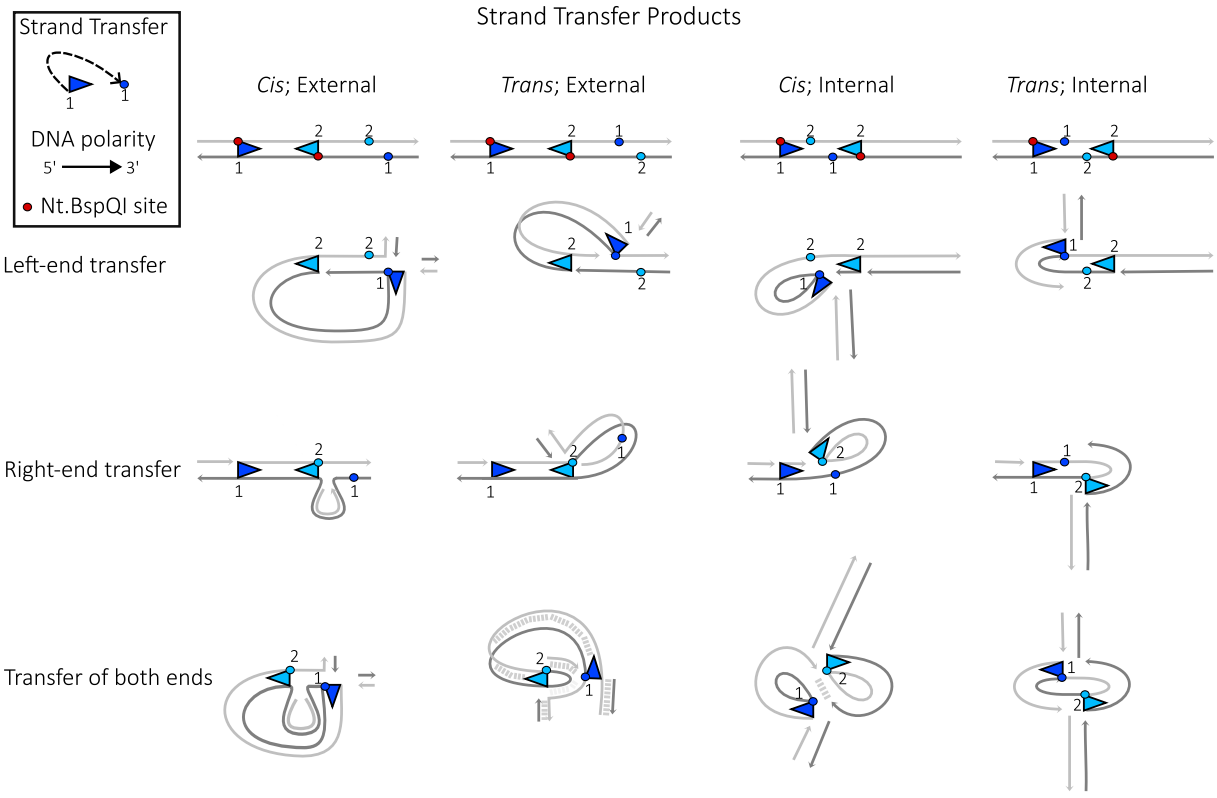

Figure 3.12: Possible intramolecular strand transfer products. The insertion of either end of the transposon or both simultaneously was modeled. For double insertions, the left and right end insertions were represented at a distance of five base pairs (bp) because, *in vivo*, Tn4430 creates 5 bp direct repeats. In addition, we represented the cases when the attack occurs either outside or inside the Mini-Tn4430. For outside transfer, we only represented transfer on the right side of the Mini-Tn4430, because transfer on the left side essentially gives the same product inverted. Finally, we imagined transfers on the same strand of DNA (*cis*) or on the opposite strands (*trans*). Light and dark blue triangles represent the transposon end. Red dots correspond to the cleavage site of Nt.BspQI. Light and dark blue circles show the possible insertion sites of the transposon.

4 | Discussion

The Tn3 family of transposons greatly impacts human health because of its implication in the dissemination of antimicrobial resistances, including against last-resort antibiotics like carbapenem and colistin. This involvement is, among other things, due to the broad expansion of this family across almost all the bacterial phyla [91].

Despite this biological importance, many aspects of the Tn3-family transposition mechanism remain still understood. Despite recent advances in the establishment of the structure of the transposase TnpA and the molecular architecture of key TnpA-DNA intermediates, the exact pathway leading to the transposition complex assembly and the mechanisms that control TnpA activity have yet to be elucidated.

During this Master's thesis, we were first interested in further optimizing the reaction conditions used in this system. TnpA-mediated DNA cleavage was examined on DNA molecules containing a derivative of Tn4430 that prefigures its natural configuration. The experiments were mainly performed using a previously characterized hyperactive mutant of TnpA, TnpA^{3X} (sometimes comparing the results with those obtained for another hyperactive variant, TnpA^{S911R}) in order to set up conditions for the study of the wild-type transposase.

Recent biochemical and structural studies revealed that both the TnpA^{WT} and hyperactive mutants of TnpA form stable dimers in solution [81, 92]. In our current models for the assembly of the transposition complex, a TnpA dimer first binds on a single transposon end to form the inactive single-end complex (SEC) and then undergoes a conformational change allowing it to bind the second end to form a paired-end complex (PEC) in which TnpA is in an active configuration. This model of synapsis-by-naked-end-capture (S-NEC) suggests that Tn4430 transposition might be controlled by an "assembly site occlusion" (ASO) mechanism in which stable binding of a TnpA dimer on both transposon ends would preclude the assembly of the activated PEC. To test this hypothesis DNA cleavage reactions were performed at different ratios between TnpA^{3X} and the DNA substrate. These assays (cf. Section 3.1.2) show the importance of testing the best concentrations to use for the substrate and the protein when the cleavage activity of the protein is to be tested. Even if these experiments need to be done again, we could already say that at the concentrations used in most of our experiments (cf. Section 3.1.1, 3.2, 3.3.2, and 3.4), where 300 nM of transposase is incubated with 6.49

nM of the substrate, there seems to be an ASO mechanism occurring (Figure 3.3). This is because we have an excess of transposase in comparison to the transposon ends.

DNA supercoiling, for some transposon families, is an important factor in the transposition mechanism because it allows a proper synapsis of the transposon ends. In the case of Tn4430, no obvious difference in the TnpA^{3X} activity was observed when comparing the supercoiled plasmid with the OC or the linear PCR form (Figures 3.7 and 3.9). These results are interesting and intriguing because they do not confirm previous results obtained with the Tn4430 [83] or in other transposon families [64]. There are two different possibilities. First, if it is proven that in Tn4430, TnpA has higher activity with relaxed substrates than with negatively supercoiled substrates. This could be explained by the fact that during transposition the final architecture of the PEC imposes a positive cross on the DNA substrate between the two transposon ends [81], therefore on a negatively supercoiled substrate, the negative cross has to be changed to a positive one, making it energetically unfavorable. This is in contrast with the relaxed form, where the formation of a positive cross does not need energy. Thus, the system could have limits in terms of energy if the substrate is supercoiled. On the other hand, in the case of Hsmar1, the supercoiled substrate appears to enhance the transposase activity because it brings both transposon ends together into a configuration that makes easier the assembly of the transpososome. Finally, with the experiments done during this project, we cannot affirm that a topological filter model is occurring in this transposon family.

Moreover, reactions with the single-TIR plasmids showed cleavage activity of the transposase (cf. Section 3.2). This suggests that, even with one transposon end, the Tn4430 transposase can be activated. These results are interesting because the cryo-EM PEC structure has shown that the cleavage activity of the transposase occurs in *trans*, meaning that one transposase monomer binds one transposon end and cleaves the other one. We could also say that maybe the reason why we observe activity with only one end is because the transposition complex was formed with two different plasmids, each one containing one TIR. Moreover, this activity can be confirmed because, with a TIR-less plasmid, the TnpA was not active. In addition, the results also showed a preference for the left TIR, which is very intriguing because, contrary to other transposons (including some members of the Tn3 family) [93], both Tn4430 ends are identical. This could be due to i) the transposon end orientation has an impact on the activity, which would need to be further investigated; or ii) the external sequences flanking the TIRs have an impact because maybe they are different between them. Indeed, in the cryo-EM structures, we could observe that the TnpA does non-specific interactions with these sequences, and thus there could be some interactions with the left TIR that are more favorable [81].

During some cleavage reactions, additional bands were observed that did correspond to TIR DNA cleavage products generated by TnpA (Figures 3.5(a – b), 3.6(b), 3.7(b), 3.9(a), and A.3(b)). The bands profile was different depending on the starting substrate. They were present in all the reactions performed with the relaxed OC form of the plasmids (pGIMF001, pGILC001, and pGILC002) and with the linear substrate, but they could only be observed in

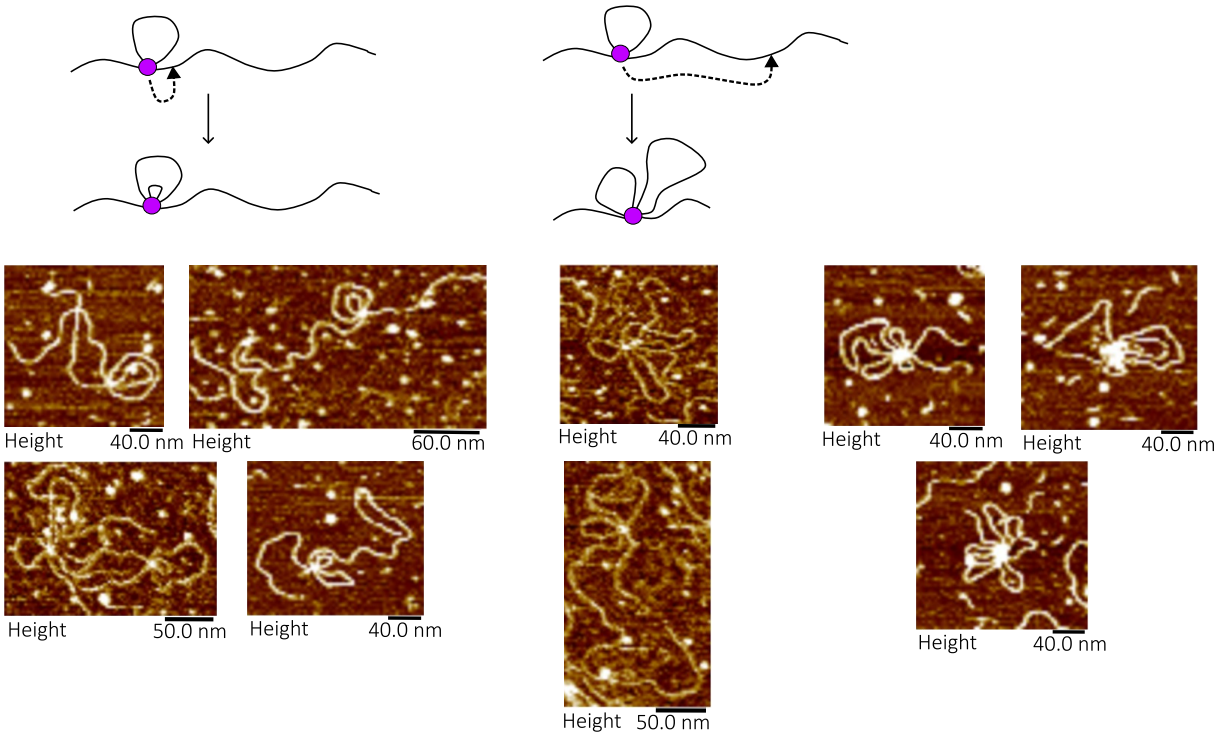

Figure 4.1: AFM imaging of products from a reaction between the linearized pGIMF001 plasmid and the TnpA^{3X} mutant. The purple circle represents the transposase. Dark lines correspond to the DNA. Dashed lines show the possible auto-insertions. (M. C. Fernandez, unpublished data).

one of the supercoiled forms, with pGILC001. The bands were further studied with reactions done with other restriction enzymes on the linearized pGIMF001 plasmid (Figure 3.10). A first attempt to characterize these extra-product bands was performed by cutting the reaction of the Mini-Tn4430 linear substrate with specific restriction enzymes (Figure 3.10). We have hypothesized that the strand transfers, either at the right, at the left, or at both ends, would be resistant to the digestion by BtsI-v2 or BsrDI, but we found that the activity of these enzymes is not affected and thus give bands that migrate close to the linear substrate (when cutting with both enzymes), or between the Mini-Tn4430 and the Mini-Tn4430::left flank fragment (if strand transfer at the right end), or between the substrate and the Mini-Tn4430::right flank fragment (if strand transfer at the left end). Furthermore, there were other low-migrating products that are still not identified but they could be products from intermolecular strand transfers.

AFM imaging offers a great opportunity to visualize the nucleoproteic complexes or cleavage and strand transfer products created by the TnpA on plasmid substrates containing the Mini-Tn4430. It allows us to get static snapshots of the reaction steps and products while the biochemical analyses give us insight into the dynamic of the reaction. Several such AFM experiments were carried out with the linear PCR substrate (M. C. Fernandez, unpublished data) and in previous works [94, 95], which gave us an indication of the different possibilities for the topology of these additional bands (Figure 4.1). Therefore, we decided to represent the auto-insertions that can occur intramolecularly because they seem to correspond more closely to

the AFM imaging data (Figure 3.12). As the number of possible insertion scenarios is vast, we selected a few specific cases of intramolecular auto-insertions based on our data and assumptions. However, we do not discard the possibility of having auto-insertions intermolecularly as they could explain some of the products in Figure 3.10.

As mentioned in the results, some unknown products (Figure 3.9) migrate close to known cleavage products (Line 3 and Line 5). This can happen for two segments of the same size when one is relaxed and the other one contains a loop. As can be seen in our representation of hypothetical strand transfer products (Figure 3.12), such loop could be easily created by intramolecular strand transfer events. Additionally, such events cause segmentation of the linear plasmid that could explain the rest of the unknown bands.

Although intramolecular/intermolecular strand transfer products should be inhibited because of the transposition immunity of Tn4430, they are plausible because our experiments were carried out with the deregulated hyperactive (T+/I–) TnpA^{3X} mutant which has a decreased immunity phenotype.

In order to determine whether these products are the results of auto-insertion events to define their nature more precisely we could use sequencing. Design of the primers and interpretations of the results might be complicated due to the many different possibilities of topologies but, with the designs, in Figure 3.12, we can start thinking about the specific primers that would help us find the best possible results.

5 | Conclusion and Perspectives

This Master's thesis has allowed us to have more knowledge and understanding of the transposition mechanism of the Tn4430. It has given us more information about the best temperature for the *in vitro* transposition reaction, as well as new insight into a possible role of overproduction inhibition during the transposition complex assembly. It will be very interesting and appropriate to pursue these experiments in order to assess what is the best ratio of protein and substrate concentration. This will allow us to better analyze and evaluate the cleavage activity of the Tn4430 transposase. Moreover, to further optimize the biochemical assay, it will be necessary to set up the appropriate pH for the reaction.

This project did not confirm the topological filter mechanism hypothesized in a previous study because we did not observe a difference in TnpA activity between a negatively supercoiled and a relaxed substrate. In the future, further experiments to test the activity will be necessary, and maybe try to understand which parameter(s) is(are) the one giving these different results. We have hypothesized that it could potentially be due to the starting plasmid. In particular, the plasmid used in previous experiments had a complex topology with many topoisomers which was not the case here.

To conclude the study on the impact of the donor DNA topology, it would be interesting to perform the same reaction but this time with a positively supercoiled plasmid as a substrate. This could be done with the use of a reversed gyrase, a topoisomerase that generates positively supercoiled DNA in thermophilic archaea.

Furthermore, to go deeper into the impact of the transposon topology on TnpA activity, the cleavage reaction should be done with a plasmid containing direct repeats instead of inverted repeats and assess if a miss orientation of the TIR allows the transposase to form the transpososome or not.

Finally, some experiments gave additional bands that are likely to be strand transfer products. The study of their topology and origin has already started but needs to be taken further. A sequencing approach could help to achieve this but it would require having the proper primers and to determine the best sequencing method. Another approach that should be pursued would be AFM because it has already given us insight into the double loops with the linear substrate, so it could also give us more information about other types of topologies.

6 | Materials and Methods

Table 6.1: Solutions

Solution	Composition
Buffer A (TP A)	50 mM Tris pH8, 1 M NaCl, 10% glycerol, 20 mM imidazole
Buffer B (TP B)	50 mM Tris pH8, 1 M NaCl, 10% glycerol, 500 mM imidazole
Cleavage buffer	5 M NaCl, HEPES 10X pH8, 1 M $MnCl_2$, 0.1 M DTT, H_2Odd
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	500 mM HEPES, 2 M NaCl, 1 M L-arabinose
LB liquid culture medium	5 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone
LB liquid culture medium	5 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone, 20 g/L agar
Lysis buffer	50 mM Tris pH8, 1 M NaCl, 10% glycerol
Orange G	0.25% bromophenol, 0.25% xylene cyanol, 30% glycerol
STE Buffer 10X	100 mM Tris-HCl, 1 M NaCl, 10 mM EDTA, pH8
STOP buffer	20 mg/ml proteinase K, 10% SDS, 0.5 EDTA, H_2Odd
SYBER-Gold buffer	10 μl for 100 ml of TAE buffer
TAE (Tris-acetate-EDTA) buffer	88 mM Tris base, 88 mM orthoboric acid, 2 mM EDTA pH8
Wash EDTA	20 mM sodium phosphate pH7.4, 0.5 M NaCl, 0.05 M EDTA

Table 6.2: PCR primers

PCR primer	Sequence (5' to 3' orientation)	Use
Fw3-MCF	TGACCATGATTACGCCAAGC	To Amplify linear substrate (2239 bp) for cleavage reaction from pGIMF001
Rv3-MCF	AAGTTGGCCGCAGTGTTATC	To Amplify linear substrate (2239 bp) for cleavage reaction from pGIMF001

6.1 Substrates

Table 6.3: Substrates

Plasmid	Description
pGIMF001	Central plasmid of this master thesis
pGILC001	Derivative from pGIMF001 containing the right transposon end. Obtained by digesting pGMF001 with XbaI and NheI
pGILC002	Derivative from pGIMF001 containing the left transposon end. Obtained by digesting pGMF001 with BamHI and BglII
pGILC003	Derivative from pGIMF001 from which both transposon ends have been deleted. Obtained by digesting pGILC001 with BamHI and BglII

6.1.1 pGIMF001

Plasmid pGIMF001 is a fully customized plasmid of 3.9 kb carrying a 683-bp Mini-Tn4430 element that is delineated by properly oriented left and right terminal inverted repeats (TIR, 38 bp) of Tn4430. This plasmid harbors recognition sites of the ssDNA endonuclease Nt.BspQI on the opposite side of the DNA cleavage sites of the TnpA (Figure A.1). Since TnpA introduces ssDNA breaks at the 3' end of the transposon, digestion of the reaction with Nt.BspQI generates double-strand breaks that are easily recognizable by standard gel electrophoresis.

The plasmid's open circular form was generated using the ssDNA endonuclease Nt.BsmAI.

Finally, the plasmid's fragment form of 2339 bp was generated by a Polymerase Chain Reaction (PCR). In a PCR tube were added 33 μl of distilled water, 10 μl of Q5 Buffer, 2.5 μl of the forward and reverse primers (Table 6.2), 0.5 μl of dNTPs, and 0.5 μl of Q5 polymerase. In another Eppendorf, colonies of bacteria were taken from a fresh plate and diluted in 200 μl of distilled and sterile water. Then 1 μl of this mix was taken and mixed with the 49 μl. Each sample was put at a denaturation temperature of 98°C for 30 seconds. Then, 25 cycles of denaturation at 98°C for 10 seconds, hybridization at 72°C for 30 seconds, and polymerization at 72°C for 1 minute and 55 seconds. Finally, the elongation was at 72°C for 7 minutes.

6.1.2 Plasmid purification

The plasmids of interest were extracted from *Escherichia coli* TOP10 strain which contains an ampicillin resistance gene (AmpR). The bacteria were streaked on LB-agar plates with ampicillin and incubated overnight at 37°C. An isolated colony was selected from the plate and inoculated in a certain volume of LB liquid pre-culture with ampicillin, depending on the

necessary quantity and kit used, at 37°C overnight. For example for the minipreps, the volume was between 20 and 32 ml of LB and for the maxiprep was 100 ml.

1 ml of overnight bacteria culture was added to an Eppendorf and centrifuged at 10,000 x g for 30 seconds. The supernatant was discarded. This step was repeated 3 to 4 times in other to collect more cells. After this, the pellet was resuspended in 250 μl of Resuspension Buffer. Once the pellet was resuspended, 250 μl of Lysis Buffer were added and the Eppendorf was gently inverted 4 to 6 times. This step must be done in less than 5 minutes and it must not be vortex. Then, 350 μl of Neutralization Buffer were added and gently inverted 4 to 6 times and centrifuged for 10 minutes at $>$ 14,000 x g. After this, a SmartPure column was inserted into a collection tube, and the supernatant was added to the column. Centrifuge for 60 seconds at 6,000 x g and the flowthrough was discarded. Once it has finished, 650 μl of Wash Buffer were added to the column and centrifuged for 60 seconds at 12,000 x g. The flowthrough was discarded and the step was repeated one more time. After that, it must be centrifuged once more for 1 minute at 12,000 x g in order to remove residual liquid. Finally, the column was transferred to an Eppendorf and 50 μl of water were added. Leave it for 1 minute at room temperature and then centrifuge for 1 minute at 12,000 x g. The column was discarded and the purified plasmid DNA was stored at -20°C.

6.2 Transposase purification

6.2.1 Expression vectors

Table 6.4: Expression vectors

Vector	TnpA	Origin
pGIAD003	WT	A. Draime
pGIML3X/M-H	3X	M. Lambin
pGIML007	S911R	M. Lambin
pGIAD002	D751N/S911R	A. Draime

These vectors allow the expression of TnpA. They were expressed in *Escherichia coli* TOP10 F- strains. The expression vector consists of an L-arabinose inducible pAra promoter with a tetracycline resistance gene (TcR) and the transposase sequences inserted, with a translation fusion, upstream MycHis6 tag [12].

6.2.2 Induction

Prepare between 500 ml and 750 ml of LB in Erlenmeyer of 2 L. In general, the volume used was 600 ml.

Escherichia coli TOP10 strains containing the vector of interest were stripped onto LB-agar plates with tetracycline and were incubated at 37°C overnight. For each strain used, an isolated colony was collected and inoculated in X ml LB pre-culture multiplied by 0.03, containing tetracycline, at 37°C overnight. In the case of 600 ml, the pre-culture contained 18 ml of LB and 18 μl.

The pre-culture was added to the 600 ml LB liquid culture with 600 μl of tetracycline and 600 μl benzyl alcohol (final concentration of 10 mM). The latter was used because the transposase was not very soluble and quite unstable. After the inoculation, 1 ml of the liquid LB culture was harvested to determine the optical density (OD) at 600 nm, which should be at least 0.04. The culture was then incubated at 37°C with 150 rmp agitation. The OD was assessed regularly during this period. When the OD was between 0.5 and 0.7 (around 3h hours later), 1 ml of the culture was recovered and washed two times in 1 ml of TN and the pellet was stored at 20°C, which will serve as a pre-induction control (T0). Then, the culture was placed on ice at 4°C for ten minutes in order to reduce the temperature in a homogeneous way to 20°C. After the temperature was reduced, the induction was initiated by adding 1.2 ml of L-Arabinose (20%). The culture was incubated for two hours at 22.5°C with 120 rpm agitation. When the two hours have passed, another 600 μl of L-arabinose was added and the culture was incubated under the same conditions for two more hours. Once finished, 1 ml of the culture was recovered and washed two times in 1 ml of TN and stored at -20°C to use as a post-induction control (T4).

Then, the culture was centrifuged for 10 minutes at 4°C, 4000 rpm, and the supernatant was discarded. The pellet was then resuspended in 100 ml of lysis buffer and re-centrifuged for 10 minutes at 4°C, 4000 rpm.

Finally, the pellet was stored at -80°C in the centrifugation tube [12].

6.2.3 Lysis

The pellet was defrosted on ice. After defrosting, it was resuspended in 15 ml of Tp A (cf. Table 6.1) with cOmplete EDTA-free (protease inhibitor). One tablet of the protease inhibitor was added per 30 ml of Tp A (Table 6.1).

Cell lysis was performed by centrifugation with glass beads. In 2 ml sterile tube with a screw cap, 200 μl of glass beads were added, and then 1 ml of the solution. These tubes were centrifuged twice at 6 m/s for 25 seconds. The different fractions of the samples were placed on ice between each step.

After lysis, the samples were centrifuged for 15 min at 4°C, 13000 rpm and, the supernatant was collected in a tube and recentrifuged for 15 minutes at 4°C, 13000 rpm. The supernatant was placed in a syringe, passed through a 0.45 μm filter, and placed at 4°C for purification [12].

6.2.4 Purification

Purification was done thanks to the affinity of the N-terminal MycHis6 tag fused to the protein for a nickel column. The entire purification was done at 4°C using the AKTA Prime machine. All the solutions needed were first filtered and then degassed.

The purification starts with a washing step with water. Then, different solutions were used in order to equilibrate the column: TP A, TP B, 0.1 M NiSO4, wash EDTA (Table 6.1), and 20% ethanol. Once the column was equilibrated, 25 ml of water were inserted followed by 25 ml of TP A. After that, the filtered lysate was loaded, 4 ml per 4 ml. Once all the lysate was loaded, the program begins. It will first do a washing step with a mixture of TP A and TP B, with an increased proportion of TP B. This will generate a gradient of imidazole because both buffers contain a certain concentration of imidazole (Table 6.1). The gradient goes from 20 mM to 500 mM. This allows the discarding of the nonspecific binding between contaminating proteins and the column at low concentrations whereas, at high concentrations, there will be the elution of the recombinant TnpA that was bound by its tag. The elution of the transposase was monitored via a spectrometer which allows the detection of the elution peak corresponding to the elution of the TnpA [12].

6.2.5 SDS-Page gel

For each purification, culture samples harvested before (T0) and after 4 h (T4) of induction were compared to the elution fractions by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The T0 and T4 were suspended in 80 μl and 160 μl of water, respectively, and 5 μl of Crack 5X (Table 6.1). For each elution sample, 20 μl were collected and 5 μl of Crack 5X were added. All the samples were heated at 95°C for 10 minutes.

Then, all the samples were loaded onto a pre-cast 4-12% gel, and the gel was subjected to a current of 120 V for 1 hour. Finally, the gel was revealed by immersion in Instant Blue [12].

6.2.6 Protein reconcentration

Fractions containing the most TnpA and the least contaminants, seen thanks to the SDS-PAGE, were collected. They were reconcentrated using AMICON Ultra 10 kDa column. The column was centrifuged at 4°C, between 4,000 and 5,500 rpm for 30 to 40 minutes. Once the optimal volume was achieved, the column was refilled with lysis buffer and centrifuged again at 4°C, 5,000 rpm for 30 minutes. This step was done twice and once it was finished, the remaining liquid in the column was separated into a falcon and stored at -80°C [12].

6.2.7 Bradford assay

The concentration of transposase was determined by a Bradford assay. First, a calibration curve was done from known concentrations of Bovine Serum Albumin (BSA). The blank was

the lysis buffer because it was the one used for the reconcentration. Then, the different concentrations of BSA were 320, 160, 80, 40, 20, and 10 μg/ml.

In order to assess the concentration a plate of 96 well was used in order to measure the absorbance. The 3 first lines were the 3 replicas of the calibration curve and the fourth line was the protein sample. In each well, 40 μl of Protein Assay Dye reagent concentration, 150 μl of water, and 10 μl of either the protein sample or the determined BSA concentration were added. Finally, the absorbance values of our samples were compared to the different BSA concentrations and then the concentration can be determined [12].

6.3 Cleavage assay

The cleavage assay was a biochemical assay that allows us to study the TnpA activities on a DNA molecule carrying functional or mutant derivatives of Tn4430, "Mini-Tn4430". It was the central assay of this master thesis.

6.3.1 Cleavage test

This test was performed in one Eppendorf per condition with a final volume of 30 μl. Under standard conditions, the Eppendorf contained 6 μl of glycerol (final concentration of 22%), 6 μl of Cleavage Buffer (Table 6.1), 2 μl of purified plasmid (final concentration of 16.67 ng/μl), and the volume was completed with distilled water (around 2 μl). In another Eppendorf, the transposase was prepared. Depending on the initial concentration, a certain quantity was added and mixed with a buffer containing 30% glycerol to have 8 μl of TnpA at 300 nM as the final concentration. The transposase was then added to the Eppendorf containing the mix and incubated at 34°C for different incubation times, from 0 to 120 minutes or 24 hours, depending on the experiment, to have the kinetics of the reaction. After incubation, 30 μl of STOP Buffer were added and the tube was incubated for another 30 minutes at 37°C.

For the kinetics, the reactions were performed in the same Eppendorf tube. For example for 5 cleavage reactions, the final volume was 150 μl and after each time point, 30 μl from the Eppendorf were taken and were added to another tube to stop the reaction, while the remaining of the reaction was continued at 34°C up to the last time point.

6.3.2 DNA product purification

After the cleavage reaction, the final volume was 60 μl. To this volume were 300 μl of Binding Buffer from the kit. The 360 μl were mixed and placed in a column and centrifuged for 1 minute at 16,000 x g. The flow-through was discarded, the column was washed with 200 μl of Wash Buffer and centrifuged for another 1 minute at 16,000 x g. The washing step was repeated and finally, 6 μl were added to the column and centrifuged at 16,000 x g for 1 minute in order to elute our DNA in a new Eppendorf.

6.3.3 Digestion with restriction enzymes and agarose gel

Once the DNA was purified, to the tube containing the 6 μl of DNA were added 11 μl of water, 2 μl of Restriction Buffer, and 1 μl of Restriction enzyme (either Nt.BspQI, BsrDI, BtsI-v2). The samples were incubated at 50°C for 1 hour. After the incubation, 13 μl of each sample were taken and added with 3 μl of orange G (Table 6.1). Each sample was loaded into a 1.2% agarose gel and subjected to a current of 140 V for 1 hour and 50 minutes. Finally, the gel was incubated in the dark for 30 minutes in SYBR Gold buffer and revealed by fluorescence using a fluorescence scanner (Amersham Typhoon).

6.3.4 Quantification of the cleavage reactions

The quantification was performed with the use of Image J and Matlab. First, each column, except the control columns, was selected in order to take out the values of each line that we were certain of their nature (i.e. relaxed OC form, linear fragment, backbone, and Mini-Tn4430). The values of each band at each column were added together. Each value of each line was divided into the value of the respective column. Finally, for most of the assays (cf. Section 3.2 and 3.3) the value of each product appearing at time 0 (e.g. the linear product) was subtracted from the rest of the values of these products.

A | Appendices

A.1 Supplementary figures

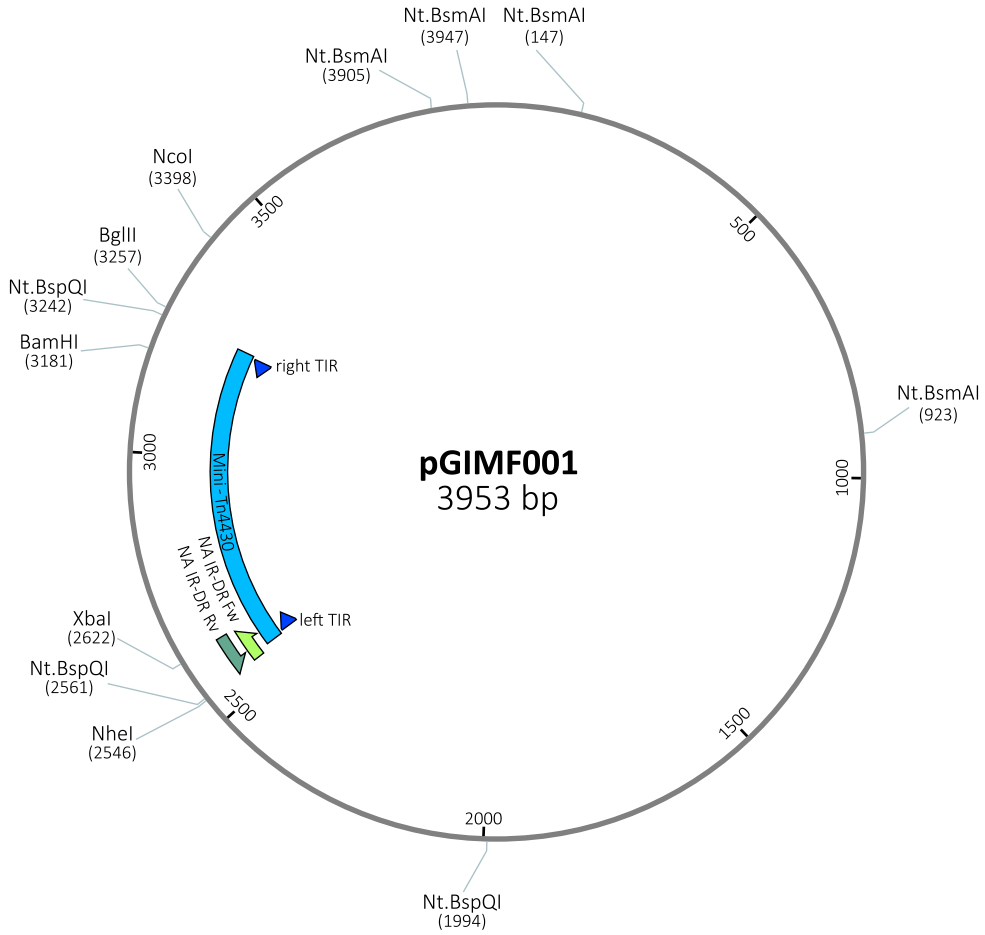

Figure A.1: Circular map of pGMF001 plasmid. Light blue represents the Mini-Tn4430. Dark blue arrows show the transposon ends. Dark and light green indicate the two synthetic oligonucleotides used to create the Mini =/Tn4430 with direct repeats. The sites of the different restriction enzymes and single-stranded endonucleases are shown.

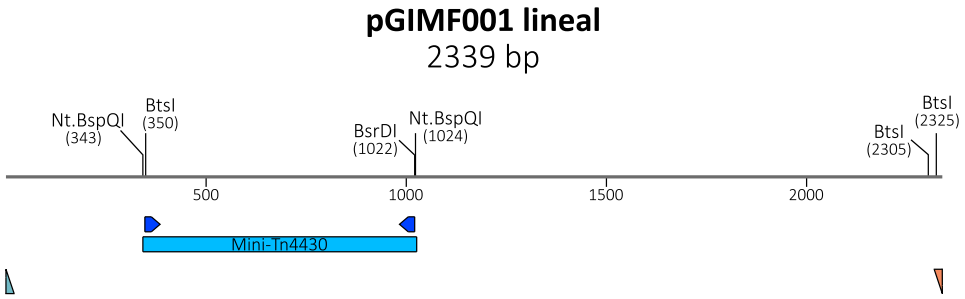

Figure A.2: Linear map of the pGMF001-derived substrate carrying Mini-Tn4430. Light blue represents the Mini-Tn4430 element. Dark blue arrows show the transposon terminal inverted repeats (TIR). Green and orange triangles show the position of the forward and reverse primers, respectively. The sites of the different restriction enzymes and single-stranded endonucleases are indicated.

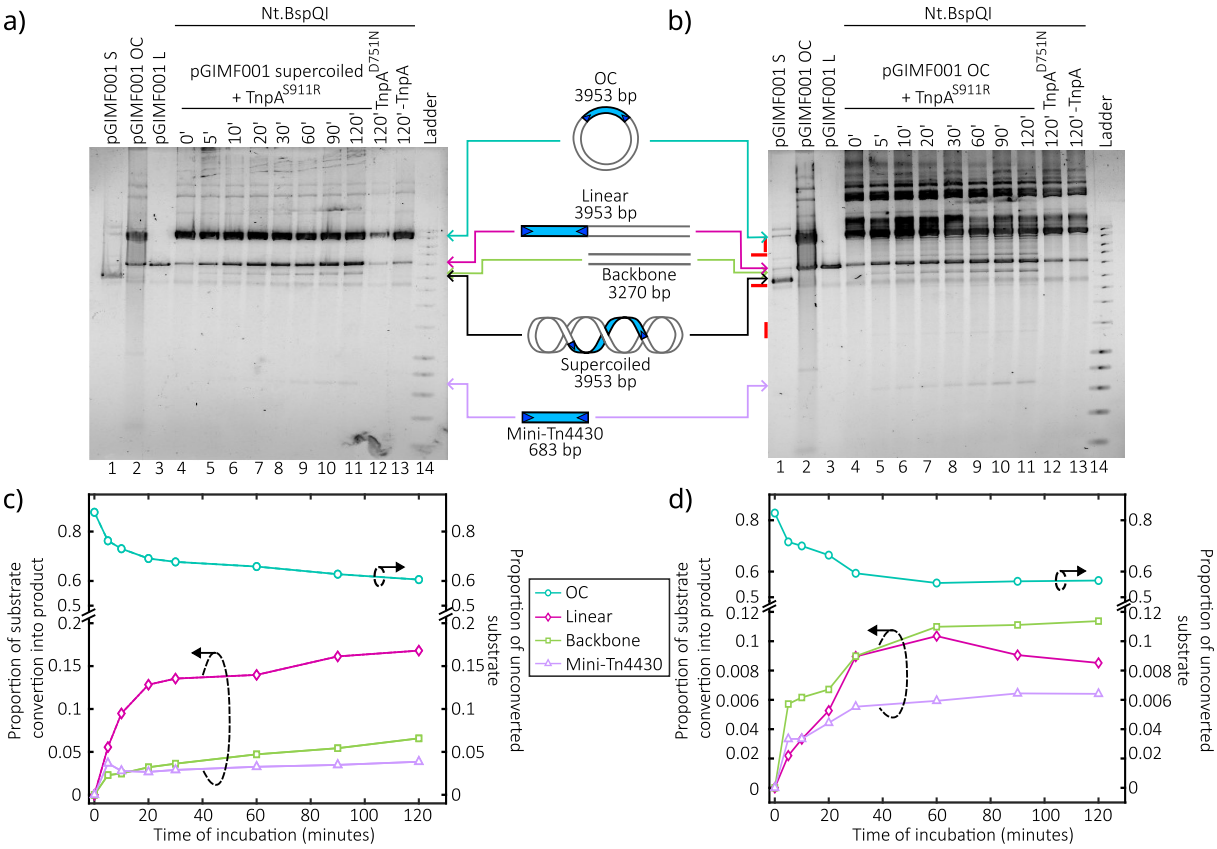

Figure A.3: Comparison of the cleavage reaction between the TnpA^{S911R} mutant and the supercoiled pGIMF001 plasmid (6.49 nM) (a) or the relaxed open circular (OC) pGIMF001 plasmid (6.49 nM) (b). Lanes 1 – 3 correspond to the three different forms of our plasmid: supercoiled (S), OC, and linear (L). The last two were obtained by digesting S with Nt.BsmAI and NcoI, respectively. Lanes 4 – 14 show the products of the cleavage reaction at different incubation times. Lanes 12 and 13 are the negative controls. The reaction was done at 34°C. 1.2% agarose gel and SYBR gold staining. c) and d) Quantification of the gels a) and b), respectively, performed using Image J (cf. Section 6.3.4). The cleavage products (linear, backbone, and Mini-Tn4430) use the left Y-axis, and the OC substrate uses the right Y-axis.

A.2 Creation of pGIVGG001 vector

Table A.1: Oligonucleotides

Oligonucleotide	Sequence (5' to 3' orientation)	Use
NA IR-DR Fw	AAAGCTCTTCGCAGTCTTAGCGTGGTTTTTTTCCGAAAT GCTGGCGGTACCCCCAAAATCAGAGTTAAAAAA	To create the right transposon end
NA IR-DR-Rv	CTAGTTTTTTAACTCTGATTTGGGGGTACCGCCAGCATTT CGGAAAAAAACCACGCTAAGCACTGCGAAGAGCTTTCTAG	To create the right transposon end

The supercoiled pGIMF001 plasmid purified from a bacterial colony was digested with two restriction enzymes (BmtI-HF and XbaI, Figure A.1) in order to take out the left transposon end. The oligonucleotides were designed with a change of the first base of the restriction site of both enzymes to have a control and if they are used, they would not be able to recognize the site and we can know if our construction worked.

In an Eppendorf, 10 μl of the DNA were mixed with 5 μl of CutSmart Buffer, 1 μl of each enzyme, and 33 μl of H_2O. This mix was incubated at 37°C for 15 minutes. In parallel, two oligonucleotides were ordered which correspond to the right transposon end (Table A.1). In an Eppendorf 5 μl of H_2O and STE 10x Buffer (Table 6.1) were mixed with 20 μl of each oligonucleotide. The reaction is incubated at 95°C for 10 minutes and the tube was left to cool down slowly with the machine.

Once both reactions have been done. For the cloning, two methods were tested. First, a Gibson assembly was performed. In a PCR tube, 0.02 pmols of the annealed oligonucleotides were mixed with 0.02 pmols of the digested plasmid and H_2O. The tube was incubated in a PCR machine for 15 minutes at 50°C. Secondly, a simple ligation with the T4 ligase was performed. 0.02 pmols of the digested plasmid were mixed with 0.06 pmols of the annealed oligonucleotides, 2 μl of T4 DNA ligase Buffer, 1μl of T4 DNA ligase, and H_2O. The reaction was left at room temperature for 10 minutes.

Finally, once both methods were done. Both samples were desalinated in a petri dish with water and a filter of 0.25 μm, for one hour. Once the hour has passed, the samples were collected. Next, the electroporation was done with electrocompetent *Eschericha coli* TOP10 strains. After, the bacteria were then spread out on LB-agar plates with ampicillin and incubated at 37°C overnight in order to select the clones that could grow, which are the ones that should contain our construction. Finally, precultures were done with liquid LB culture and ampicillin in order to extract the plasmid and do i) a digestion reaction with our constructions and one of the restriction enzymes used to delete the transposon end (BmtI-HF), and ii) send the constructions to sequencing (see primers in Table 6.2) to confirm that our cloning was correctly done.

Unfortunately, the constructions did not work, the results were either a simple deletion of the left transposon or a religation of the left transposon end. The problems could have

been the quantity of DNA vector and insert that we used. Another possibility is that the oligonucleotides were not able to bind themselves.

Bibliography

[1] B. McClintock, "The origin and behavior of mutable loci in maize," *Proceeding of the National Academy of Science*, vol. 36, pp. 344–355, 1950.

[2] S. Ravindran, "Barbara McClintock and the discovery of jumping genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 50, pp. 20 198–20 199, Dec. 2012.

[3] N. Fedoroff, S. Wessler, and M. Shure, "Isolation of the transposable maize controlling elements Ac and Ds," *Cell*, vol. 35, no. 1, pp. 235–242, Nov. 1983.

[4] H.-P. Döring and P. Starlinger, "Barbara McClintock's controlling elements: Now at the DNA level," *Cell*, vol. 39, no. 2, pp. 253–259, Dec. 1984.

[5] N. L. Craig, "A Moveable Feast: An Introduction to Mobile DNA," in *Mobile DNA*, 3rd ed. American Society for Microbiology, 2014, pp. 3–39.

[6] I. R. Arkhipova and I. A. Yushenova, "Giant Transposons in Eukaryotes: Is Bigger Better?" *Genome Biology and Evolution*, vol. 11, no. 3, pp. 906–918, Mar. 2019.

[7] S. R. Partridge, S. M. Kwong, N. Firth, and S. O. Jensen, "Mobile Genetic Elements Associated with Antimicrobial Resistance," *Clinical Microbiology Reviews*, vol. 31, no. 4, pp. e00 088–17, Aug. 2018.

[8] A. Toussaint and C. Merlin, "Mobile Elements as a Combination of Functional Modules," *Plasmid*, vol. 47, no. 1, pp. 26–35, Jan. 2002.

[9] A. B. Hickman and F. Dyda, "Mechanisms of DNA Transposition," in *Mobile DNA*, 3rd ed. American Society for Microbiology, 2014, pp. 531–553.

[10] L. A. Pray, "Transposons: The Jumping Genes," *Nature education*, vol. 1, no. 204, 2008.

[11] R. K. Aziz, M. Breitbart, and R. A. Edwards, "Transposases are the most abundant, most ubiquitous genes in nature," *Nucleic Acids Research*, vol. 38, no. 13, pp. 4207–4217, Jul. 2010.

[12] E. Nicolas, C. A. Oger, N. Nguyen, M. Lambin, A. Draime, S. C. Leterme, M. Chandler, and B. F. J. Hallet, "Unlocking Tn3-family transposase activity in vitro unveils an asymetric pathway for transposome assembly," *Proceedings of the National Academy of Sciences*, vol. 114, no. 5, Jan. 2017.

[13] E. Nicolas, C. A. Oger, C. Stulemeijer, N. Aryanpour, N. Nguyen, M. Chandler, and B. F. J. Hallet, "A replication hijacking mechanism for Tn3-family replicative transposition," In Prep.

[14] L. S. Frost, R. Leplae, A. O. Summers, and A. Toussaint, "Mobile genetic elements: The agents of open source evolution," *Nature Reviews Microbiology*, vol. 3, no. 9, pp. 722–732, Sep. 2005.

[15] P. M. Bennett, "Genome Plasticity," in *Genomics, Proteomics, and Clinical Bacteriology*, N. Woodford and A. P. Johnson, Eds. Totowa, NJ: Humana Press, 2004, pp. 71–113.

[16] D. Lereclus, J. Mahillon, G. Menou, and M.-M. Lecadet, "Identification of Tn4430, a transposon of Bacillus thuringiensis functional in Escherichia coli," *Molecular and General Genetics MGG*, vol. 204, no. 1, pp. 52–57, Jul. 1986.

[17] B. Piégu, S. Bire, P. Arensburger, and Y. Bigot, "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*, p. 20, 2015.

[18] M. J. Curcio and K. M. Derbyshire, "The outs and ins of transposition: From Mu to Kangaroo," *Nature Reviews Molecular Cell Biology*, vol. 4, no. 11, pp. 865–877, Nov. 2003.

[19] D. J. Finnegan, "Eukaryotic transposition elements and genome evoolution," *Elsevier Science*, vol. 5, no. 4, pp. 103–107, 1989.

[20] T. Wicker, F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, and A. H. Schulman, "A unified classification system for eukaryotic transposable elements," *Nature Reviews Genetics*, vol. 8, no. 12, pp. 973–982, Dec. 2007.

[21] J. L. Goodier, "Restricting retrotransposons: A review," *Mobile DNA*, vol. 7, no. 1, p. 16, Dec. 2016.

[22] T. H. Eickbush and V. K. Jamburuthugoda, "The diversity of retrotransposons and the properties of their reverse transcriptases," *Virus Research*, vol. 134, no. 1-2, pp. 221–234, Jun. 2008.

[23] D. J. Finnegan, "Retrotransposons," *Current Biology*, vol. 22, no. 11, pp. R432–R437, 2012.

[24] A. Kumar and J. Bennetzen, "Plant Retrotransposons," *Annual review of genetics*, vol. 33, pp. 479–532, Feb. 1999.

[25] A. Beauregard, M. J. Curcio, and M. Belfort, "The Take and Give Between Retrotransposable Elements and their Hosts," *Annual Review of Genetics*, vol. 42, no. 1, pp. 587–617, Dec. 2008.

[26] C. Feschotte and E. J. Pritham, "DNA Transposons and the Evolution of Eukaryotic Genomes," *Annual Review of Genetics*, vol. 41, no. 1, pp. 331–368, Dec. 2007.

[27] K. Mizuuchi, "Transpositional Recombination: Mechanistic Insights from Studies of Mu and Other Elements," *Annual Review of Biochemistry*, vol. 61, pp. 1011–1051, 1992.

[28] E. Nicolas, M. Lambin, D. Dandoy, C. Galloy, N. Nguyen, C. A. Oger, and B. Hallet, "The Tn3-family of Replicative Transposons," in *Mobile DNA III*, 3rd ed. Washington DC: American Society for Microbiology, 2014, pp. 693–726.

[29] J. Bender and N. Kleckner, "Genetic Evidence That Tn10 Transposes by a Nonreplicative Mechanism," *Cell*, vol. 45, pp. 801–815, 1986.

[30] A. B. Hickman and F. Dyda, "DNA Transposition at Work," *Chemical reviews*, vol. 116, no. 20, pp. 12 758–12 784, Oct. 2016.

[31] F. Heffron, B. J. McCarthy, H. Ohtsubo, and E. Ohtsubo, "DNA sequence analysis of the transposon Tn3: Three genes and three sites involved in transposition of Tn3," *Cell*, vol. 18, pp. 1153–1163, 1979.

[32] M. Chandler, F. de la Cruz, F. Dyda, A. B. Hickman, G. Moncalian, and B. Ton-Hoang, "Breaking and joining single-stranded DNA: The HUH endonuclease superfamily," *Nature Reviews Microbiology*, vol. 11, no. 8, pp. 525–538, Aug. 2013.

[33] N. D. Grindley, K. L. Whiteson, and P. A. Rice, "Mechanisms of Site-Specific Recombination," *Annual Review of Biochemistry*, vol. 75, no. 1, pp. 567–605, Jun. 2006.

[34] M. Krupovic, K. S. Makarova, P. Forterre, D. Prangishvili, and E. V. Koonin, "Casposons: A new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity," *BMC Biology*, vol. 12, no. 1, p. 36, Dec. 2014.

[35] A. B. Hickman, M. Chandler, and F. Dyda, "Integrating prokaryotes and eukaryotes: DNA transposases in light of structure," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 45, no. 1, pp. 50–69, Feb. 2010.

[36] I. V. Nesmelova and P. B. Hackett, "DDE Transposases: Structural Similarity and Diversity," *Advanced drug delivery reviews*, vol. 62, no. 12, pp. 1187–1195, Sep. 2010.

[37] M. Nowotny, "Retroviral integrase superfamily: The structural perspective," *EMBO Reports*, vol. 10, no. 2, pp. 144–151, Feb. 2009.

[38] A. B. Hickman and F. Dyda, "Mechanisms of DNA Transposition," *Microbiology spectrum*, vol. 3, no. 2, pp. 1–22, Apr. 2015.

[39] J. A. Shapiro, "Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements." *Proceeding of the National Academy of Science of the United States of America*, vol. 76, no. 4, pp. 1933–1937, 1979.

[40] E. Nicolas, M. Lambin, D. Dandoy, C. Galloy, N. Nguyen, C. A. Oger, and B. Hallet, "The Tn3-family of Replicative Transposons," *Microbiology Spectrum*, vol. 3, no. 4, p. 3.4.14, Jul. 2015.

[41] P. Rousseau, C. Tardin, N. Tolou, L. Salomé, and M. Chandler, "A model for the molecular organisation of the IS911 transpososome," *Mobile DNA*, vol. 1, no. 1, p. 16, Jun. 2010.

[42] W. S. Reznikoff, "Transposon Tn5," *Annual Review of Genetics*, vol. 42, pp. 269–286, 2008.

[43] A. B. Hickman, A. Regier Voth, H. Ewis, X. Li, N. L. Craig, and F. Dyda, "Structural insights into the mechanism of double strand break formation by Hermes, a hAT family

eukaryotic DNA transposase," *Nucleic Acids Research*, vol. 46, no. 19, pp. 10 286–10 301, 2018.

[44] C. M. Singer, D. Joy, D. J. Jacobs, and I. V. Nesmelova, "Rigidity and flexibility characteristics of DD[E/D]-transposases Mos1 and Sleeping Beauty," *Proteins*, vol. 87, no. 4, pp. 313–325, 2019.

[45] E. W. May and N. L. Craig, "Switching from Cut-and-Paste to Replicative Tn7 Transposition," *Science*, vol. 272, pp. 401–404, 1996.

[46] J. E. Peters, "Tn7," *Microbiology Spectrum*, vol. 2, no. 5, 2014.

[47] K. A. Majorek, S. Dunin-Horkawicz, K. Steczkiewicz, A. Muszewska, M. Nowotny, K. Ginalski, and J. M. Bujnicki, "The RNase H-like superfamily: New members, comparative structural analysis and evolutionary classification," *Nucleic Acids Research*, vol. 42, no. 7, pp. 4160–4179, Apr. 2014.

[48] R. T. Clubb, H. Savilahti, K. Mizuuchi, A. M. Gronenborn, and G. M. Clore, "A novel class of winged helix-turn-helix protein: The DNA-binding domain of Mu transposase," *Structure*, vol. 2, no. 11, pp. 1041–1048, 1994.

[49] F. Dyda, M. Chandler, and A. B. Hickman, "The emerging diversity of transpososome architectures," *Quarterly Reviews of Biophysics*, vol. 45, no. 4, pp. 493–521, Nov. 2012.

[50] C. Claeys Bouuaert and R. Chalmers, "Hsmar1 Transposition Is Sensitive to the Topology of the Transposon Donor and the Target," *PLoS ONE*, vol. 8, no. 1, p. e53690, Jan. 2013.

[51] G. Blundell-Hunter, M. Tellier, and R. Chalmers, "Transposase subunit architecture and its relationship to genome size and the rate of transposition in prokaryotes and eukaryotes," *Nucleic Acids Research*, vol. 46, no. 18, pp. 9637–9646, Oct. 2018.

[52] W. Zhou, G. Liang, P. L. Molloy, and P. A. Jones, "DNA methylation enables transposable element-driven genome expansion," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 32, pp. 19 359–19 366, Aug. 2020.

[53] C. Claeys Bouuaert, K. Lipkow, S. S. Andrews, D. Liu, and R. Chalmers, "The autoregulation of a eukaryotic DNA transposon," *eLife*, vol. 2, p. e00668, Jun. 2013.

[54] A. Pflieger, J. Jaillet, A. Petit, C. Augé-Gouillou, and S. Renault, "Target Capture during Mos1 Transposition," *The Journal of Biological Chemistry*, vol. 289, no. 1, pp. 100–111, Jan. 2014.

[55] J. Sakai and N. Kleckner, "The Tn10 Synaptic Complex Can Capture a Target DNA only after Transposon Excision," *Cell*, vol. 89, pp. 205–214, 1997.

[56] K. Lipkow, N. Buisine, and R. Chalmers, "Promiscuous Target Interactions in the mariner Transposon Himar1," *Journal of Biological Chemistry*, vol. 279, no. 47, pp. 48 569–48 575, Nov. 2004.

[57] D. Z. Naigamwalla and G. Chaconas, "A new set of Mu DNA transposition intermediates: Alternate pathways of target capture preceding strand transfer." *The EMBO Journal*, vol. 16, no. 17, pp. 5227–5234, Sep. 1997.

[58] K. C. Neuman, "Single-molecule Measurements of DNA Topology and Topoisomerases," *The Journal of Biological Chemistry*, vol. 285, no. 25, pp. 18 967–18 971, Jun. 2010.

[59] S. M. Mirkin, "DNA Topology: Fundamentals," in *eLS*, 1st ed., John Wiley & Sons, Ltd, Ed. Wiley, May 2001.

[60] T. A. Baker and K. Mizuuchi, "DNA-promoted assembly of the active tetramer of the Mu transposase." *Genes & Development*, vol. 6, no. 11, pp. 2221–2232, Jan. 1992.

[61] Z. Wang and R. M. Harshey, "Crucial role for DNA supercoiling in Mu transposition: A kinetic study." *Proceedings of the National Academy of Sciences*, vol. 91, pp. 699–703, 1994.

[62] M. G. Surette and G. Chaconas, "A protein factor which reduces the negative supercoiling requirement in the Mu DNA strand transfer reaction is Escherichia coli integration host factor | Elsevier Enhanced Reader," *The Journal of Biological Chemistry*, vol. 264, no. 5, pp. 3028–3034, 1989.

[63] R. M. Chalmers and N. Kleckner, "IS10/Tn10 transposition efficiently accommodates diverse transposon end configurations." *The EMBO Journal*, vol. 15, no. 18, pp. 5112–5122, Sep. 1996.

[64] C. Claeys Bouuaert, D. Liu, and R. Chalmers, "A Simple Topological Filter in a Eukaryotic Transposon as a Mechanism To Suppress Genome Instability," *Molecular and Cellular Biology*, vol. 31, no. 2, pp. 317–327, Jan. 2011.

[65] A. Travers and G. Muskhelishvili, "A common topology for bacterial and eukaryotic transcription initiation?" *EMBO Reports*, vol. 8, no. 2, pp. 147–151, Feb. 2007.

[66] C. J. Dorman, "DNA supercoiling and transcription in bacteria: A two-way street," *BMC Molecular and Cell Biology*, vol. 20, no. 1, p. 26, Jul. 2019.

[67] M. G. Surette, S. J. Buch, and G. Chaconas, "Transpososomes: Stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA," *Cell*, vol. 49, pp. 253–262, 1987.

[68] R. Craigie and K. Mizuuchi, "Transposition of Mu DNA: Joining of Mu to target DNA can be uncoupled from cleavage at the ends of Mu," *Cell*, vol. 51, no. 3, pp. 493–501, Nov. 1987.

[69] R. Craigie, D. J. Arndt-Jovin, and K. Mizuuchi, "A defined system for the DNA strand-transfer reaction at the initiation of bacteriophage Mu transposition: Protein and DNA substrate requirements." *Proceeding of the National Academy of Science of the Unoted States of America*, vol. 82, pp. 7570–7574, 1985.

[70] Z. Yin, S. Pathania, and R. M. Harshey, "The Mu Transposase Interwraps Distant DNA Sites within a Functional Transpososome in the Absence of DNA Supercoiling," *The Journal of Biological Chemistry*, vol. 280, no. 7, pp. 6149–6156, 2005.

[71] D. B. Haniford, "Transpososome Dynamics and Regulation in Tn10 Transposition," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 41, no. 6, pp. 407–424, Jan. 2006.

[72] R. Chalmers, A. Guhathakurta, H. Benjamin, and N. Kleckner, "IHF Modulation of Tn10 Transposition: Sensory Transduction of Supercoiling Status via a Proposed Protein/DNA Molecular Spring," *Cell*, vol. 93, no. 5, pp. 897–908, May 1998.

[73] G. C. Cerqueira, A. M. Earl, C. M. Ernst, Y. H. Grad, J. P. Dekker, M. Feldgarden, S. B. Chapman, J. L. Reis-Cunha, T. P. Shea, S. Young, Q. Zeng, M. L. Delaney, D. Kim, E. M. Peterson, T. F. O'Brien, M. J. Ferraro, D. C. Hooper, S. S. Huang, J. E. Kirby, A. B. Onderdonk, B. W. Birren, D. T. Hung, L. A. Cosimi, J. R. Wortman, C. I. Murphy, and W. P. Hanage, "Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks," *Proceedings of the National Academy of Sciences*, vol. 114, no. 5, pp. 1135–1140, Jan. 2017.

[74] H. Z. Smith and B. Kendall, "Carbapenem Resistant Enterobacteriaceae," in *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2022.

[75] M. Lambin, E. Nicolas, C. A. Oger, N. Nguyen, D. Prozzi, and B. Hallet, "Separate structural and functional domains of Tn4430 transposase contribute to target immunity," *Molecular Microbiology*, vol. 83, no. 4, pp. 805–820, 2012.

[76] E. Nicolas, M. Lambin, and B. Hallet, "Target Immunity of the Tn3-Family Transposon Tn4430 Requires Specific Interactions between the Transposase and the Terminal Inverted Repeats of the Transposon," *Journal of Bacteriology*, vol. 192, no. 16, pp. 4233–4238, Aug. 2010.

[77] V. Vanhooff, C. Galloy, H. Agaisse, D. Lereclus, B. Révet, and B. Hallet, "Self-control in DNA site-specific recombination mediated by the tyrosine recombinase TnpI," *Molecular Microbiology*, vol. 60, no. 3, pp. 617–629, 2006.

[78] J. Mahillon and D. Lereclus, "Structural and functional analysis of Tn4430: Identification of an integrase-like protein involved in the co-integrate-resolution process." *The EMBO Journal*, vol. 7, no. 5, pp. 1515–1526, May 1988.

[79] O. d'Udekem d'Acoz, "Functional interaction between Tn4430 replicative transposition and DNA replication," Ph.D. dissertation, Université catholique de Louvain, Louvain la Neuve, 2022.

[80] B. Hallet, "Tn-3D: Transposition at high resolution," Unpublished Data, 2019.

[81] A. V. Shkumatov, N. Aryanpour, C. A. Oger, G. Goossens, B. F. Hallet, and R. G. Efremov, "Metamorphism of catalytic domain controls transposition in Tn3 family transposases," Molecular Biology, Preprint, Feb. 2022.

[82] C. Guynet, E. Nicolas, B. Ton-Hoang, J.-Y. Bouet, and B. Hallet, "First Biochemical Steps on Bacterial Transposition Pathways," *Methods in Molecular Biology (Clifton, N.J.)*, vol. 2075, pp. 157–177, 2020.

[83] L. Codemo, "Tn4430 transposase activity on Mini-Tn-containing DNA molecules in vitro," Master's thesis, Université catholique de Louvain, Jan. 2022.

[84] R. Chalmers, S. Sewitz, K. Lipkow, and P. Crellin, "Complete Nucleotide Sequence of Tn10," *Journal of Bacteriology*, vol. 182, no. 10, pp. 2970–2972, May 2000.

[85] D. Morisato and N. Kleckner, "TnlO Transposition and Circle Formation In Vitro," *Cell*, vol. 51, pp. 101–111, 1987.

[86] P. Hyman and S. T. Abedon, "Bacteriophage: Overview," in *Encyclopedia of Microbiology (Fourth Edition)*, T. M. Schmidt, Ed. Oxford: Academic Press, Jan. 2019, pp. 441–457.

[87] C. Claeys Bouuaert and R. Chalmers, "Transposition of the human Hsmar1 transposon: Rate-limiting steps and the importance of the flanking TA dinucleotide in second strand cleavage," *Nucleic Acids Research*, vol. 38, no. 1, pp. 190–202, Jan. 2010.

[88] P. L. Privalov, "Cold Denaturation of Protein," *Biochemistry and Molecular Biology*, p. 26, 1990.

[89] J. C. Bischof and X. He, "Thermal Stability of Proteins," *Annals of the New York Academy of Sciences*, vol. 1066, no. 1, pp. 12–33, 2006.

[90] H. W. Benjamin and N. R. Cozzarelli, "Isolation and characterization of the Tn3 resolvase synaptic intermediate." *The EMBO Journal*, vol. 7, no. 6, pp. 1897–1905, 1988.

[91] S. He, M. Chandler, A. M. Varani, A. B. Hickman, and F. Dyda, "Mechanisms of Evolution in High-Consequence Drug Resistance Plasmids," *American Society for Microbiology*, vol. 7, no. 6, pp. 1–11, 2016.

[92] M. Fernandez, A. V. Shkumatov, Y. Liu, C. Stulemeijer, S. Derclaye, R. G. Efremov, B. Hallet, and D. Alsteens, "AFM-based force spectroscopy unravels the stepwise-formation of a DNA transposition complex driving multi-drug resistance dissemination," Biophysics, Preprint, Jul. 2022.

[93] M. Arthur, C. Molinas, F. Depardieu, and P. Courvalin, "Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in Enterococcus faecium BM4147," *Journal of Bacteriology*, vol. 175, no. 1, pp. 117–127, 1993.

[94] C. Oger, "Ends-synapsis and target integration during replicative transposition of the Tn3-family transposon Tn4430," Ph.D. dissertation, Université catholique de Louvain, 2018.

[95] E. Nicolas, "Biochemical activities of the transposase TnpA in transposition and target immunity mediated by the replicative transposon Tn4430," Ph.D. dissertation, Université catholique de Louvain, 2011.

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