UCLouvain

FACULTE DES SCIENCES

Faculté des sciences

Finding transposition partners of Tn4430 transposon using proximity labelling

Author: Marius SWINNEN

Supervisors: Prof. Bernard HALLET
PhD Candidate Claire STULEMEIJER

Readers: Prof. Henri BATOKO
Prof. Pascal HOLS

Laboratory of Biochemistry and Genetics of Microorganisms (BGM)

Master in Biochemistry, Molecular and Cell Biology (Research Focus)
Academic year 2023-2024

Acknowledgments

Embarking on this master thesis journey has been a transformative experience, and I owe its success to the invaluable support, guidance, and encouragement of numerous individuals who have played a pivotal role in shaping this work.

Foremost, my sincere appreciation goes to my supervisor, Bernard Hallet, whose contagious passion for the expansive realm of transposition fueled my curiosity and dedication. He always managed to communicate his excitement for the project, his ideas, and experiments for future research.

I extend heartfelt gratitude to Claire, whose generosity in sharing her expertise and time has been indispensable. Her patience, valuable advice, and accessibility provided a steady foundation whenever doubts arose. I am particularly grateful for the autonomy she granted me, allowing me to experiment by myself, to make mistakes, and to learn from them.

A warm thank you is extended to the entire lab community. From my first day in the lab, I have felt a welcoming atmosphere. Special acknowledgment goes to Brigitte, Damien, and Cyril for their tireless efforts within the lab, assisting with queries, and aiding in the navigation of lab resources.

I am indebted to Nicolas, Loïc, Alexandre, and Shany for their engaging discussions during team meetings. Their keen interest in my project, coupled with insightful questions, has significantly contributed to refining, and expanding the scope of my work.

Finally, I express my deepest gratitude to my family for their unwavering support throughout my academic journey. Their encouragement, genuine interest in my pursuits, and their questions every Friday when going back home. A special thank you is reserved for my dad, whose attentive ear, and efforts to comprehend the intricacies of my project have been invaluable.

Table of content

Abstract

This master thesis aims to identify potential transposition partners of the bacterial transposon Tn4430. Aligned with the broader goal of comprehending the transposition dynamics of Tn4430 and related Tn3-family transposons, we established a proximity labelling assay in E. *coli*. This assay aimed to pinpoint host proteins in the vicinity of the transposase of Tn4430. Validating the identified proteins as interactants was achieved through bacterial-two-hybrid assays. This analysis identified multiple potential interactants of TnpA, with several that were previously shown to play a role in the transposition mechanism of several other elements. Further exploration is needed to ascertain the functional roles of these interactions in the transposition mechanism of Tn4430.

I. Introduction

I.1 Transposable elements

Transposable elements, commonly referred to as transposons, group a broad variety of genetic elements that have the capacity to move and integrate within genomes. At the primary level, these elements exhibit hallmark structural features such as specific DNA sequences marking the extremities and at least one gene coding for a protein, generically referred as a transposase, that is responsible for their mobility. Variations in their configuration, including the presence of additional passenger elements and/or transposase accessory genes, add layers of complexity to their diversity. The dynamic nature of transposons, allowing for the insertion, deletion, or duplication of DNA segments, imparts them with a profound influence over the genetic makeup of organisms (Ayarpadikannan and Kim 2014; Dubin, Mittelsten Scheid, and Becker 2018; Feschotte and Pritham 2007; Pourrajab and Hekmatimoghaddam 2021).

One striking illustration of this evolutionary relevance are Alu elements. These primate-specific repeats comprise about 10-11% of the human genome and continue to insert in modern humans (Deininger 2011; Häsler and Strub 2006). Extended homology of Alu elements is a source for non-allelic homologous recombination either leading to duplication or deletion of DNA fragments (Häsler and Strub 2006). Somatic insertion of these elements can lead to both positive effects such as genetic diversity and negative effects such as diseases by insertional mutagenesis. Alu elements were also reported to have wide-ranging effects on gene expression as they became effectors of gene expression (Brosius 1999; Tomilin 1998). Numerous other instances underscore the pivotal role of transposons in evolution. Consider, for instance, the peppered moth's adaptation during the Industrial Revolution, wherein transposon-driven modifications in its camouflage genes facilitated its adjustment to shifting environments by altering its coloration (Hof et al. 2016).

But their most worrisome characteristic probably lies in their implication in the dissemination of antibiotic resistance genes among bacteria (Ellabaan et al. 2021; Hedges and Jacob 1974; Noel, Petrey, and Palmer 2022). Transposons, including those belonging to the widespread Tn3 family, have gained notoriety for their role in promoting the

proliferation of resistance to different classes of antibiotics, including to carbapenems, the “last chance” antibiotics (Pasquali et al. 2005). Nonetheless, the importance of transposons transcends antibiotic resistance genes alone. These versatile genetic elements can serve as carriers for any arrays of passenger genes, encompassing heavy metal resistance genes (Cervantes, Chávez, and Vaca 1991), toxin-antitoxin systems (Kamruzzaman, Wu, and Iredell 2021), virulence factors (Morales et al. 2023), and more, making them a driving force of bacterial adaptation.

I.1.1 Diversity of transposons

In recent years, the number of identified transposons has surged in tandem with the number of sequenced genomes. Their remarkable diversity in terms of genetic composition and transposition mechanisms underscore the necessity for a systematic classification. Within the scope of our research, which is focused on the identification of potential transposition partners of Tn4430 transposase (TnpA), the classification based on the mechanism of transposition emerges as the most pertinent. In their review, Hickman and Dyda describe the mechanisms of the four known types of transposition reactions catalysed by (1) RNase H-like transposases (also known as DD(E/D) enzymes); (2) HUH single-stranded DNA transposases; (3) serine transposases; and (4) tyrosine transposases (Hickman and Dyda 2015a).

From a mechanistic point of view, there are only a few ways in which transposases catalyse the required DNA strand breakage and re-joining reactions. This translates in a few different structures of catalytic domain found in transposases (see Figure **I.1.1-1)**. In this section, we describe the chemistry of cleavage and strand transfer for those four types of transposition reaction.

I.1.1.1 RNase H-like catalytic domains

RNase H-like transposases, commonly called DDE/DDD transposases contain a catalytic domain with three acidic residues (DDE or DDD) that coordinate two bivalent metal ions. This allows to activate either a molecule of water or a 3'-OH group of a nucleophile for a nucleophilic attack on the phosphodiester bond (Hickman and Dyda 2015a). RNase H-like transposases are thus able to catalyse two types of reactions. First, a nucleophilic attack by an activated water molecule on the phosphate, breaking a phosphodiester bond and generating a 3′-OH and a 5′-phosphate group at the cleavage site. Second, a transesterification reaction between the nucleophile 3'-OH group of a terminal nucleotide and another DNA strand. This reaction leads to the simultaneous cleavage of one strand and the joining one to another. This reaction is used in many different transposition mechanisms to integrate the transposon into a new site but also to form hairpin as well as circular intermediates.

Both reactions of cleavage and strand transfer catalysed by DDE/DDD transposases are highly dependent on the divalent ion. Mg^{2+} and Mn^{2+} are the only two cations able to support all steps of transposition. Interestingly, Ca^{2+} which does not differ significantly in size, generally does not support hydrolysis and transesterification reactions (Rosta, Yang, and Hummer 2014). This specificity highlights the importance of bivalent ions to precisely position the reactive groups, activate the nucleophile and stabilize the transition state of the SN_2 reaction (Hickman and Dyda 2015a). This superfamily of transposons, to which Tn4430 belongs, will be more comprehensively detailed in **I.1.2**.

I.1.1.2 HUH catalytic domain

The HUH catalytic domain is used by many DNA transposons to cut ssDNA. HUH catalytic domains are used by a variety of proteins initiating processes such as plasmid rolling circle replication or the conjugative transfer of plasmids between cells. The HUH fold contains one ("Y1 Transposases") or two ("Y2 Transposases") nucleophilic tyrosine residues in its active site. Single stranded DNA cleavage happens through the formation of a 5′-phosphotyrosine covalent intermediate. The intermediate is then resolved with the attack by a terminal 3′-OH group of another DNA strand on the 5′-phosphotyrosine

linkage. Like RNase H-like catalytic domains, some of the HUH nucleases use divalent metallic ions to support the chemical steps of the transposition mechanism. However, the range of ions that are used is wider, suggesting that the catalytic site is more tolerant than RNase H-like catalytic site.

I.1.1.3 Serine transposases

Most of bacterial transposons encoding for a serine transposase use a circular intermediate of the transposable elements to recombine with the target DNA molecule. Those transposases are thought to share many catalytic features with serine site-specific recombinases. DNA cleavage of those recombinases involves an active site serine nucleophile that attacks the phosphodiester bond generating a covalent 5′-phosphoserine intermediate and a free 3′-OH group. The resolution of the covalent intermediate by a terminal 3′-OH group of another strand results in strand transfer. Unlike the two precedent types of catalytic domains, no divalent metal ions are required, and key roles of the catalysis is played by arginine residues in the active site (Hickman and Dyda 2015a; Keenholtz et al. 2011, 2013).

I.1.1.4 Tyrosine transposases

Many conjugative transposons use tyrosine transposases to mobilize. These transposases are thought to be related to well-studied tyrosine recombinases both in terms of structure and mechanism. Like serine transposases and recombinases, tyrosine transposases cleave DNA using the active site tyrosine residue to make a nucleophilic attack on a phosphodiester bond and generate a phosphotyrosine intermediate. However, in opposition to the three other types of transposases seen before, the cleavage polarity is inverted, thus leading to a 3′-phosphotyrosine linkage.

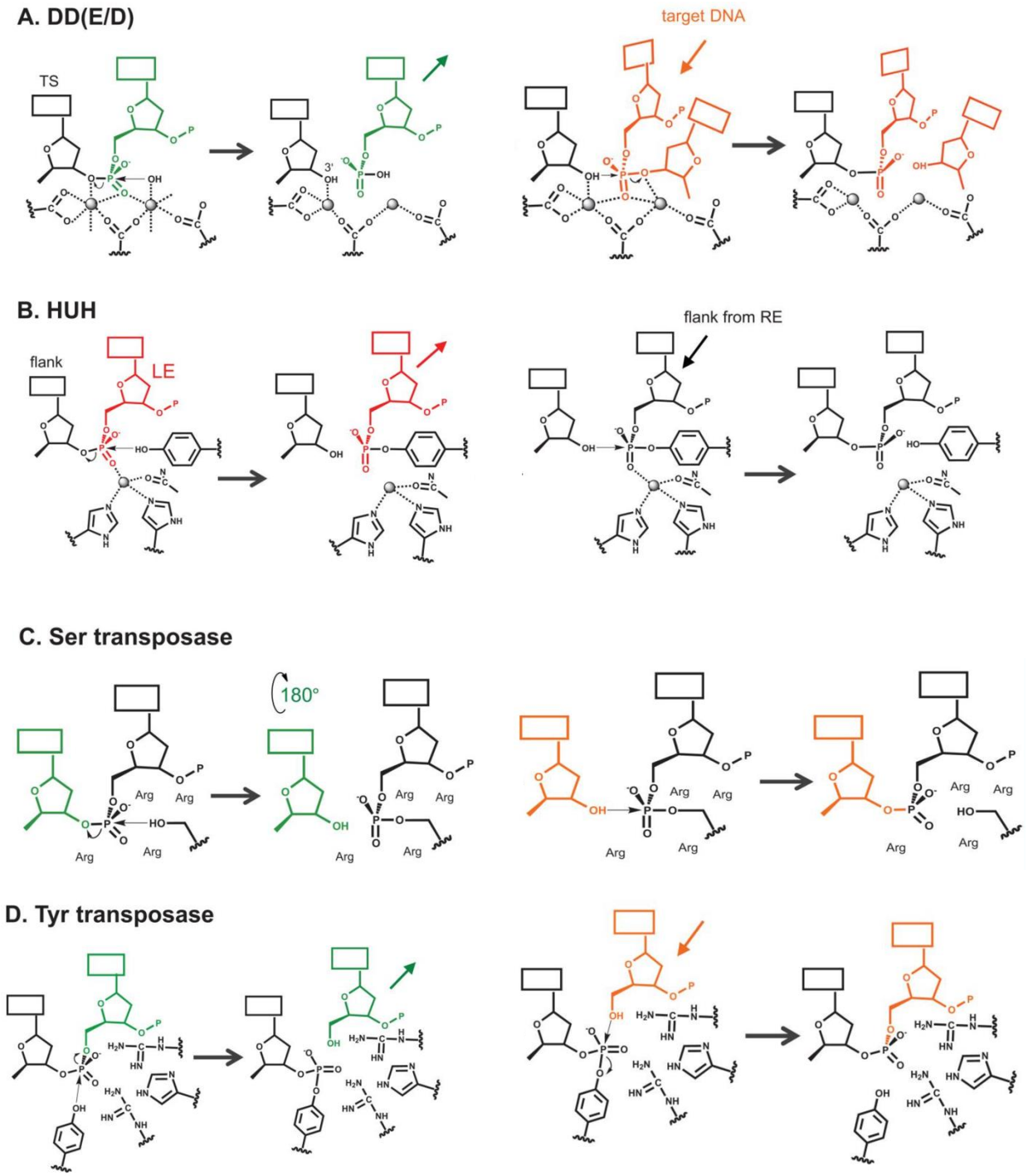


Figure I.1.1-1 - Basic chemical reactions catalysed by DNA transposases. (A) An RNase H-like active site, based on structures of PFV intasomes. The green DNA represents the cleaved dinucleotide, and orange is the target strand. Spheres indicate bound metal ions. (B) HUH nuclease active site acting on single-stranded DNA (based on PDB ID 2X06 of IS*Dra2* TnpA). Shown is the reaction that occurs at the transposon Left End (LE). After cleavage, the DNA flanking the LE (black) remains in the active site; upon exchange of α-helices between the two active sites of the dimeric transposase, the cleaved LE moves to the other monomer where it is joined to the cleaved RE to form a circular excised transposon (not shown). At the same time, the flanking DNA from the RE of the transposon switches active sites (as shown here in black) and subsequent joining results in a sealed donor backbone. (C) DNA cleavage catalysed by a serine recombinase. The active serine is surrounded by many Arg residues. Upon 180° rotation of one dimer within a tetramer, one strand rotates out of the active site (green) while another rotates in (orange). (D) DNA cleavage catalysed by a tyrosine recombinase. Figure from Hyckman and Dyda, 2015.

I.1.2 DDE/D family of transposons

As described in **I.1.1**, transposases within this family exhibit a highly conserved DDE/DDD triad of acidic residues and share an RNase H-like fold domain which brings the three conserved residues into spatial proximity, thereby facilitating their catalytic activity.

Transposition mediated by DDE/D transposases is a two-step process **(Figure I.1-1)** (Nesmelova and Hackett 2010). Initially, TnpA catalyzes a nucleophilic attack involving a water molecule and a phosphate group located at the transposon end **(Figure I.1-1A)**(Nesmelova and Hackett 2010). This action yields a 3'-OH group at each end of the transposon. The activation of the water molecule relies on the coordinated involvement of two divalent metal ions (Mg^{2+} or Mn^{2+}) within the active site (Hickman and Dyda 2015b). Subsequently, the second step involves a DNA strand transfer reaction **(Figure I.1-1B)**, wherein the exposed 3'-OH ends of the transposon become fused to the target DNA (Nesmelova and Hackett 2010). This nucleophilic attack by the two 3'-OH groups usually takes place at staggered positions of the target separated by a few base pairs, typically ranging from 2 to 9 bp (Curcio and Derbyshire 2003). This specific distance exhibits variability depending on the transposon under consideration. Staggered cleavage of the target results in short target sequence duplications that flank the transposon in its new genomic locus, thus serving as a distinctive hallmark of DDE/D transposition.

While all DDE/D transposases execute fundamentally similar chemical reactions, the mechanisms they employ to generate the DNA substrate for strand transfer can exhibit significant variations among different transposons, defining different transposition pathways. Many RNase H-like transposases catalyze “non-replicative” or “cut-and-paste” transposition. These enzymes must catalyze a double strand break (DSB) to liberate the mobile element from its donor site. This transposition pathway clearly differs from replicative transposition mechanisms where no DSBs occur. One of the most studied elements using the replicative transposition pathway is bacteriophage Mu. This pathway is known as the "paste-and-copy" pathway and was first proposed by James Shapiro in the context of the bacteriophage Mu and several other replicative transposons (Shapiro

1979). Since, this transposition mechanism has been studied and demonstrated in much detail.

In Shapiro's model, the transposase catalyzes the hydrolysis of phosphodiester bonds within the donor DNA, yielding 3'-OH ends capable of catalyzing transesterification reactions on the target DNA **(see Figure I.1.2-1A and B)**. This catalytic action gives rise to a distinct structural entity termed a "Shapiro intermediate" **(Figure I.1.2-1C)**. Subsequently, the host cell's replication machinery comes into play, synthesizing complementary strands of the transposon. This process culminates in the formation of a co-integrate, where the donor and target DNAs become linked to each other by two copies of the transposon. The resolution of this co-integrate intermediate can be facilitated by a site-specific protein encoded by the transposon itself, typically a serine (TnpR), a tyrosine resolvase (TnpI) or a heteromeric resolvase that combines both a tyrosine recombinase (TnpS) and an additional helper protein (TnpT), thereby completing the transposition process (Nicolas et al. 2015).

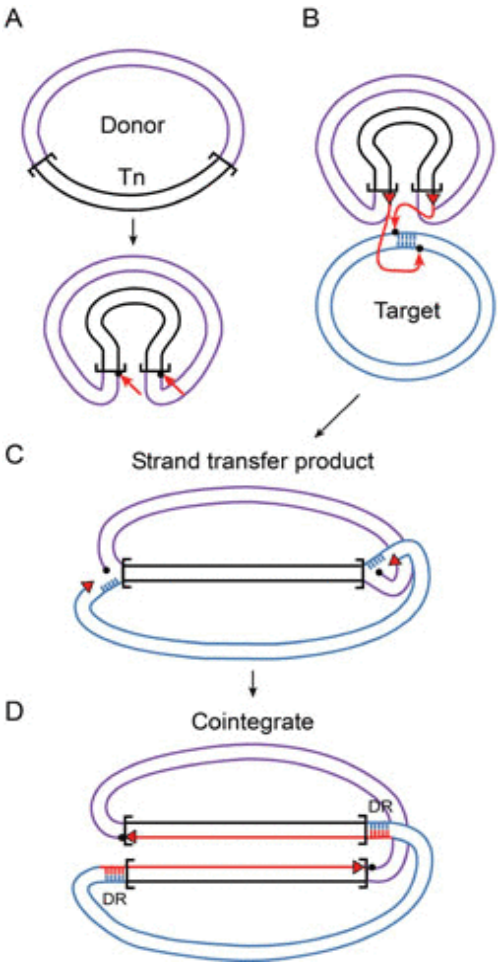


Figure I.1.2-1 - Model for the "paste-and-copy" transposition mechanism. Transposition starts with the nicking and release of 3′ OH at both ends of the transposon by the transposase (TnpA) (**A**). 3′ OH ends then make a nucleophilic attack in phosphodiester bonds of both target DNA strands (**B**). Host cell replication machinery synthesize complementary strands of the transposon to form a cointegrate (**D**). From Nicolas et al. 2015.

I.1.3 Tn3-family of DDE/D transposons

This master thesis work focuses on Tn4430 which belongs to the Tn3-family of DDE/D transposons. Tn3-family will therefore be further described in this section.

Among replicative transposons, the Tn3-family emerges as a vast and extensively distributed group of transposable elements, with representatives found across nearly every bacterial phylum (Nicolas et al. 2015). These transposons are characterized by the presence of Inverted Repeat (IRs) sequences (typically of 38 bp) situated at both ends of the transposon that play a pivotal role in the transposition process by facilitating the recognition and binding of the transposase. They also share an unusually lengthy transposase, often exceeding 900 amino acids, which stands in contrast to the transposases found in other well-studied models like Mu (663 aa) or Tn7 (702 aa) (Nicolas et al. 2015). The remarkable length of these proteins might be linked to the multifaceted roles they assume during replicative transposition. In fact, TnpA of Tn4430 plays a role in the contact and recognition of transposon ends, the interaction with target DNA, the assembly of the transpososome, as well as management of target immunity mechanisms (Lambin et al. 2012). This latter characteristic is also present in other members of the Tn3-family and corresponds to a process that prevents multiple insertions into the same target DNA molecule. Target "immunity" is conferred by the IRs and is specific to each element of the family indicating the TnpA is implicated in the process.

In opposition to Tn7 and Mu where transposition steps are carried by multiple proteins, Tn3-family has only one transposase protein which catalyze every step of transposition. While Tn3 members exhibit notable consistency concerning their transposition functions, they display substantial diversity in the passenger genes they carry. Most notably, they have gained notoriety for their pivotal role in dissemination of multidrug resistance within bacterial populations (Pasquali et al. 2005; Tian et al. 2023).

Within this family of transposons, transposases exhibit significant divergence from one another, as elucidated through phylogenetic analysis (**Figure I.1.3-1**), where seven distinct subgroups have been identified. Our model transposon is the Tn4430, which is now considered as a paradigm for understanding transposition mechanism of the Tn3-family.

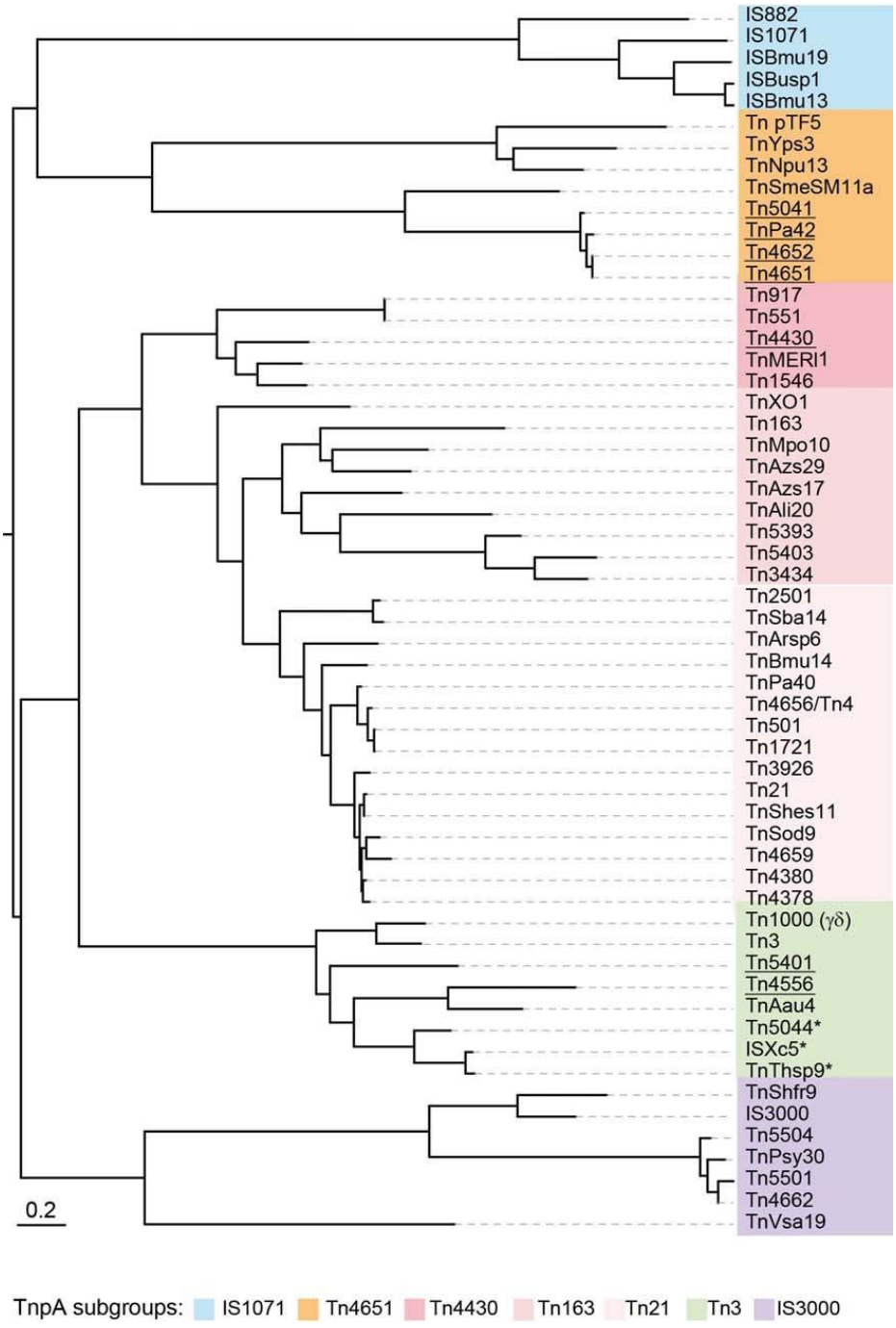


Figure I.1.3-1 - Phylogenetic tree of the Tn3-family transposase proteins. The different clusters and subgroups identified within the family are boxed with different colors as indicated. Transposons that contain a cointegrate resolution module working with a tyrosine recombinase are underlined. Transposons that encode a "long", C-terminally extended resolvase of the serine-recombinase family are marked with an asterisk. The length of the branches is proportional to the average number of substitutions per residue. Figure from Nicolas et al. 2015.

I.1.4 Tn4430 as a model for Tn3-family of transposons

Tn4430 is a 4149-bp transposon belonging to the Tn-3 family. It was initially isolated from *Bacillus thuringiensis* but is functional in *Escherichia coli* (Lereclus et al. 1986; Mahillon and Lereclus 1988). It is characterized by two 38-bp inverted repeats (IRs) and encodes two proteins: a DDE/D transposase (987 amino acids) (**Figure I.1.4-1**) and a site-specific tyrosine recombinase (285 amino acids)(Mahillon and Lereclus 1988). With no identified passenger genes, it is classified as a "cryptic" transposon.

Recent advances in understanding the transposition mechanism of the Tn3 family have been made possible by identifying deregulated TnpA mutants activated in transposition activity (T^+) and impaired in target immunity (I^-) (Lambin et al. 2012). These deregulated mutants were compaired with wild-type TnpA ($TnpA^{WT}$) and *in vitro* experiment showed that they assembled in a paired-end complex (PEC) with two IR ends in opposition to $TnpA^{WT}$ which assembled in a single-end complex (SEC). Further experiments showed that formation of PEC correlated with the activation on DNA cleavage at transposon ends (Fernandez et al. 2023). These data, with the fact that $TnpA^{WT}$ only forms PEC in presence of pre-cleaved transposon ends suggest that TnpA activation involves a conformational change and promotes the formation of the paired-end complex. Formation of PEC with mutants impaired in target immunity raise the question of target DNA and/or host factors role in the assembly of the active transposition complex.

These results were supported by the recent resolution of TnpA structure by cryo-electron microscopy (cryo-EM) (Shkumatov et al. 2022) in its apoform (**Figure I.1.4-1B**) and bound to transposon ends. The structural insights reveal that TnpA forms a compact dimer in its apoform and undergoes a conformational change upon binding to its IR ends. The resulting paired-end complex, exhibits activity in DNA cleavage and strand transfer (Nicolas et al. 2017a). In this process, one subunit of the TnpA dimer binds to the specific DNA sequence of the IR, while the other subunit catalyzes the cleavage reaction, releasing a 3'OH group. This mechanism, common among DDE/D-family transposases, facilitates a coordinated catalysis of both ends.

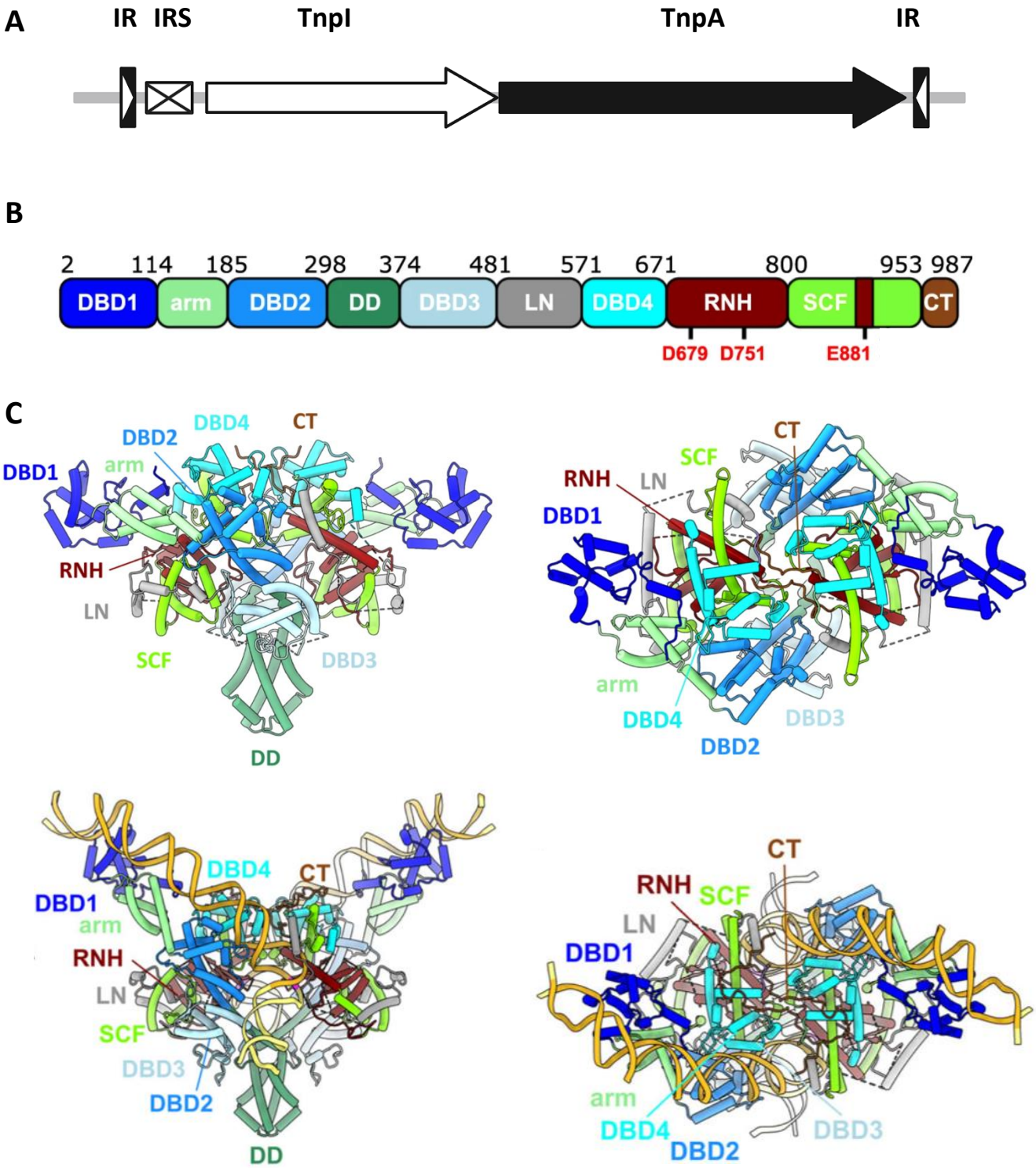


Figure I.1.4-1 - Structure and functions of Tn4430 and its transposase. A Schematic structure of Tn4430 delineated by 38bp IRs surrounding the internal recombination site (IRS) and genes coding for TnpI and TnpA. **B** Schematic subdivision of TnpA into structural domains. DBD1-4 DNA binding domains 1–4, DD dimerization domain, LN linker domain, RNH RNase H-like catalytic domain, SCF scaffold domain, CT C-terminal tail. Positions of the catalytic DDE triad within the RNAse H fold (RNH) indicated (red). **C** Structure of $TnpA^{WT}$ in apo conformation shown in cartoon representation and colored by structural domains. **C** Structure of $TnpA^{S911R}$ complex with two DNA molecules containing the transposon end in cartoon representation and colored by structural domains. **B, C** and **D** From Shkumatov et al., 2022.

Further investigations both *in vitro* and *in vivo* have uncovered a strong link between Tn4430 targeting and DNA replication. *In vitro* studies have underscored the robust affinity of TnpA for DNA substrates mimicking replication forks and *in vivo* experiments have demonstrated that Tn4430 cannot insert into non-replicating targets, highlighting the critical dependence on target replication for successful transposition (Nicolas E. PhD thesis, Oger C. PhD thesis; Nicolas et al., in prep.). Consistently with these results, data have evidenced a predilection for Tn4430 insertion within chromosomal regions where replication forks meet (Udekem d'Acoz, O. PhD thesis).

These findings challenge the previously accepted model of replicative transposition within the Tn3-family. Until recently, these transposons were believed to share a replicative transposition mechanism similar to bacteriophage Mu, which relies on a "replication hiring" mechanism. To transpose, Mu requires two self-encoded proteins, MuA and MuB. Like TnpA and Tn4430, MuA first binds to both ends of Mu genome. For transpososome assembly, Mu requires two *E. coli* proteins: IHF and HU which binds specific DNA regions of Mu sequence. The second self-encoded protein, MuB plays roles in all stages of transposition but play an essential role in target capture. Replication is recruited after strand transfer and destabilization of the transpososome complex by a molecular chaperone (ClpX). The destabilized ClpX-transpososome complex is then disassembled and exchanged for the translation initiation factor 2 (IF2). Finally, replication restart proteins promote the binding of replicative helicase (DnaB) and the assembly of DNA polymerase III holoenzyme to restart the replisome (Harshey 2014). In contrast, based on the results, a novel transposition mechanism termed "replication hijacking" has been proposed for Tn4430 and other members of the Tn3 family. In opposition to Mu, where replication machinery recruitment occurs post-insertion of the transposon, this new model postulates a mechanism where the element directly inserts into replicating forks thereby recruiting the host's replication machinery at the time of integration into the target.

I.1.5 Host factors and processes in replicative transposition

The replication “hijacking” mechanism suggests functional and/or physical interactions between the transpososome and the host replication machinery. The possibility of such interaction is supported by several studies on other transposable elements that have demonstrated interactions between transposition and host cellular processes, especially DNA replication. This can be illustrated by the replication machinery recruitment by bacteriophage Mu (Harshey 2014), the requirement of both IHF and HU host proteins (see above).

Another example is the Tn7 bacterial transposon which has been reported to recognize structures associated with DNA replication. Tn7 uses a self-encoded protein called TnsE. This protein was shown to be a DNA-binding protein and TnsE mutants with increased DNA-binding properties led to higher transposition frequencies. Like results obtained in our laboratory with TnpA (Nicolas et al. 2017b), TnsE binds fork-like DNA structures with high affinity. In addition, Tn7 insertion into *E. coli* chromosome occur in the terminal region which was also observed with Tn4430 (Peters and Craig 2000, 2001). One of the hypotheses proposed to explain this specific targeting is that replication structures could become accessible when the progression of a replication fork is prohibited in the terminus region (Peters and Craig 2001).

IS608 uses single-stranded DNA (ssDNA) transposition intermediates. Several cellular processes generate or use ssDNA like DNA replication, conjugative plasmid transfer, transformation, or DNA repair. This link between IS608 transposition and ssDNA availability was investigated and results showed that transposition of IS608 takes advantage of the presence of ssDNA on the template strand, used for Okazaki fragment synthesis (Ton-Hoang et al. 2010).

This master thesis seeks to delve deeper into the "replication hijacking" model and to identify potential cellular partners of TnpA. Since precise attributes of these potential interactions remain undisclosed, encompassing the possibility of both functional and physical interactions, we developed a proximity labelling assay based on a novel method described by Varnaitė, R. and MacNeill, S. (2016).

I.2 Proximity labelling

Protein-protein interactions (PPIs) play a central role in the regulation of cellular functions. Various methods have been employed to investigate these potential molecular interactions. In our quest to identify potential transposition partners of TnpA during replicative transposition, we have chosen to adopt an innovative approach known as proximity labeling (Varnaitė and MacNeill 2016). This cutting-edge technique offers distinct advantages that distinguish it from traditional methods such as co-immunoprecipitation (Co-IP), affinity purification (AP), and yeast-two-hybrid (Y2H) assay. Traditional methods, while valuable, often fall short in capturing weak and/or transient interactions and hardly allow to work in physiological conditions (Qin et al. 2021a). For instance, Co-IP or AP-based methods require the lysis of cells to release the bait protein, potentially compromising the preservation of physiological conditions and *in vivo* interaction status (Liu et al. 2020; Qin et al. 2021b). Meanwhile, the yeast-two-hybrid approaches are best suited for either testing suspected interactions or for large scale interactomics studies.

In contrast, the proximity labelling approach offers the unique advantage of probing interactions within their native cellular context and detecting interactions between molecular processes without necessarily involving physical interactions (Branon et al. 2018a; Kido et al. 2020a; Qin et al. 2021a; Samavarchi-Tehrani, Samson, and Gingras 2020; Santin et al. 2018). This technique allows for the investigation of even weak and transient interactions that might elude detection by more traditional methods (**Figure I.2-1**). The main drawback that comes with this approach is the non-specific probing of "random" proteins. For this reason, proteins identified by proximity labelling must be validated with complementary approaches.

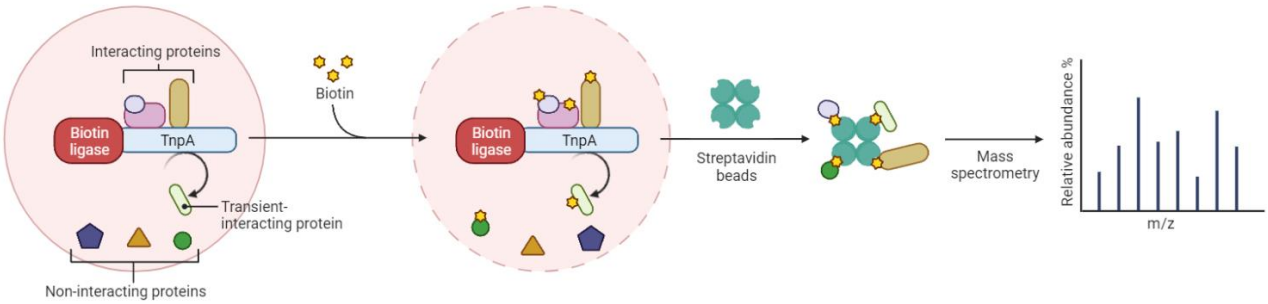


Figure I.2-1 - Proximity labelling workflow. Schematic representation of the proximity labeling approach and workflow. A labeling enzyme (biotin ligase) is fused to the protein of interest (here, TnpA). The biotin ligase catalyzes the synthesis of a reactive biotin intermediate which diffuses and labels neighboring proteins by reacting with primary amines (lysine residues). Biotinylated proteins are captured with streptavidin beads, washed, and then analyzed by mass spectroscopy.

In recent years, enzyme-catalyzed proximity labeling has emerged as a groundbreaking strategy for labeling, capturing, and identifying proteins that directly interact with the protein of interest (POI). Within this approach, two distinct systems have gained prominence: the APEX system and the BioID system. Each of these systems offers unique advantages and operates through different biochemical mechanisms, providing researchers with versatile options for investigating proximity-based interactions.

The APEX system features a monomeric 28 kDa ascorbate peroxidase, known for its ability to catalyze the conversion of exogenous biotin-phenol in the presence of hydrogen peroxide (Kalocsay 2019; Xu, Fan, and Hu 2021). This enzymatic reaction leads to the formation of a biotin-phenoxyl intermediate, a reactive species which forms covalent bonds with electron-rich amino acids, particularly tyrosine residues.

In contrast, the BioID system, introduced in 2012 (Roux et al. 2012), utilizes a mutated variant of the *Escherichia coli* biotin ligase BirA, with a molecular weight of 35 kDa. This engineered enzyme catalyzes the synthesis of biotinoyl-5'-AMP (bio-AMP) from biotin and ATP which is then released and reacts with vicinal lysine residues (**see Figure I.2-2**). BioID and the newly engineered biotin ligases (BioID2, miniTurbo, AirID) differ from the APEX system in several ways. Specifically, APEX requires the addition of hydrogen peroxide for instantaneous labeling, akin to capturing a snapshot of the interactions. In contrast, BioID requires only the non-toxic supply of biotin and labels proteins over more extended time frames.

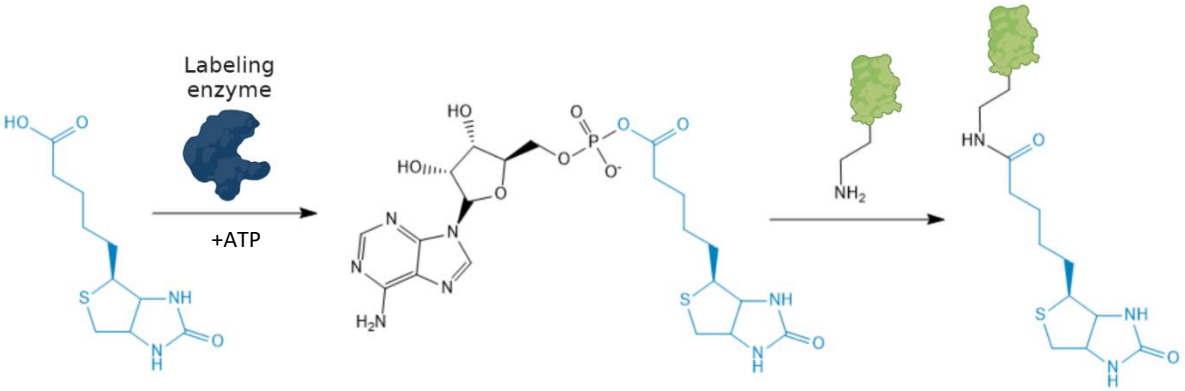


Figure I.2-2 - Labeling mechanism of biotin ligases. In presence of ATP, labeling enzymes (dark blue) synthesize a reactive intermediate of biotin (biotinyl-5'-AMP). This intermediate reacts with lysine residues of proteins located in a radius of 10-20 nm (green). After biotinylation, cell lysate is mixed with streptavidin beads, allowing for enrichment and identification of biotinylated proteins.

II. Objectives

Previous studies have highlighted a relationship between transposition mechanism of Tn4430 and the host cell replication machinery. These findings challenge the classical model of replicative transposition for Tn3-family members. Distinct from to the model based on the replicative transposition mechanism of bacteriophage Mu, a new "replisome hijacking" model was proposed in which the transposon directly inserts itself into replication intermediates rather than recruiting the replication machinery after strand transfer.

The aim of this master thesis is to provide more evidence to support this "replication hijacking" hypothesis by finding potential transposition interactants that could help to target replication intermediates in living cells. More generally, the proposed strategy should help to identify any host factor or cellular process that could contribute to Tn4430 transposition at any stage of the reaction, a topic that has been poorly investigated until now.

The first part of this work was to develop an experimental strategy aiming at coupling BioID-mediated proximity labelling with TnpA activity using a suicide plasmid-based bacterial conjugation assay. This assay was optimized so as to establish the TnpA peroxisome during a defined period of time within bacterial cells where Tn4430 transposition actually occurs and can be monitored.

The second part of this work was focused on verifying proteins identified as potential TnpA interactants in the proximity labelling assay. Since this assay yields a proxisome of the protein of interest, candidate proteins that directly interact with TnpA will be validated through bacterial-two-hybrid experiments.

III. Results

III.1 Rationale of the experimental strategy: coupling TnpA proximity labelling with Tn4430 transposition activity

Since transposition is a rare event occurring sporadically over time in a given bacterial population, we required a method capable of effectively tagging TnpA host cell partners over an extended period. This motivated us to use a mating out assay with the possibility performing labelling in living cells during an extended period where transposition events could be measured. In the same context, with the aim of studying TnpA interactions with rare transposition events, the APEX system providing a snapshot of molecular interactions at a specific moment seemed less suitable for our needs. Conversely, the BioID system aligned more closely with our experimental requirements. Its temporal flexibility should allow for the gradual accumulation of proximity-dependent interactions, a feature particularly valuable when investigating rare events that happen linearly over time like replicative transposition.

To allow proximity labelling *in vivo* within a given period and controlled environment, we used a suicide plasmid carrying a transposon or mini-transposon along with a selectable marker is introduced into recipient cells (Figure **III.1-1**) (Demarre et al. 2005; Ferrières et al. 2010). The selectable marker enables the selective proliferation of cells that have undergone transposition events. This feature facilitates the quantification of transposition rates, making suicide conjugation assays invaluable for exploring transposon dynamics within a cell population. This mating out assay allows to perform proximity labelling *in vivo* within a given period where transposition occurs and can be quantified.

In more details, the assay relies on a donor strain (DAP-/Pir+) containing a suicide transposition vector designed to conjugate with the recipient strain (Pir-) but unable to replicate within it. Donor plasmid includes the fusion of TnpA with biotin ligase or the biotin ligase alone (used as a negative control) and is accompanied by a mini-transposon cassette housing the Kan^R gene, serving to quantify transposition frequency. Furthermore, the plasmid features an oriVR6Kgamma origin of replication, making its replication dependent on the pi protein.

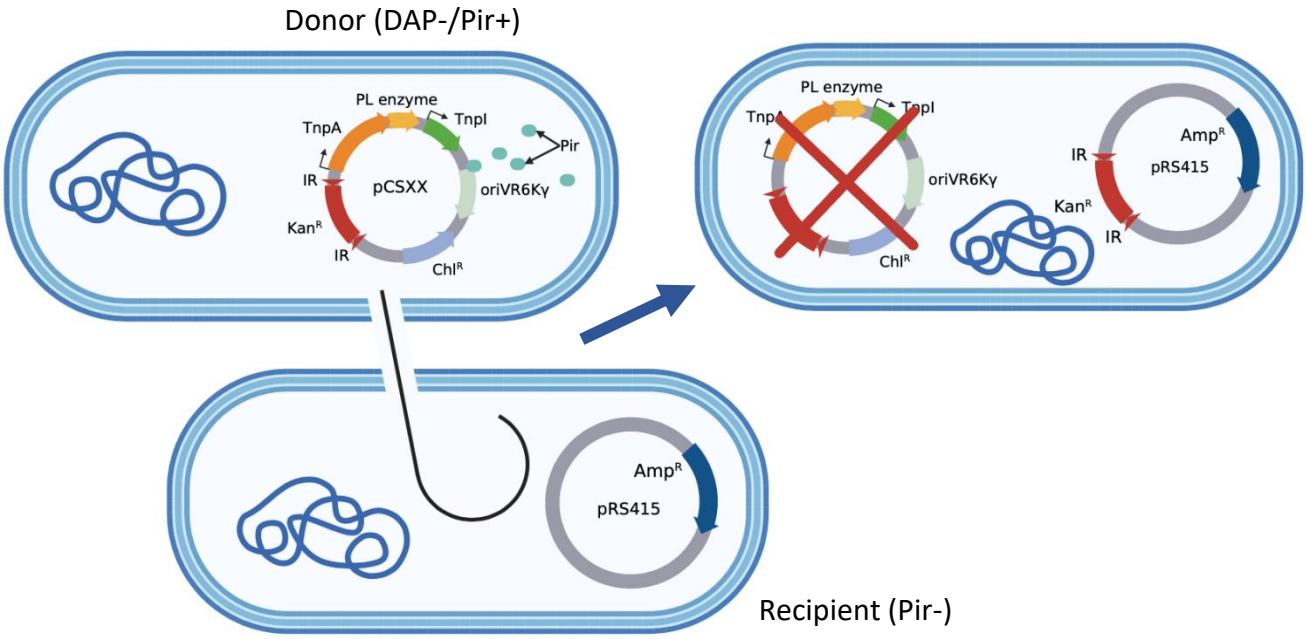


Figure III.1-1 - Conjugation assay. Schematic representation of the conjugation assay used for the biotinylation experiment. Donor strain requires diaminopimelic acid for growth (DAP-) and expresses Pir, a protein that allows the cell to replicate the suicide, transposon donor plasmid (Pir+). Recipient strain does not express Pir (Pir-) and is thus not capable of replicating the suicide plasmid.

III.2 Proximity labelling approach

III.2.1 Choice of the BioID-based labelling method

As mentioned in the introduction, previous data from the lab have led to a "replication hijacking" model of replicative transposition. In this model, Tn4430 inserts into ongoing replicating fork, thus taking advantage of the presence of an active replication machinery on the target DNA to synchronize transposition with replicative duplication of the transposon. To test this hypothesis, as well as identifying any other host factors or processes, we used the proximity labelling approach, which allows biotinylation of proteins in proximity of a given POI. Proteins tagged by biotin can then be easily enriched using streptavidin beads and analyzed by mass spectrometry. It is important to note that proximity alone does not necessarily imply a direct interaction. Therefore, to validate potential interactants identified through proximity labeling, additional rigorous methods such as bacterial-two-hybrid assays will be employed. These supplementary assays will provide a more comprehensive and reliable assessment of the true interactions between TnpA and its proximal protein environment, ensuring the robustness of our findings and contributing to the overall depth of the study.

Here, we decided to test three different biotin ligases (TurboID, miniTurbo and AirID (Branon et al. 2018a; Kido et al. 2020)) with variable labeling times (ranging from a few minutes to multiple hours), affinities, and sizes (27-35 kDa). This decision stemmed from the need to comprehensively assess the most suitable system for our specific research objectives.

III.2.2 Genetic constructs for the coupled transposition-proximity labelling assay

The proximity labelling assay hinges on the fusion of a BirA derived biotin ligase (BirA*) to the POI TnpA. To discern genuine potential interactants of TnpA from other cellular proteins, we constructed a negative control comprising the BirA* gene alone. This control is necessary to discriminate between the overall cellular background of proteins that can

be labelled by a particular BirA* derivative from the specific set of proteins that are labelled by the TnpA-BirA* fusion. The TnpA-BirA* and the BirA*only constructs were separately integrated into a conjugative donor plasmid, housing a mini-transposon cassette (Mini-TnKm). This mobile cassette included a kanamycin resistance gene (Km^R) flanked by a pair of Tn4430 inverted repeats (IR ends). The donor plasmids also contained the gene for the resolvase (TnpI, a chloramphenicol resistance gene (Chl^R), and a Pir-dependent origin of replication from R6K plamid (Figure **III.2.1-1**). TnpI and TnpA were expressed at low level by using leaky expression from the *E. coli* pLac promotor to encourage specific TnpA localization within the cells like observed in previous fluorescence microscopy experiments (d'Udekem d'Acoz O., PhD thesis, 2022).

The plasmids containing the TnpA-enzyme and enzyme-only constructs were created using Gibson assembly. TnpA-enzyme and enzyme-only constructs were already available for TurboID and mini-Turbo but new constructions we made for AirID, Post-assembly, plasmids were transformed into the MFDPir strain and rigorously validated by restriction analysis and sequencing.

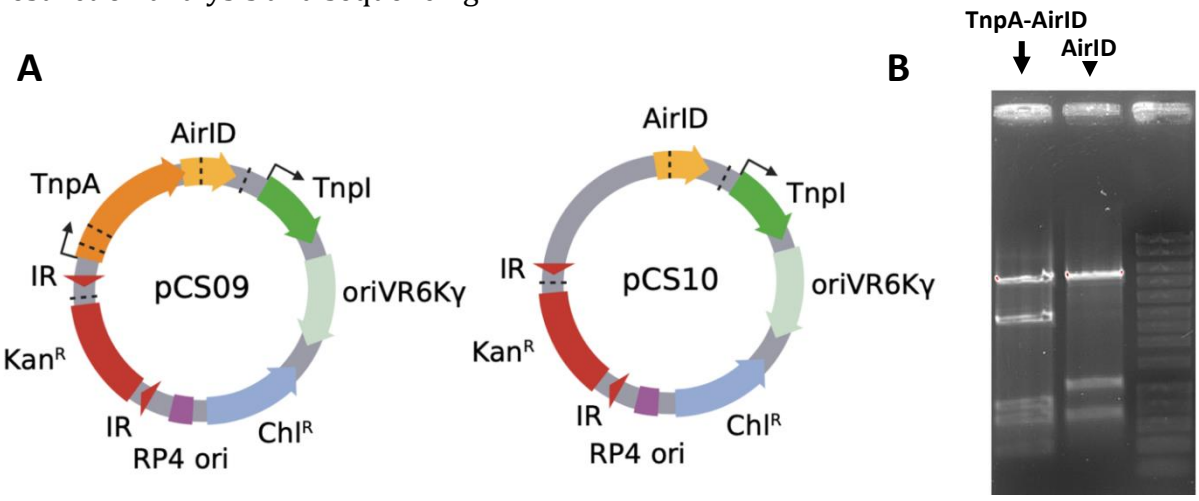


Figure III.2.1-1 - Genetic constructs based on AiID-derived BioID ligase. (A) Schematic representation of genetic constructs. TnpA-AirID serves as the test condition and contains AirID fused to the C-terminal end of the TnpA coding sequence, while the AirID coding plamid was used as a negative control to express the AirID moiety alone. Both constructs include a Mini-TnKm mobile cassette carrying KanR gene, a chloramphenicol resistance gene ChlR, the RP4 origin of conjugative transfer (RP4oriT), and a Pir-dependent origin of replication (oriVR6Kγ). Both TnpA and TnpI are under the expression of the E. coli pLac promoter. The dashed lines represent the AvaII restriction sites that were used for construct verification. Upon AvaII restriction, TnpA-AirID yields five fragments of sizes 5178, 2827, 840, 702, and 367, whereas AirID-only generates three fragments of 5178, 1073, and 702 base pairs. **(B)** 0.8% agarose gel loaded with AvaII restricted pCS09 and pCS10.

III.3 Development and optimization of the coupled BioID labelling-transposition assay

As detailed above, to maintain precise control over the transposition environment, transposition frequency, and the initial timing of the proximity labelling, we developed a mating out assay. This assay involves a donor strain (MFDpir) diaminopimelic acid auxotroph (DAP-) expressing Pi protein (Pir+) and carrying a suicide donor plasmid based on the Pi dependent origin of replication (oriVR6Kγ). This specific origin of replication allows replication of the donor plasmid in strains expressing the Pi protein only, which is not the case of the recipient strain (Pir-) harboring a transposition receiver plasmid (pRS415, **Figure III.1-1**). The suicide plasmid contains the TnpA-BirA* fusion or the BirA* alone, along with a mini-transposon cassette housing the Kan^R gene. Consequently, the insertion of this mini-transposon into the target strain confers Kanamycin resistance, enabling the quantification of transposition frequency through quantification of Kan^R transconjugants on kanamycin-containing LB plates.

III.3.1 Overnight growth of donor cells gives a strong background of endogenous protein biotinylation

In the initial set of experiments, MFDPir donor strains containing either the TnpA-BirA* plasmid or the BirA* alone construct were cultured overnight prior being harvested and mixed on biotin-containing LB plates with recipient MG1655 cells containing a 10-kb ampicillin resistant (Amp^R) ColE1-derived plasmid (pRS415) serving as transposition receiver (**Figure III.1-1**). Following the conjugation period, cells were resuspended and used for western-blot analysis using alkaline phosphatase-conjugated streptavidin to detect biotinylated proteins.

The conjugation assay was conducted using the three distinct biotin ligases (TurboID, miniTurbo, and AirID), with conjugation periods of 0, 90, and 180 min, in the presence of 50 µM biotin. **Figure III.3.1-1** shows that a notable background signal was present for all three biotin ligases at the 0-min time point. This indicates that the biotinylation signal originated from the donor cells, in which the BioID-derived enzymes used endogenous

biotin as a substrate to biotinylate a range of proteins during overnight growth preceding conjugation. Supporting this interpretation, the background signal was strongly reduced when recipient cells alone were spotted on the conjugation plates and treated in the same way as the mating mix (**Figure III.3.1-1**). In this case, endogenous biotinylation yielded a major band corresponding to biotin carboxyl carrier protein (BCCP, 22.5 kDa), the natural substrate of BirA, as expected from previously published biotinylation profiles in *E. coli* (Santin et al. 2018).

Unlike to previous reports (Kido et al. 2020b), expression of AirID in the donor cells did not exhibit a significantly reduced background signal when compared to TurboID and miniTurbo, presumably because the three ligases reached the same plateau after overnight growth despite their different specific activities. Together, these results prompted us to reduce the contribution of donor cells to the biotinylation pattern obtained after conjugation.

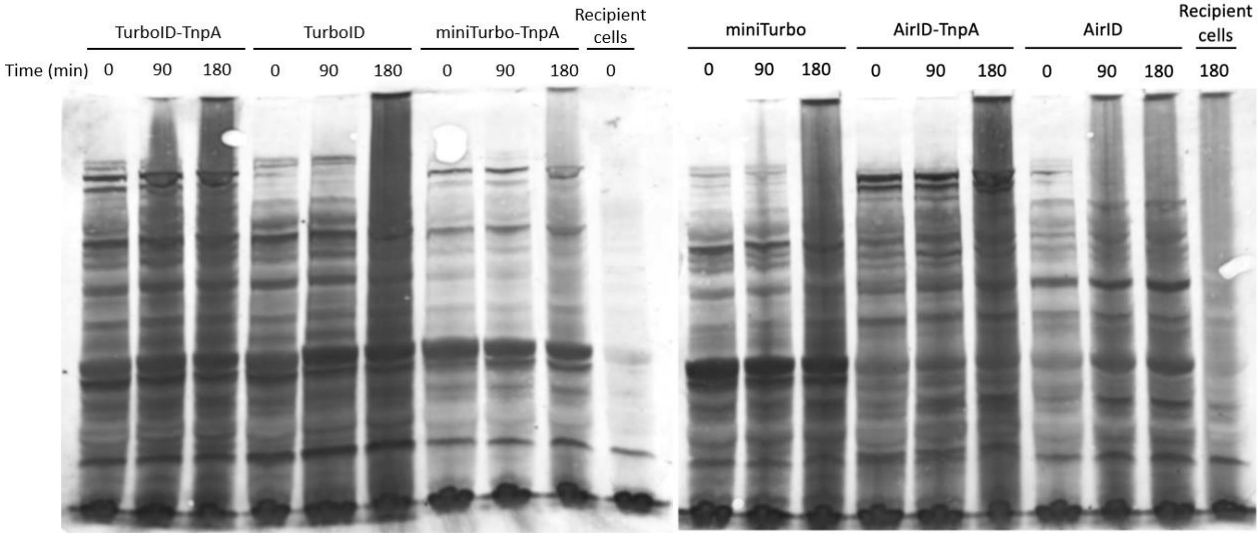


Figure III.3.1-1 – Biotinylation profiles reveal high background signal. Western blot analysis of biotinylated proteins after 0, 90 and 180 min of conjugation. Lysates were run on SDS-PAGE and analyzed with AP Streptavidin. Experiment was conducted for three biotin ligases (TurboID, miniTurbo and AirID), with test (TnpA fused with the labeling enzyme) and negative control (labeling enzyme only) conditions. Outer right lanes correspond to recipient cells only, after 0 and 180 min of incubation on the conjugation filter under the same conditions as the conjugation mix. Donor/recipient cell ratio used for the conjugation was 1:2.

III.3.2 Biotinylation background is reduced by reducing donor to recipient cells ratio and by using ampicillin to eliminate the donor from the conjugation mix

To mitigate donor cell signal interference, conjugation was performed with a modified donor:recipient cell ratio, reduced from 1:2 to 1:10 when compared to previous experiments. As depicted in **Figure III.3.3-1**, a significant increase in global biotinylation intensity was observed between 0 and 90 min of conjugation, which was not the case at lower dilution of the donor cells in the conjugation mix (compare **III.3.1-1** and **III.3.2-1**). This suggests that a major part of detected biotynilation occurred after triggering conjugation. However, it was difficult to determine whether the observed increase in biotinylation signal resulted from the recipient cells or from donor growth within the conjugation mix.

To further reduce the contribution of the donor cells in the conjugation-specific biotinylation pattern, and thus to enrich or restrict TnpA-mediated proximity labelling recovery in the recipient cells, we attempted to combine conjugation with ampicillin treatment in order to remove the donor cells from the conjugation mix.

In prevision to this approach, we performed a trial experiment to test the sensitivity of donor cells to ampicillin under the condition used for the mating out assay (Figure **III.3.3-1**). Overnight-grown donor cells were pelleted and spotted on filters placed on a petri dish with or without 250 µg/mL of ampicillin. At given time points, filters were removed, and cells were resuspended in fresh LB medium. Optical density was measured and plotted as a function of time (Figure **III.3.3-2A**). The results showed that ampicillin-treated donor cells underwent lysis after 90 min, whereas untreated cells continued to grow before declining due to the absence of diaminopimelic acid in the conjugation plates (Figure **III.3.2-2A**). Western blot detection of biotinylated proteins in donor cell extracts showed a strong reduction of the signals between 90 min and 150 min of incubation in the presence of ampicillin, aligning with the observed lysis pattern of donor cells (Figure **III.3.2-2B**).

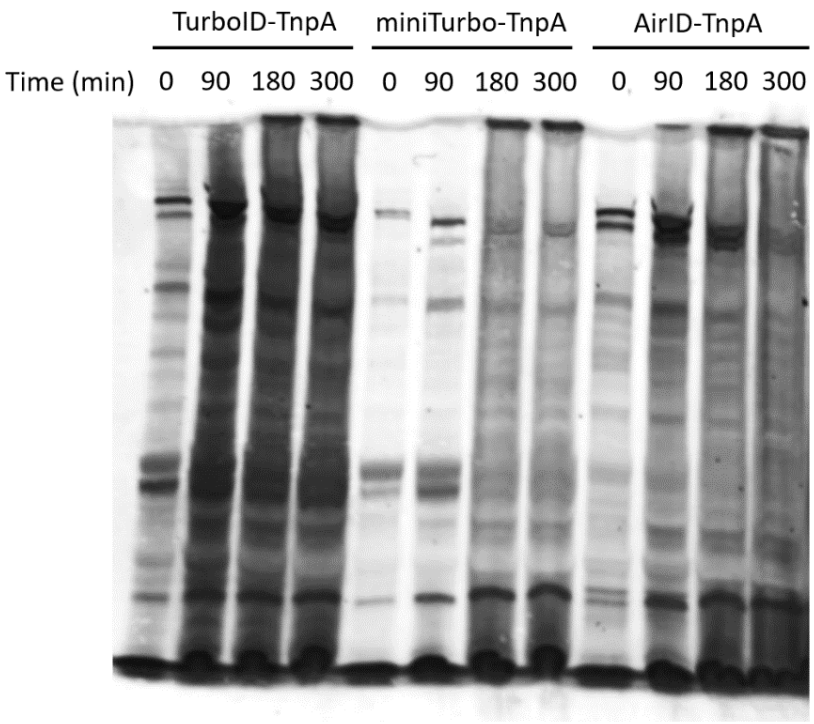


Figure III.3.2-1 – Cell ratio and ampicillin reduces overnight signal. Western blot analysis of biotinylated proteins from 0 to 300 min of conjugation. Lysates were run on SDS-PAGE and revealed with AP Streptavidin. Donor/recipient cell ratio used for the conjugation was 1:

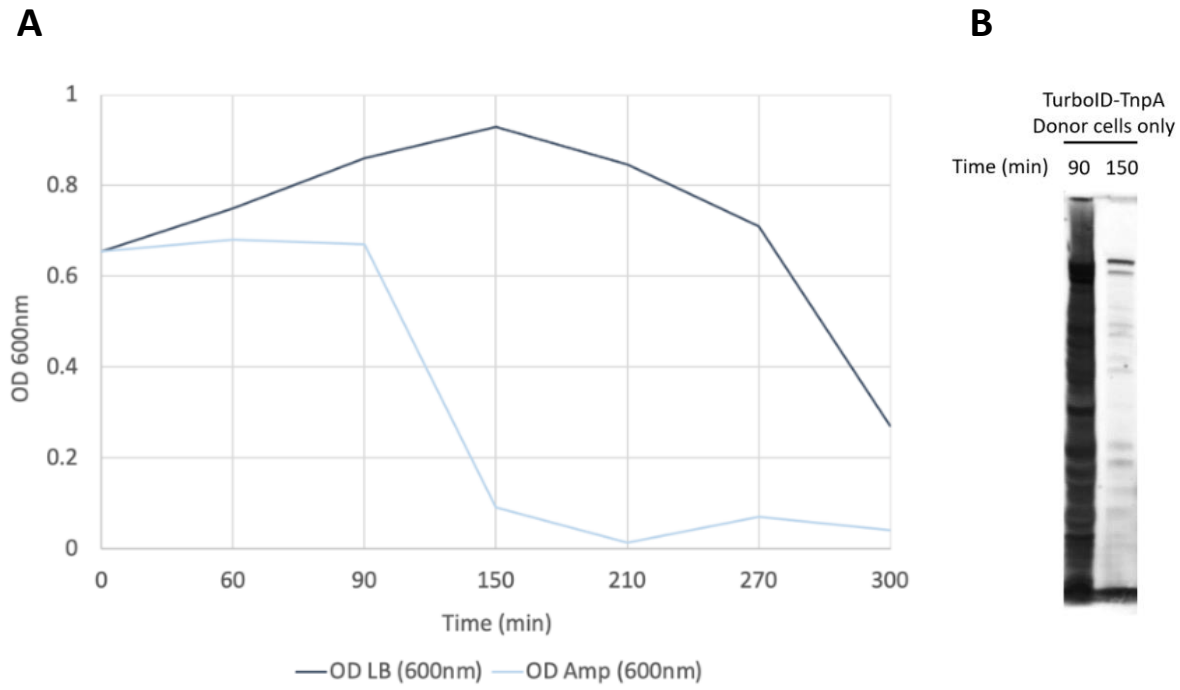


Figure III.3.3-2 – Cell ratio and ampicillin reduces overnight signal. (A) Western blot analysis of biotinylated proteins in donor cells during the ampicillin resistance test (90 and 150 min). Lysates were run on SDS-PAGE and revealed with AP Streptavidin. **(B)** Ampicillin sensibility of donor cells only was tested in conjugation conditions. Donor cells were pelleted and spotted on a filter placed on a petri dish with 250µg/mL ampicillin. At given time point, filter was removed, and cells were resuspended in 2 mL of LB. Optical density was measured and plotted as a function of time. Control SDS in supplementary figures (**see S1**).

Based on these results, mating out assays involving the three different variants of TnpA-BirA* fusions and their respective controls were repeated by adding ampicillin at different stages of conjugation (Supplementary Figure **III.3.3-1**). Western blot analysis confirmed that for the three sets of constructs, the protein biotinylation signal increased during conjugation as was observed in mating experiments where donor cells where simply diluted when compared to the recipient cells (Supplementary Figure **III.3.3-1A**, compare with Figure **III.3.3-1**). However, when the donor cells alone were treated under the same conditions as the conjugation mixes, the biotinylation signal recovered after 150 min of incubation of the filters was strongly reduced compared to that observed after the same period of conjugation between the donor and recipient cells (Supplementary Figure **III.3.3-1A**, compare left panels of the gel with the corresponding 6 right-most lanes). This indicates that the donor cells were efficiently eliminated from the conjugation mix under the different conditions tested, and that the detected biotinylation activity came from the recipient cells into which the donor plasmids have transferred by conjugation giving specific patterns for the different constructs. Intriguingly, for each of them, the biotinylation profile obtained after 150 min of incubation on the conjugation plate was essentially the same under the two conditions tested (Supplementary Figure **III.3.3-1A**, condition 1 or 2), suggesting that directly treating the conjugation mix with ampicillin or not did not markedly affect the efficiency of plasmid transfer and subsequent expression of the corresponding biotin ligase.

III.3.3 Fusion proteins of TnpA are active in transposition

Transposition frequencies were evaluated to assess potential impairment of transposition activity resulting from the fusion of biotin ligases to TnpA. For this purpose, precise dilutions of cell suspensions obtained from each mating were plated on selective LB plates. The rate of mini-TnKm transposition was determined by dividing the number of doubly resistant Amp^R-Km^R transconjugants by the total number of donor cells (Cm^R) in the sample, as outlined in the Material and Methods section. As explained previously, we normalized transposition frequencies using the number of donor cells. This adjustment was implemented after analyzing the data presented in Supplementary Figure **III.3.3-1B**, where transposition frequencies were found to be underestimated when normalized by the number of recipient cells, which are in excess under the conjugation conditions.

Similar to the results depicted in Supplementary Figure **III.2.3-1B**, elevated transposition rates were observed, even in negative controls lacking TnpA (uncorrected transposition frequency, represented in light blue) (Figure **III.3.3-1**). This, along with previous findings, prompted us to refine transposition frequencies by assessing events leading to chloramphenicol resistance in recipient cells. As events leading to chloramphenicol resistance are not indicative of genuine transposition events, we measured and subtracted them from the frequency of recipient cells acquiring only kanamycin resistance. Corrected transposition frequencies (depicted in dark blue) underscore a noticeable reduction in negative controls, a trend less pronounced for positive controls where the transposition frequency remains relatively consistent. Significantly, the fusion of TnpA-TurboID appears to exhibit reduced activity, comparable to TurboID alone, which cannot facilitate transposition due to the absence of TnpA. Consequently, TurboID could not be incorporated into our experiment due to its impact on the transposition activity of TnpA.

Moreover, the transposition activities measured with the three TnpA-BirA* fusions were not consistently significantly higher than the background levels observed with constructs containing the BirA* enzyme alone, which prompted us to conduct a molecular characterization of the selected transconjugants.

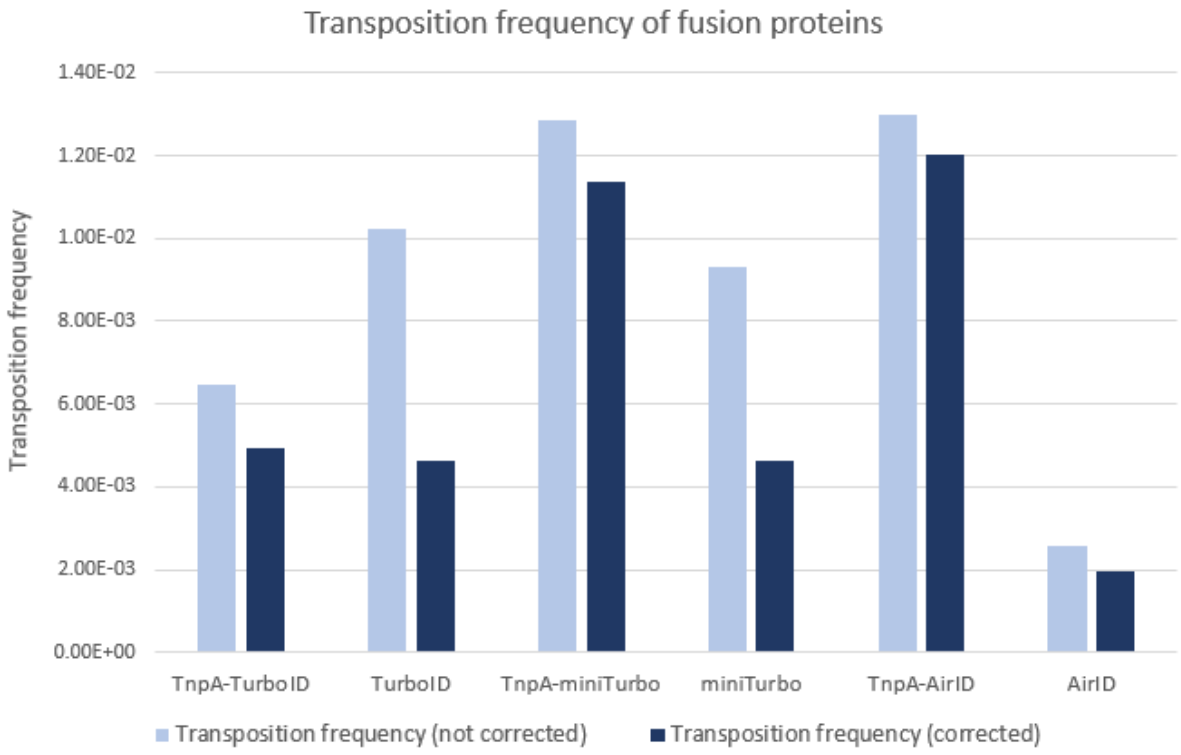


Figure III.3.3-1 – Fusion proteins are active in transposition. Transposition frequencies were calculated and plotted for every construction (positive and negative controls). Transposition frequencies were obtained by dividing the number of colonies growing overnight on kanamycin ampicillin LB plates after conjugation by colonies growing overnight on DAP chloramphenicol LB plates.

III.3.4 Molecular characterization of the transposition products

As detailed in section **III.3.3**, our focus shifted towards characterizing Kan^R/Amp^R transconjugants and analyzing transposition products. The anticipated products of transposition between the donor and target plasmid are illustrated in Figure **III.3.4-1B**, along with the depiction of the cointegrate intermediate. Although chromosomal DNA serves as a viable target for transposition, it is not explicitly represented in the figure. Furthermore, Figure **III.3.4-1C** illustrates potential transposition-independent recombination products between the donor and target plasmids, providing a plausible explanation for the observed gain of kanamycin resistance in recipient cells.

To unravel the transposition products resulting from the mating out assay, we decided to select multiple Amp^R/Kan^R transconjugants obtained with both fusion protein sand controls. These transconjugants were cultured on petri dishes containing either kanamycin and ampicillin or chloramphenicol and ampicillin to assess and evaluate their resistance phenotype. Notably, 44.1% (15/34) of transconjugants obtained with the TnpA-AirID fusion displayed resistance to chloramphenicol in addition to kanamycin. Conversely, all transconjugants obtained with AirID alone (34/34) exhibited resistance to both kanamycin and ampicillin. For a more comprehensive characterization of transposition products, some transconjugants were randomly selected, and their plasmids were individually extracted and back transformed into a TOP10 strain to ascertain the retention of antibiotic resistances. This approach enabled us to discern instances where transposition occurred into chromosomal DNA and cases where the donor plasmid might have been transiently maintained.

At the same time as transforming extracted plasmids into TOP10, isolated plasmids were also analyzed and cut with NcoI and KpnI. The mini-TnKm has two KpnI sites at the transposon ends, and NcoI sites are located outside the inverted repeat sequences (see Figure **III.3.4-1A**). Using these enzymes helps identify transposition events where the transposon and its IRS are duplicated, maintaining KpnI restriction sites but losing NcoI sites in the target. If transposition occurred, cutting the target DNA with KpnI should produce a band corresponding to the mini-TnKm, which is not the case with NcoI due to the absence of sites in the target (see Figure **III.3.4-1B**). This profile is illustrated in

Figure **III.3.4-2**, lanes 13-15. This figure represents restriction profiles of plasmids after extraction from transconjugants for TnpA-AirID (lanes 11-20) and AirID only (lanes 1-10). Five transconjugants for TnpA-AirID were resistant to kanamycin and ampicillin (lanes 11-15) and five others were resistant to chloramphenicol as well (lanes 16-20). As described before, profiles observed in lanes 13-15 (Figure **III.3.4-2**) correspond to transposition, suggesting that most events leading to kanamycin resistance in presence of the TnpA-AirID fusion are transposition-dependent.

As expected, chloramphenicol resistant TnpA-AirID transconjugants showed different restriction profiles. In fact, restriction by KnpI did not lead to the apparition of a band corresponding to the mini-Tn*Km.* Three transconjugants (Figure **III.3.4-2**, lanes 17, 19 and 20) showed bands from donor and receiver plasmids and kept their resistances after resistance back transformation. This could indicate some recombination events between the two plasmids, rescuing the donor plasmid (see Figure **III.3.4-1C**). The two other lanes (Figure **III.3.4-2,** 16 and 18) showed no resistances after back transformation (indicated with X), which could indicate similar recombination events with the chromosome rather than the receiver plasmid.

Similarly, restriction analysis of AirID only transconjugants showed two distinct profiles. First, lanes 4,6 and 10 lost their resistances upon back transformation and showed a restriction profile corresponding to the receiver plasmid only. These could indicate a recombination event with the chromosome. Second, plasmids restricted in remaining lanes led to the maintenance of resistances after back transformation, and consistently showed additional bands of various size as compared to the "receiver plasmid only" profile (Figure **III.3.4-2**, lanes 1-3, 5, 7-9). In these lanes, we can easly observe the band corresponding to mini-Tn*Km* when restricted by NcoI (less clear in restriction by KnpI). This result, taken with the fact that mini-Tn*Km* becomes associated with the recipient plasmid in absence of TnpA, suggests that a recombination event must have happened. In case of recombination, conservation of band **B** and/or apparition of a new band (near **A**) depends on the recombination site between the two plasmids (depicted in **III.3.4-1C**). Note that one of the band appearing has a similar size as to band **A** which is part of the TnpA-AirID construct but the absence of band **C (Figure III.3.4-2)** in the KpnI restriction of these colonies attests that it is not due to inversion of constructs.

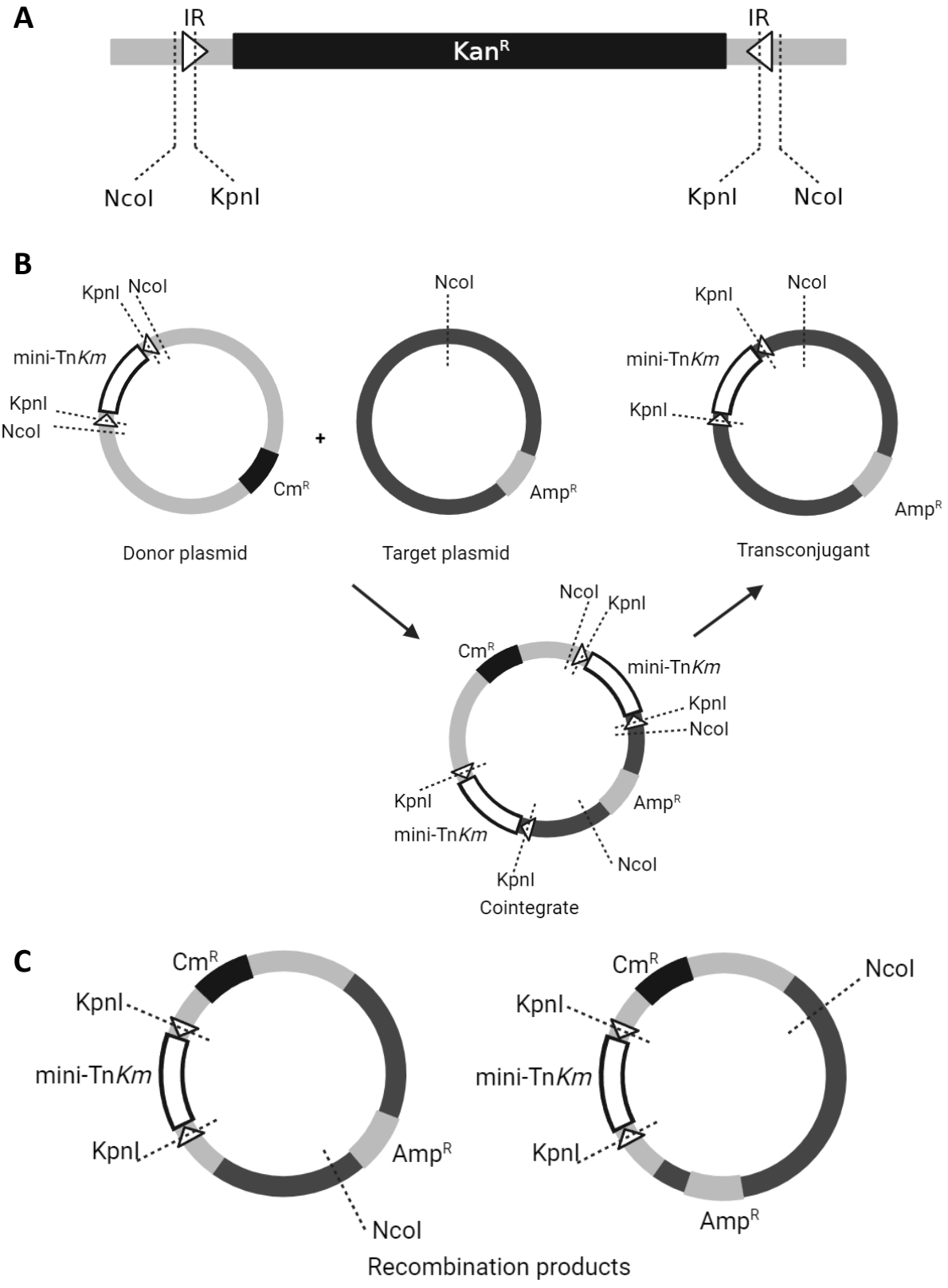


Figure III.3.4-1 – Schematic representation of the mini-Tn*Km*, transposition event and hypothetical recombination events. (A) Schematic representation of the mini-Tn*Km* with NcoI and KpnI restriction sitespresent in the donor plasmid. **(B)** Schematic representation of the donor and target plasmid with expected transposition product and the cointegrate intermediate. **(C)** Schematic representation of possible transposition-independent recombination products.

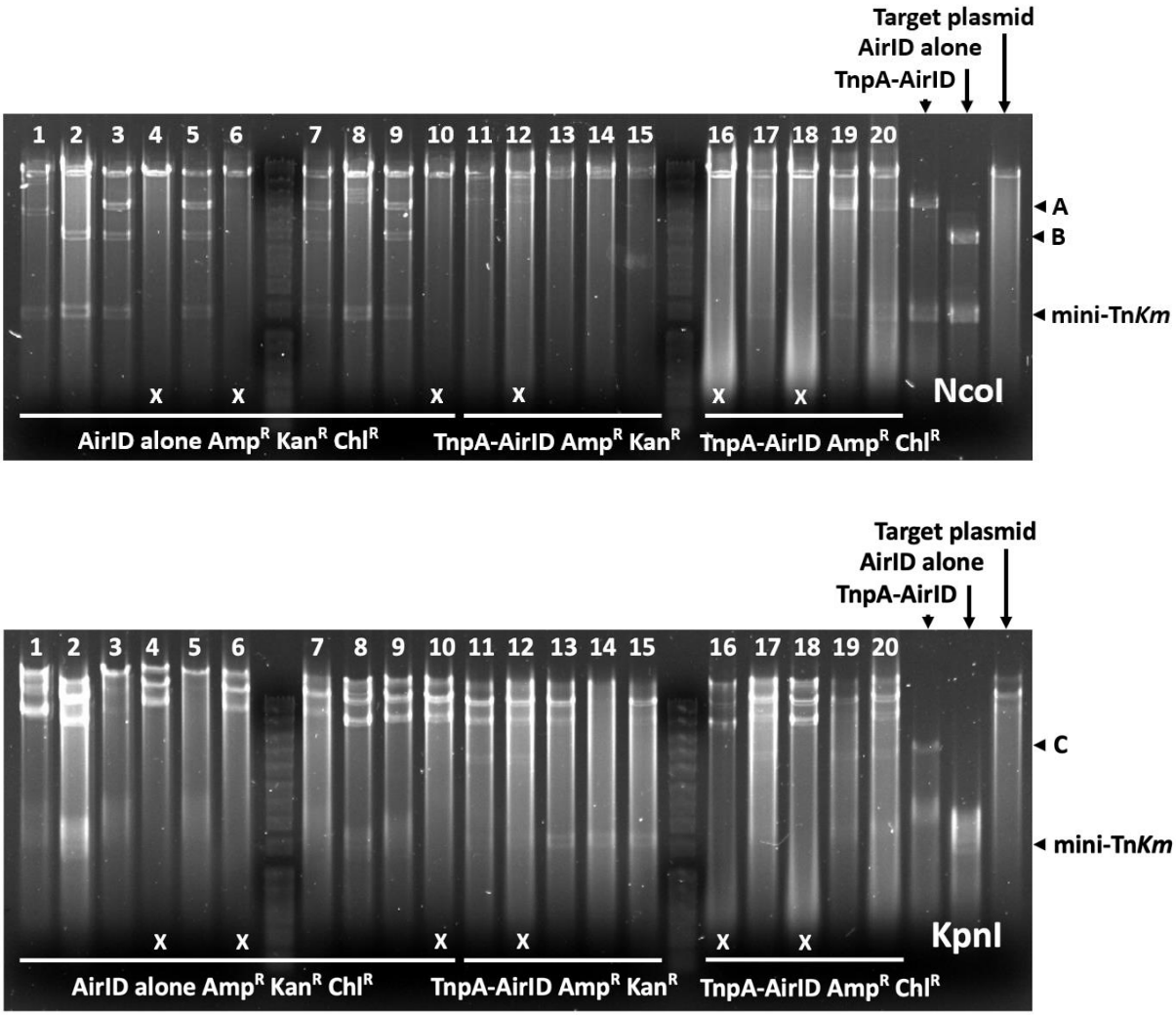


Figure III.3.4-2– Restriction analysis of plasmids from Amp^R/Kan^R transconjugants resulting from the mating out assay. Plasmid DNA from 20 AmpR-KanR transconjugants was isolated by back transformation and digested with NcoI (A) and KpnI (B). The digestion profiles of the TnpA-AirID, AirID alone and target plasmids is shown in the right-most lanes of the gels as indicated. Lower legend indicates from which donor plasmid construct the colonies originate and their resistance phenotype before back transformation. "X' indicates colonies that lost resistances after back transformation.

III.3.5 Testing the stability of proximity labelling constructs in *E. coli*

Prior to performing proximity labelling experiments, we tested the stability of the TnpA-BirA* fusion. Unstable constructs could potentially yield false positive, compromising the integrity of the entire experimental dataset. To rigorously assess this stability, we added a FLAG sequence in the C-terminal part of TnpA- AirID or AirID alone constructs using the Gibson assembly method (see Material and Methods). This sequence allowed us to perform a western blot analysis and to detect the expression of fusion proteins (Figure **III.3.5-1**).

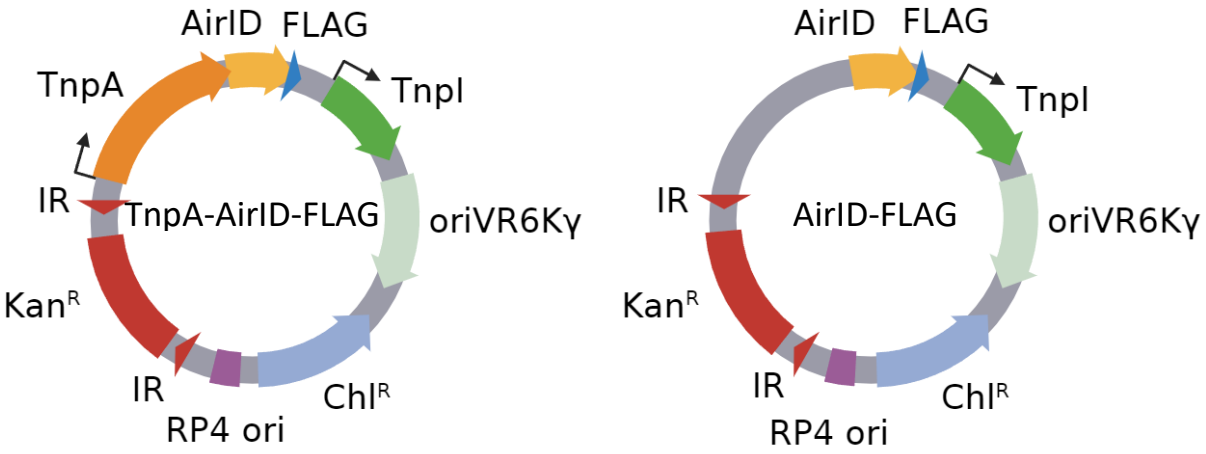


Figure III.3.5-1 - Genetic constructions. Schematic representation of genetic constructs. TnpA-AirID-FLAG serves as the test condition, while AirID-FLAG is the negative control. Both constructs include a mini-Transposon carrying *Kan*^{R} gene, chloramphenicol resistance, RP4 ori, and a Pir-dependent origin of replication (oriVR6Kγ). Both TnpA and TnpI are under the expression of pLac promoter. Designed with BioRender.

Cell lysates were analyzed under conditions as close as possible to experimental conditions chosen to perform the conjugation-coupled proximity ligation assay by performing a conjugation between donor MFDPir cells harboring the AirD vectors and recipient MG1655 cells containing the ampicillin-resistant pRS415 plasmid. Conjugation was performed on LB plates containing 50 µM of Biotin and 250 µg/mL of ampicillin as described above (see **III.3.2**). Here, incubation was continued to up to 5 h in the presence or absence of 0.5 mM of IPTG to evaluate the effect of protein expression level on stability (Figure **III.3.5-2**).

The western blot analysis is summarized in **Figure III.3.5-2 to III.3.5-5**. The first experiment aimed at testing the TnpA-AirID construct stability after conjugation **(Figure III.3.5-2A)**. In the test condition (TnpA-AirID), a band around 140 kDa was visible, corresponding to the fusion of TnpA (113kDa) and AirID (35kDa). This band vanished after 2h30' of conjugation, suggesting potential degradation of the construct. Similar degradation was noticed in the negative control (AirID only), where the 35 kDa band corresponding to AirID started to show degradation after 4h of conjugation. The diminishing intensity of bands related to both constructs correlated with the emergence of this new, intensifying band of low molecular mass over time. The molecular weight of this band did not correspond to the FLAG sequence alone or the FLAG sequence and biotin ligase.

The experiment was repeated with induced expression of our constructs **(Figure III.3.5-2B)** to see if we could increase the amount of expressed TnpA-AirID construct through induction. The results exhibited a comparable profile between induced and non-induced conditions with the exception that bands corresponding to both constructs and degradation appeared earlier. IPTG induction seems to "accelerate" degradation, which could be caused by the precipitation of proteins due to its overexpression.

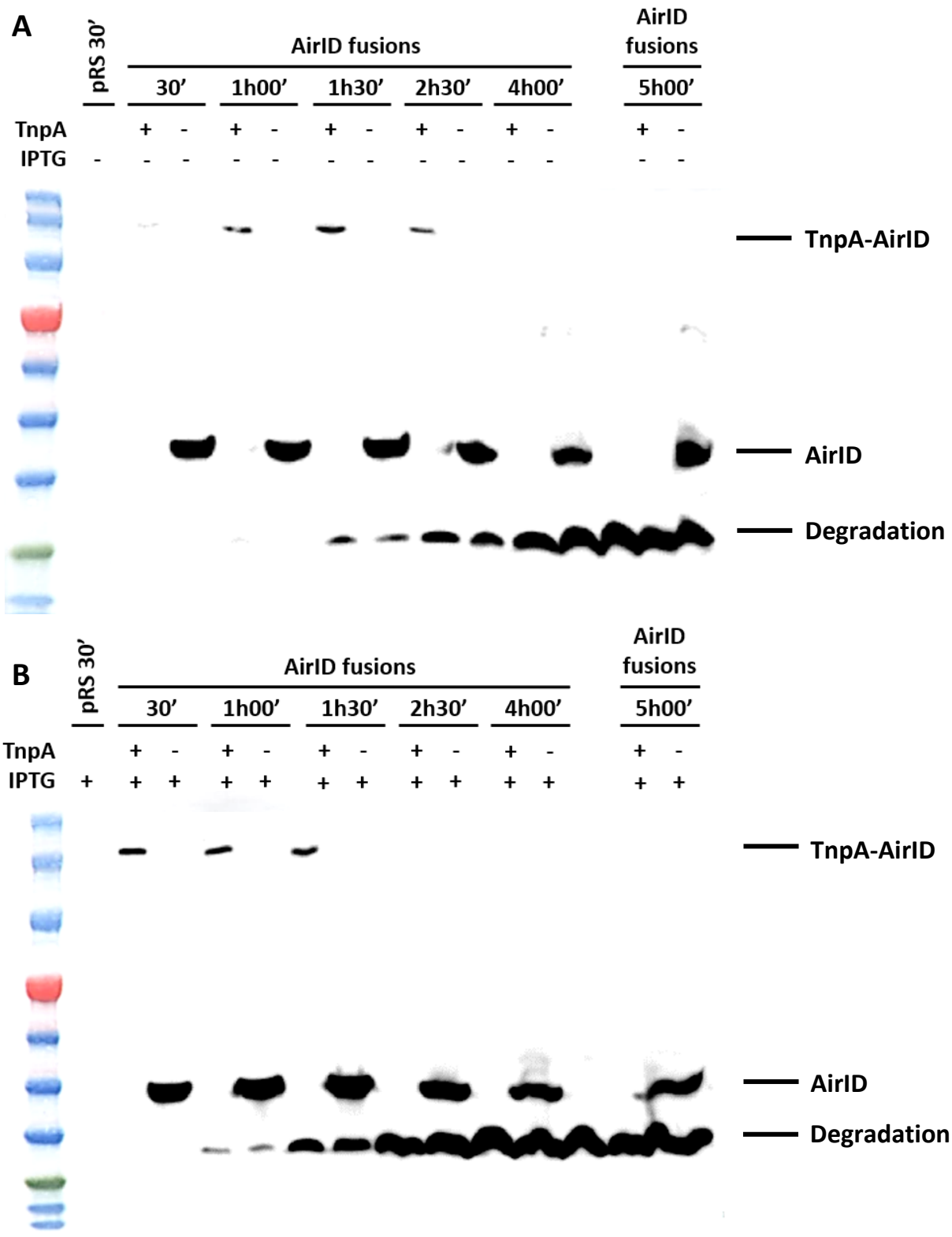


Figure III.3.5-2 – AirID fusion proteins show signs of degradation after 1h30. (A) Western blot analysis was conducted to assess the stability of the AirID constructs after conjugation on a LB plate containing ampicillin and DAP, ranging from 30 minutes to 5 hours. The primary antibody targeted the FLAG sequence integrated into the constructs, and detection was performed using a secondary antibody conjugated with horseradish peroxidase. TnpA (+) lanes correspond to TnpA-AirID and (-) corresponds to AirID alone. Lanes labeled "pRS" represent recipient cells without constructs. **(B)** Western blot analysis of cellular extracts of cells conjugated as mentioned above on LB + ampicillin + IPTG plate.

To see whether degradation was due to the presence of ampicillin in the conjugation plates and subsequent lysis of donor cells or whether it is intrinsic to the protein constructs, we checked the stability of the TnpA-AirID fusion protein on petri dishes containing DAP only without ampicillin (Figure **III.3.5-3**).

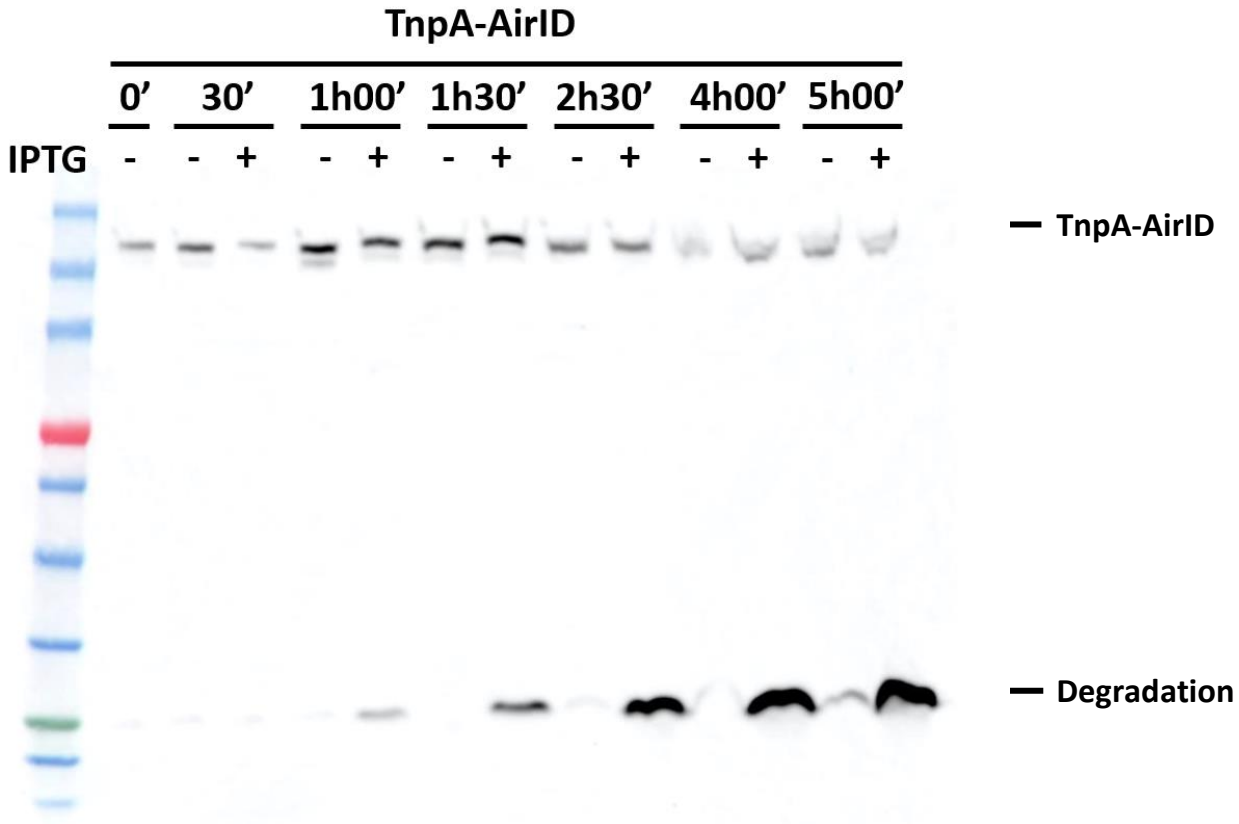


Figure III.3.5-3 – TnpA-AirID fusion show signs of degradation in absence of ampicillin during conjugation. Western blot analysis was conducted to assess the stability of the AirID constructs after conjugation on LB plates with DAP and without ampicillin, ranging from 30 minutes to 5 hours. The primary antibody targeted the FLAG sequence integrated into the constructs, and detection was performed using a secondary antibody conjugated to horseradish peroxidase. IPTG (+) or (-) correspond to cells conjugated on plates with or without IPTG.

Surprisingly, the TnpA-AirID fusion exhibited signs of degradation even in the absence of ampicillin (Figure **III.3.5-3**). The band corresponding to the degradation product showed an apparent increase in intensity when expression was induced by IPTG, although the band corresponding to the TnpA-AirID fusion was of similar intensity in both conditions. This suggests that overexpression of TnpA-AirID increased its degradation rate, possibly by causing its precipitation.

Given that degradation of our construct could not be solely attributed to the addition of ampicillin, we investigated whether conjugation itself could impact the expression of our construct. For this purpose, we conducted an assay simulating conjugation conditions without the addition of our recipient cell. Donor cells were spotted on an LB + DAP petri dish for 30 minutes to 5 hours to monitor TnpA-AirID expression levels (**Figure III.3.5-4**). The band corresponding to the TnpA-AirID construct exhibited an increase in protein levels, with almost no discernible degradation band throughout the entire experiment. These findings collectively suggest a conjugation-dependent degradation of the TnpA-AirID construct after 2 hours and 30 minutes of conjugation.

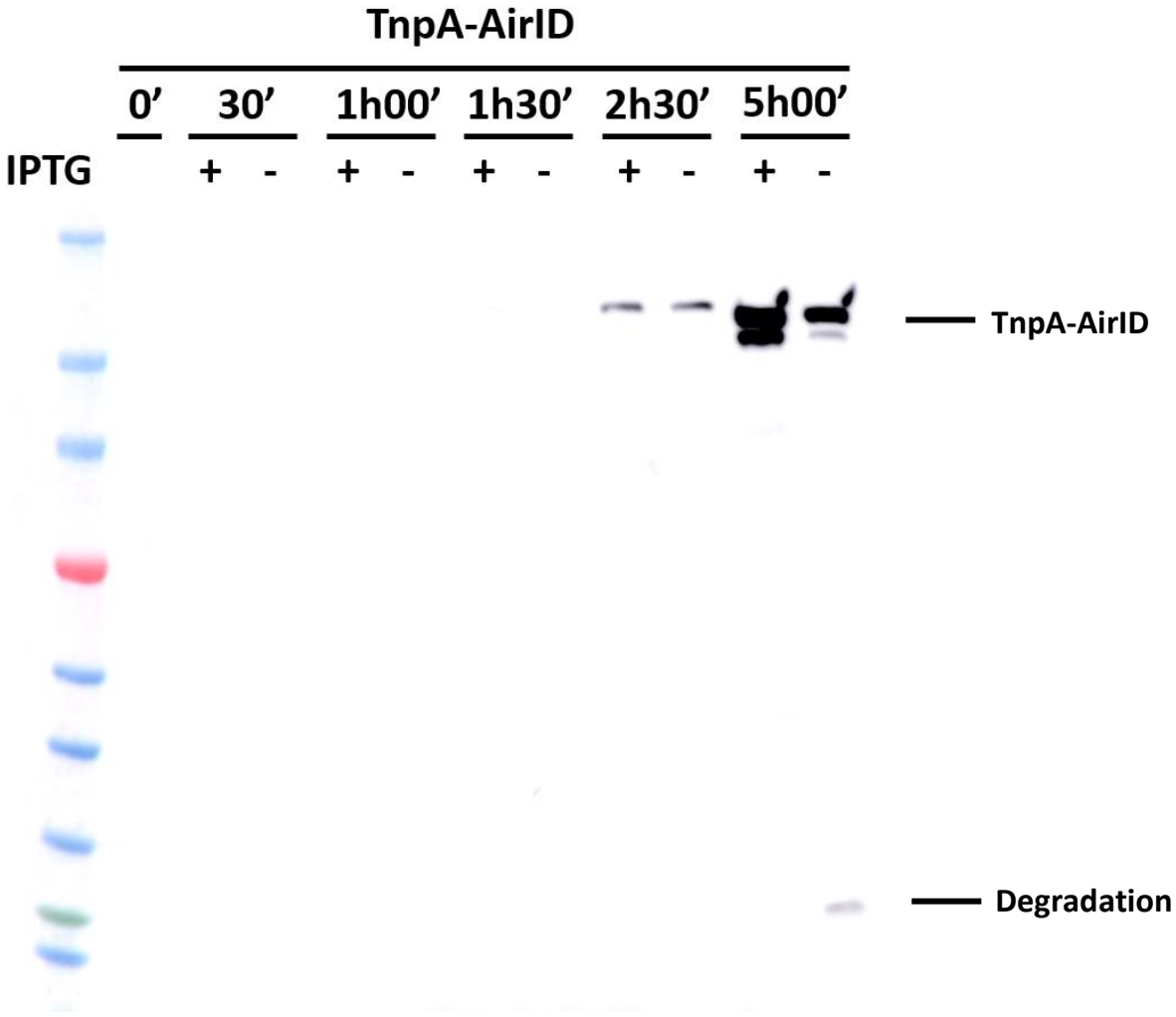


Figure III.3.5-4 – TnpA-AirID degradation depends on conjugation. Western blot analysis was conducted to assess the stability of the AirID constructs in donor cells in the absence of recipient cells and ampicillin over 30 min to 5 h of growth on plates. The primary antibody targeted the FLAG sequence integrated into the constructs, and detection was performed using a secondary antibody conjugated with horseradish peroxidase. (+) lanes correspond to induced expression by IPTG and (-) corresponds to leaky expression (without IPTG).

Finally, to assess the stability of the equivalents TnpA-miniTurbo constructs (C. Stulemeijer, unpublished; see Material and Methods), we conducted a similar experiment with both the TnpA-miniTurbo fusion protein and miniTurbo alone. The western blot analysis (Figure **III.3.5-5**) showed no detectable bands corresponding to the full-length proteins although a low molecular-weight band similar to that observed with the AirID constructs emerged over time, suggesting the potential accumulation of degraded products.

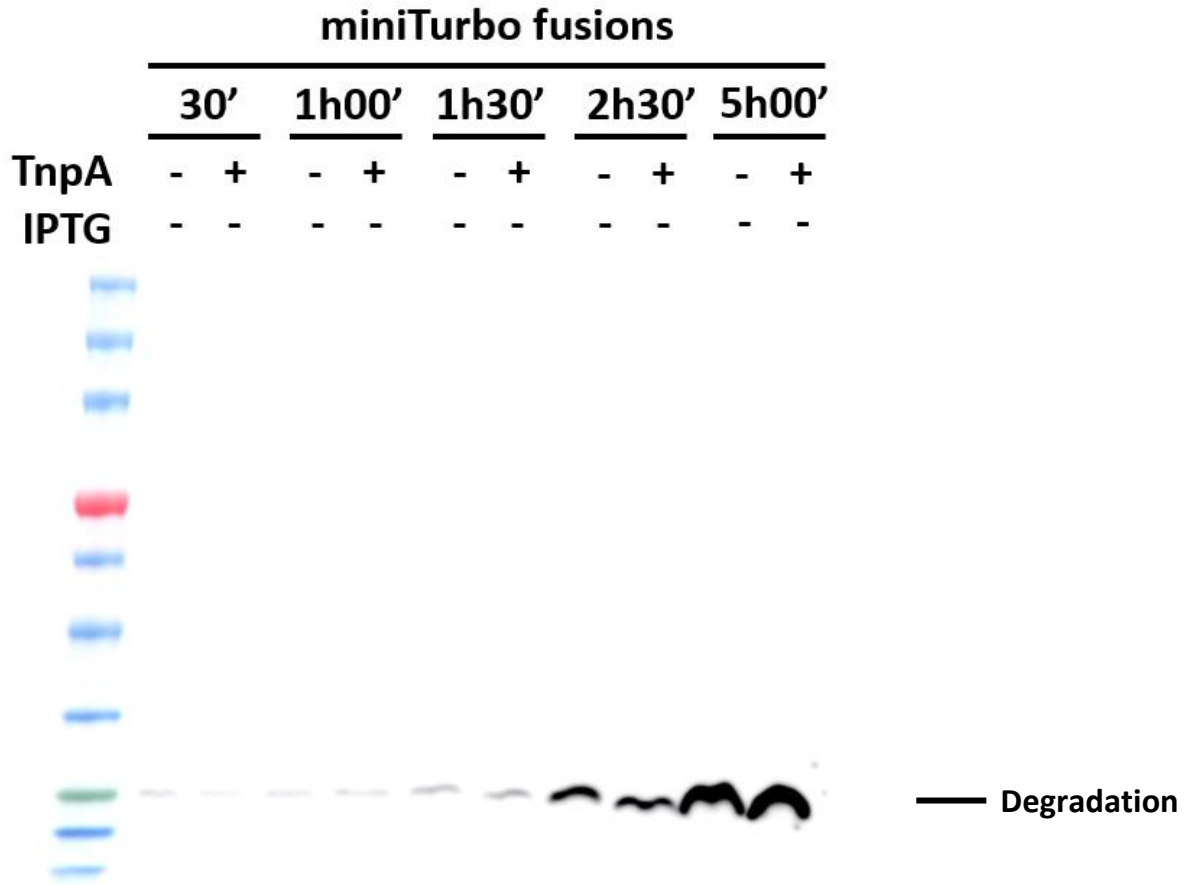


Figure III.3.5-5 – miniTurbo fusion proteins are not stable under conjugation conditions. Western blot analysis was conducted to assess the stability of the miniTurbo-derived proteins under the experimental conditions used for coupled conjugation-proximity labelling assay over 30 mins to 5 h of incubation. The primary antibody targeted the FLAG sequence integrated into the constructs, and detection was performed using a secondary antibody conjugated to horseradish peroxidase. TnpA (+) lanes correspond to TnpA-miniTurbo and (-) corresponds to miniTurbo alone.

III.4 Streptavidin enrichment of proteins biotinylated by AirID and miniTurbo fusions

In preparation for mass spectrometry analysis of biotinylated proteins resulting from the proximity labelling experiments, it was imperative to set up purification conditions to segregate tagged proteins from their non-tagged counterparts. To achieve this, we enriched biotinylated proteins from conjugated cell lysates using streptavidin-coated magnetic beads (see Material and Methods). This methodology capitalizes on the robust and specific interaction between biotin and streptavidin.

III.4.1 Successful enrichment of biotinylated proteins

Cells from both positive and negative controls were subjected to mechanical lysis followed by dialysis to eliminate residual free biotin molecules preventing saturation of the streptavidin-conjugated beads. These dialysed lysates were then incubated on the magnetic beads with gentle shaking for 1h30 at room temperature. Rigorous washing steps were employed to ensure the removal of non-tagged proteins, effectively eliminating any contaminants that might have adhered due to nonspecific interactions.

The western blots presented in Figure **III.4.1-1** depict the biotinylation profiles obtained at each stage of the enrichment process. As above, biotinylated proteins were visualized using streptavidin conjugated to alkaline phosphatase. Cell lysates were spotted on the gel prior to binding onto the beads (pre-binding) as well as the supernatants remaining after incubation with the beads (post-binding) to check for unbound proteins in the mix.

A decrease of total biotinylation signal was observed in the post-binding samples, indicating that a substantial fraction of biotinylated proteins was retained onto the beads upon binding to streptavidin. Minimal loss of biotinylated proteins occurred during washing steps reflecting the high affinity between biotin and streptavidin interaction and poor non-specific binding activity of the beads.

This robust interaction between biotin and streptavidin posed challenges for the elution step. However, the focus of this experiment was to set up conditions for preparing samples for mass spectrometry analysis, during which elution of the biotinylated species were conducted by directly digesting the proteins bound onto the beads. Consequently, optimization of elution conditions was not the primary concern of this study, and for western blot analysis, tagged proteins were simply eluted by heating the beads in the presence of SDS and 2-mercaptoethanol.

For both the miniTurbo and AirID fusions, the biotinylation profile of the elution fractions closely resembled that of cell lysates with some variations in the relative intensities of the bands likely reflecting different specificities between labelled proteins (Figure **III.4.1-1**).

Most interestingly, in both biotinylation systems, the elution pattern given by the TnpA-BirA* fusion differed from that obtained for the unfused BirA* suggesting different specificities in proximity labelling activities (Figure **III.4.1-1**, compare pre-binding lanes). In particular, the biotinylation profile of the TnpA-BirA* derivatives showed a specific band corresponding to the fused TnpA protein. This specific product presumably arose from auto-biotinylation of the corresponding protein. However, control samples did not show any specific band corresponding to the unfused BirA* ligase. This result was not surprising since AirID was previously found to not auto-biotinylate itself, probably due to a lack of accessible lysine residue (Kido et al. 2020a) which could also be the case for miniTurbo. Likewise, other discrete bands were absent from one sample and present in the other sample or showed marked differences in intensity, indicative of differential labelling. This was most obvious for the AirD-derived couple of proteins, which were therefore chosen for the large-scale proximity ligation sample preparation.

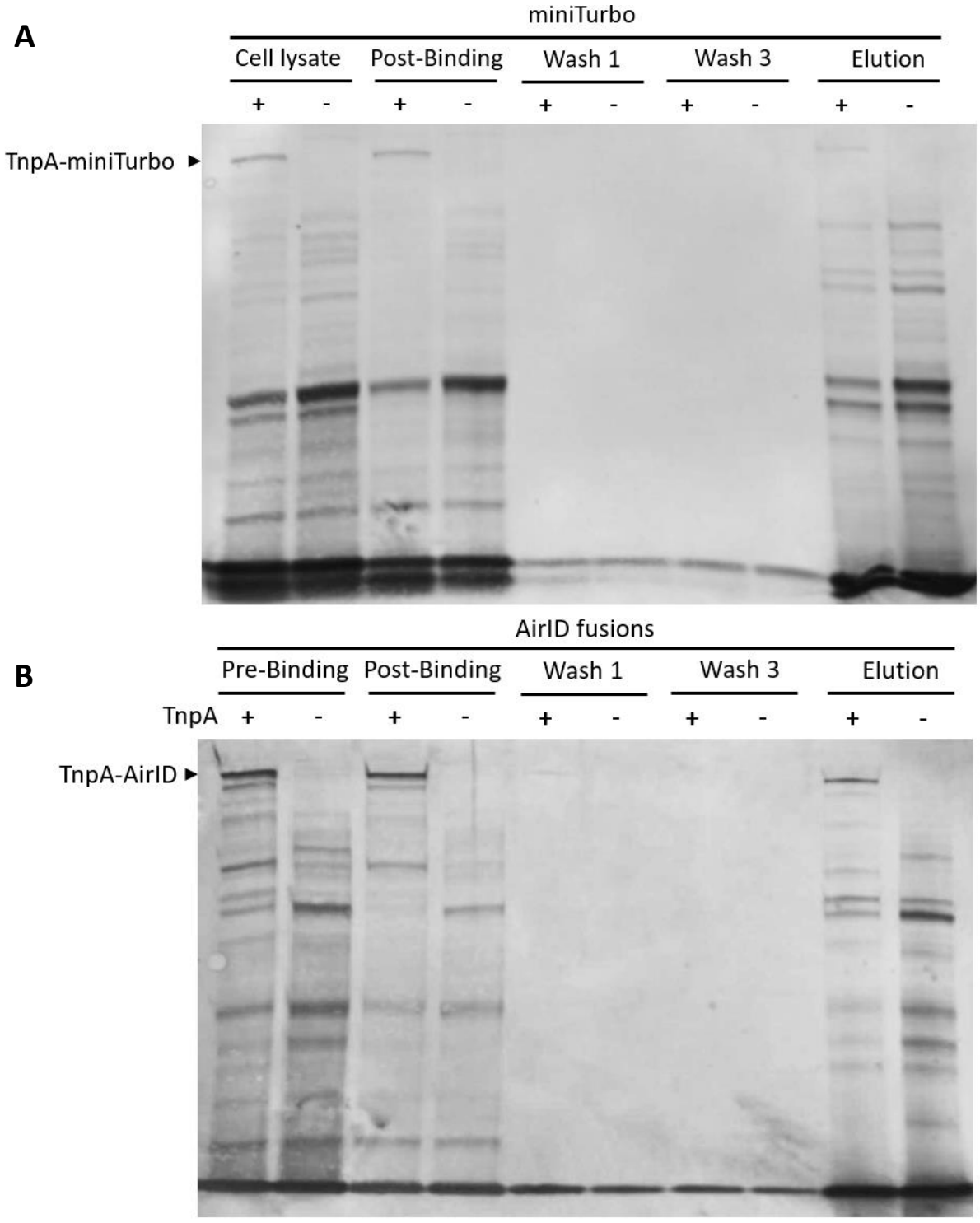


Figure III.4.1-1 – Biotinylated proteins enrichment. (A) Western blot analysis of proteins biotinylated with miniTurbo fusion proteins at different steps of the enrichment process. The samples analyzed include "Pre-binding represents the cell lysate after dialysis and prior to incubation with streptavidin beads. The "post-binding" samples correspond to the supernatant after incubation with the beads. The "Wash 1 and 3" samples correspond to the first and last washes of streptavidin beads before the elution step. Elution was performed by adding 5% SDS to the beads and heating samples at 95°C for 10 minutes. Control (+) lanes correspond to TnpA-BirA* and (-) corresponds to BirA* alone. **(B)** Western blot analysis of proteins biotinylated with AirID fusion proteins at different steps of the enrichment process. Coomassie blue gels corresponding to western blots are available in supplementary data. Coomassie blue gels corresponding to western blots above are found in Supplementary Figure **III.4.1-1**.

III.4.2 Sample preparation for mass spectrometry analysis

As a quality control of the samples sent to mass spectrometry analysis, we performed similar western blot and SDS page analysis as previously described. In this context, we conducted an experiment with increased protein concentrations and magnetic bead quantity, allowing us to visualize elution profiles on Coomassie blue gels.

Higher protein concentration resulted in the emergence of additional discernible bands, as depicted in the "Elution" lanes of **Figure III.4.2-1A**. Crucially, the gel representing all proteins (Coomassie blue in **Figure III.4.2-1B**) showed no visible reduction in protein concentration post-binding to the beads, indicating that biotinylated proteins constitute only a minor fraction of the total protein content. The minimal protein loss during the washing steps suggests a robust and specific binding of biotinylated proteins to the beads. Analysis of the elution profile of biotinylated proteins **(Figure III.4.2-1A)** revealed numerous bands with relatively high intensity. When subjected to Coomassie blue staining **(Figure III.4.2-1B)**, only a few bands were detected, aligning with the high-intensity bands of biotinylated proteins. This finding underscores the absence of contamination by non-biotinylated proteins during the enrichment procedure and successful purification of biotinylated proteins from cell lysate.

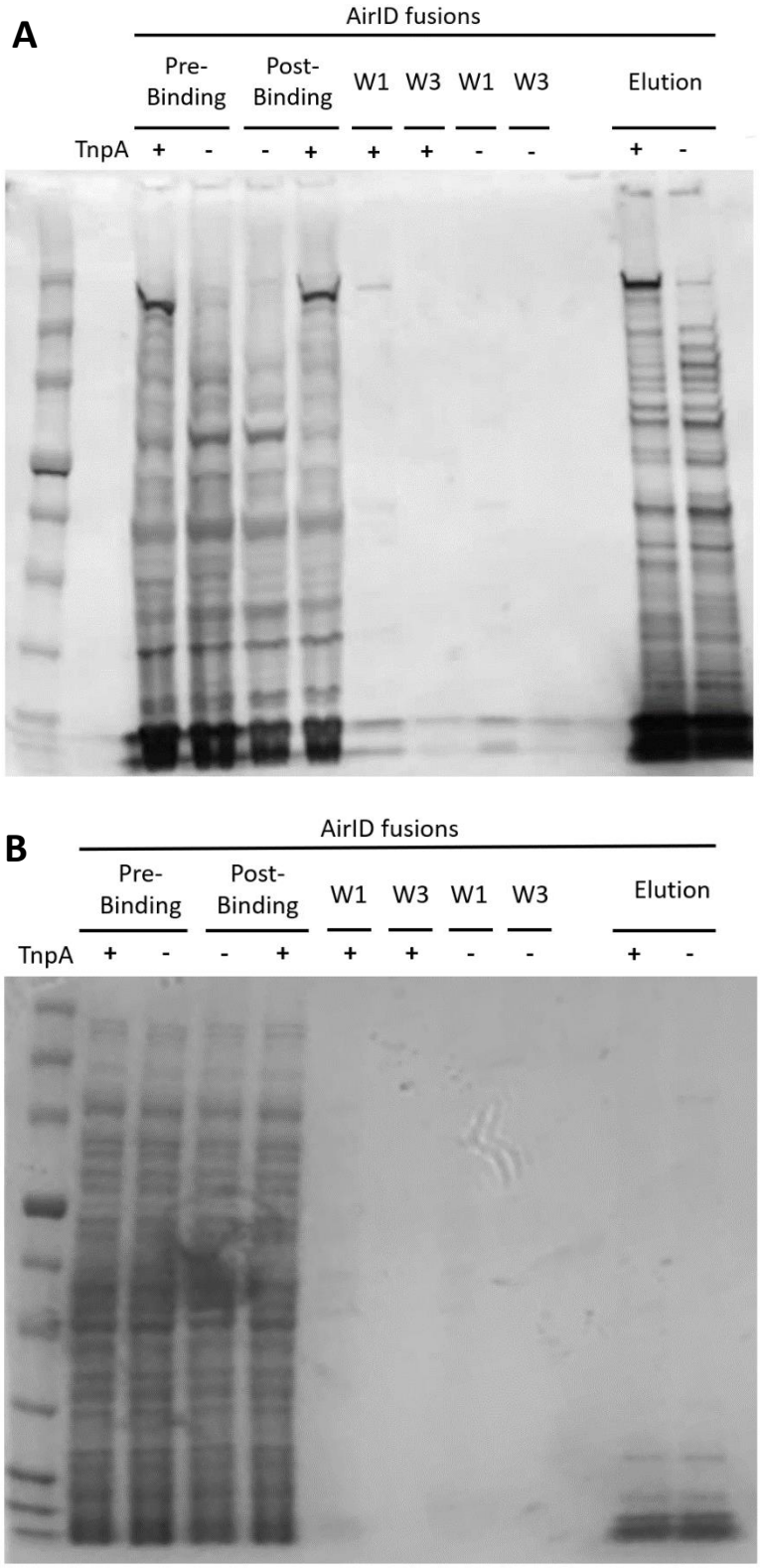


Figure III.4.2-1 - Biotinylated proteins enrichment. (A) Western blot analysis was performed to examine biotinylated proteins at different steps of the enrichment process. The samples analyzed include "pre-binding," which represents the cell lysate after dialysis and prior to incubation with streptavidin beads. "Post-binding" samples correspond to the supernatant after incubation with the beads. "W1" and "W3" samples correspond to the first and last washes of streptavidin beads before the elution step. Elution was performed by adding 5% SDS to the beads and heating samples at 95°C for 10 minutes. **(B)** Coomassie blue-stained gel corresponding to western blot of **(A)**.

III.5 Mass spectrometry

Large-scale AirID-prepared samples have been sent to the mass spectrometry platform where biotinylated proteins have been digested directly on the magnetic beads, and the resulting peptides were analysed by LC-MS/MS spectrometry.

Quantitative analysis measured the total intensity of all identified peptides for each protein (peptide peak) expressed as the iBAQ (Intensity Based Absolute Quantification) value. While iBAQ values provides insights about the relative abondance of different proteins within a specific sample, it is not appropriate to directly compare the ocurance of a given set of proteins present in the sample under different conditions. In particular, the distribution of iBAQ values revealed that the average abondance of biotinylated proteins collected with AirID alone was higher than the average distribution obtained for the TnpA-AirID fusion (Figure **III.5-1**). Consequently, to address this limitation and to analyze the data from a quantitative perspective, we normalized iBAQ values by the total abundance of protein in each sample, resulting in riBAQ (relative intensity Based Absolute Quantification) (see Supplementary Data). The normalized riBAQ values were instrumental in determining which proteins were enriched in the TnpA-AirID-collected sample compared to the control, as evidenced by previous studies (Krey et al. 2014; Lu et al. 2007; Schwanhäusser et al. 2011).

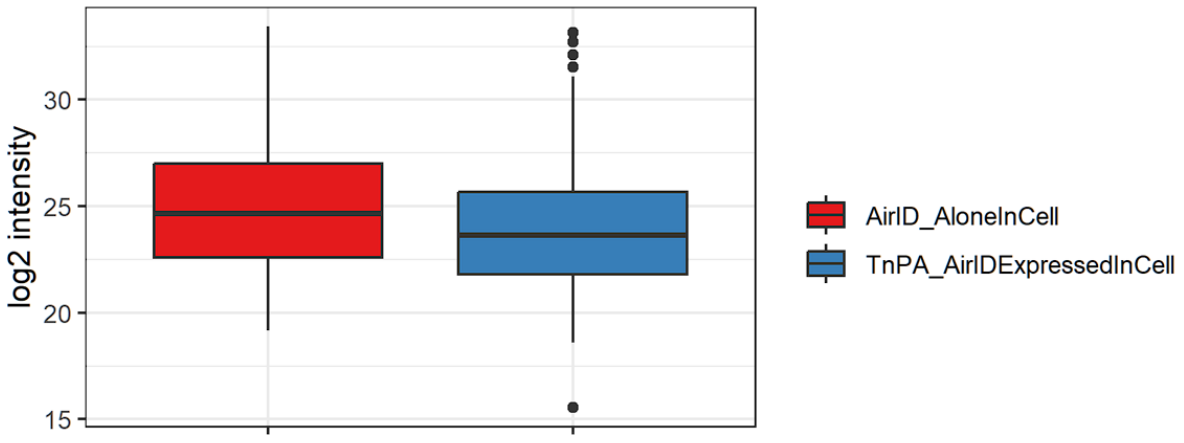


Figure III.5-1 – Samples intensity distribution. Boxplot representing total intensity of positive (blue) and negative (red) control samples.

In the absence of replicates, we complemented this partial quantitative analysis with an arbitrary selection of potential candidates based on functional criteria with a specific focus on proteins involved in DNA interactions, components of replication and DNA repair mechanisms, and intriguing yet uncharacterized proteins (Table **III.5**, DNA binding proteins in green, proteins associated with DNA replication/or repair in blue, proteins associated with translation in orange, proteins associated with cell division in yellow, and uncharacterized proteins in grey). This selection was guided by its relevance in the context of DNA replication, aligning with the targeting mechanism of TnpA in Tn4430. Potential interaction between these different proteins and TnpA was further investigated using the bacterial-two-hybrid method.

allR	HTH-type transcriptional repressor AllR
cbpA	Curved DNA-binding protein
cspC	Cold shock-like protein CspC
fis	DNA-binding protein Fis
ftsA	Cell division protein FtsA
gntR	HTH-type transcriptional regulator GntR
helD	DNA helicase IV
hns	DNA-binding protein H-NS
himA	Integration host factor subunit alpha
hupA	DNA-binding protein HU-alpha
hupB	DNA-binding protein HU-beta
lrp	Leucine-responsive regulatory protein
minE	Cell division topological specificity factor
recA	DNA repair protein RecA
recO	DNA repair protein RecO
rplE	Large ribosomal subunit protein uL5
rpsS	Small ribosomal subunit protein uS19
rpsT	Small ribosomal subunit protein bS20
stpA	DNA-binding protein StpA
tufA	Elongation factor Tu 1
tufB	Elongation factor Tu 2
ydiZ	Uncharacterized protein YdiZ
yejK	Nucleoid-associated protein YejK
yfaA	Uncharacterized protein YfaA
yjgR	Uncharacterized protein YjgR
yjhU	Uncharacterized protein YjhU

Table III.5 – List of selected proteins

III.6 Using Bacterial Two Hybrid to validate potential interaction candidates

In order to discern authentic interactants from incidental proteins located in close proximity to our protein of interest, we employed a robust validation strategy utilizing the bacterial two-hybrid system. This method hinges on the reconstitution of a split adenylate cyclase enzyme (CyaA), fragmented into its C-terminal T18 and N-terminal T25 segments (Battesti and Bouveret 2012; Karimova et al. 1998). These fragments were strategically fused to the proteins of interest, serving as a 'bait' and 'prey', and co-transformed into a bacterial host strain devoid of endogenous adenylate cyclase activity.

Upon interaction between the bait and prey proteins, the T18 and T25 fragments spatially converge, allowing for the reconstitution of a catalytically active CyaA adenylate cyclase. This reconstituted enzyme, in turn, catalyses the synthesis of cAMP from ATP. The generated cAMP acts as a signalling molecule, activating a catabolite activator protein (CAP). This activation triggers the transcription of a reporter gene, such as lac, catalyzing lactose degradation. On indicator McConkey plates, fermentation of lactose into lactic acid leads to acidification of the media, which is visualized by a distinct colour change around bacterial colonies, whereas on plates containing the chromogenic X-Gal analog of lactose, β-galactosidase activity releases the indole moiety of the substrate, colouring colonies in blue.

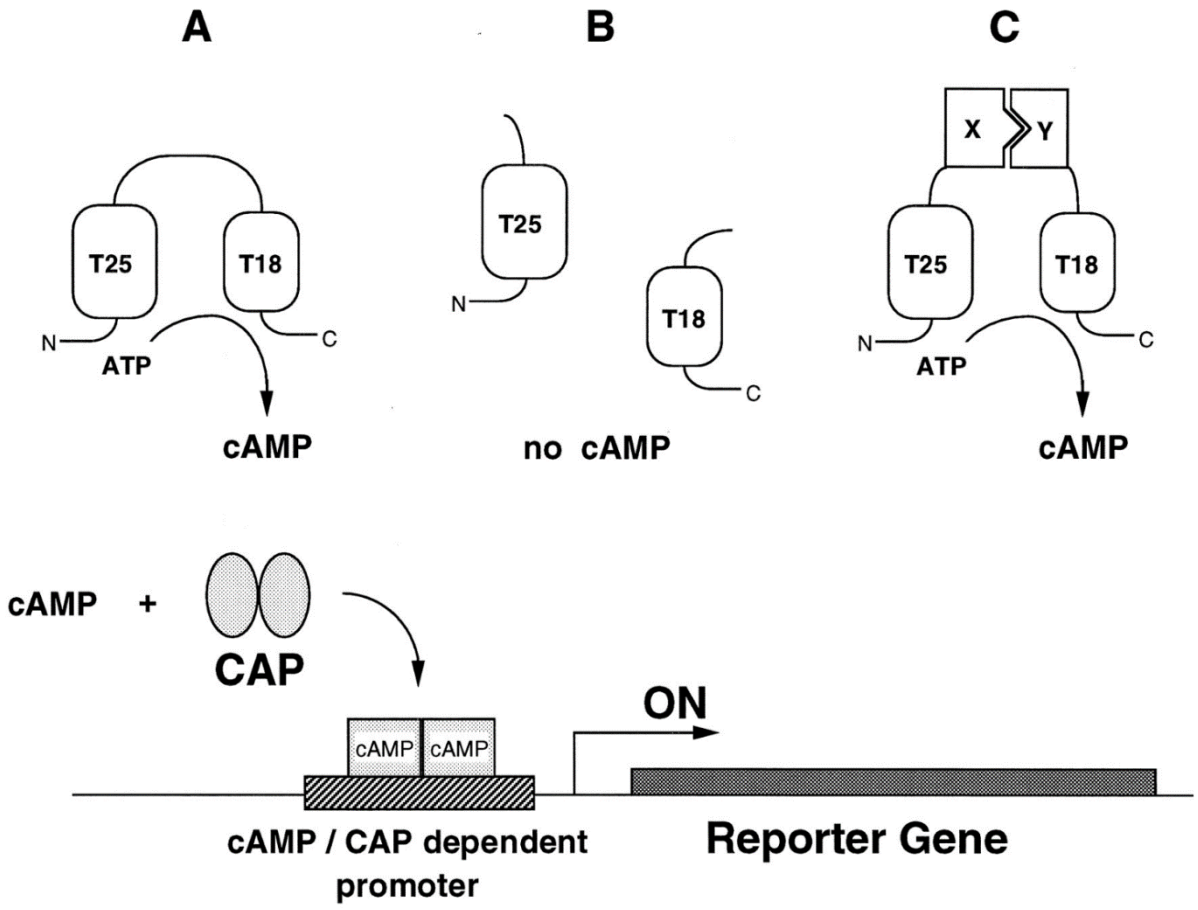


Figure III.6-1 - Principle of an *E. coli* two-hybrid system based on functional complementation of CyaA fragments. **(Upper)** The two boxes represent the T25 and T18 fragments corresponding to amino acids 1–224 and 225–399 of the CyaA protein. **(A)** Full-length catalytic domain (residues 1–399), when expressed in *E. coli*, exhibits a basal calmodulin-independent activity that results in cAMP synthesis. **(B)** the two fragments, T25 and T18, when co-expressed as independent polypeptides, are unable to interact and no cAMP synthesis occurs. **(C)** The two fragments, fused to two interacting proteins, X and Y, are brought into close proximity, resulting in functional complementation followed by cAMP production. **(Lower)** Schematic of the readout of the complementation. cAMP, synthesized in an *E. coli* cya strain by the complementing T25 and T18 pairs, binds to the catabolite gene activator protein, CAP. The cAMP/CAP complex then can recognize specific promoters and switch on the transcription of the corresponding genes. From Karimova et al., 1998.

III.6.1 Bacterial Two Hybrid constructs

To implement this methodology, the creation of distinct constructs for each protein under investigation was indispensable (see Table **III.5**). These constructs were devised by fusing the protein with the T25 fragment of the adenylate cyclase enzyme. The construction process commenced with the PCR amplification of the protein sequences directly from the E. coli chromosome, facilitated by specially designed primers. Subsequently, Gibson assembly was employed to seamlessly integrate these amplified sequences with the backbone containing the T25 fragment.

Prior experiments conducted in the laboratory showed that optimal results were attained when the T25 fragment was fused to the C-terminal extremity of the protein rather than at the N-terminal. The integrity of each construct was subjected to rigorous validation through thorough sequencing, confirming the accurate fusion of the T25 fragment with the respective proteins of interest. The resulting constructs served as the foundation for subsequent interaction assays within the bacterial two-hybrid system.

III.6.2 Bacterial Two Hybrid results

Different combinations between the chromosomal protein-T25 preys and the TnpA-T18 bait were transformed into BTH101 strain and separate co-transformants of each pair of constructs were cultured in fresh LB. 2 µL of cell culture were then spotted on McConkey plates and incubated for 22-24 hours. The experiment was validated by using negative controls in which the either the chromosomal prey protein and/or the TnpA bait was replaced by the unfused T25 and T18 fragment. A positive control with TnpA fused to both T18 and T25 fragments individually (TnpA dimerization). Unfortunately, the two first replicates of the experiment were not reproducible, giving distinct patterns of positive and negative signals for the same couples of proteins (Figure **III.6.2-1** and Supplementary Figure **III.6.2-1**). We therefore decided to repeat the experiment for every protein that showed signs of interactions at least once.

The results of the first experiments are presented in Figure **III.6.2-1** as an example and show cultures that were spotted on McConkey plates. Each protein-T25 constructs were co-transformed with a plasmid containing the TnpA-T18 bait and a plasmid containing the T18 on its own (Empty-T18). As shown in Figure **III.6.2-1**, multiple combinations of protein-T25 preys and TnpA-T18 bait yielded signal on McConkey plates. Unfortunately, as mentioned above, signal was not constant across replicates within the same experiment. Similar results were obtained during the second experiments with other protein-T25 baits leading to positive signal (Supplementary Figure **III.6.2-1**). Notably, it seemed that most candidates gave a signal when co-expressed with TnpA-T18 and not 18 alone which could suggest interactions with the TnpA.

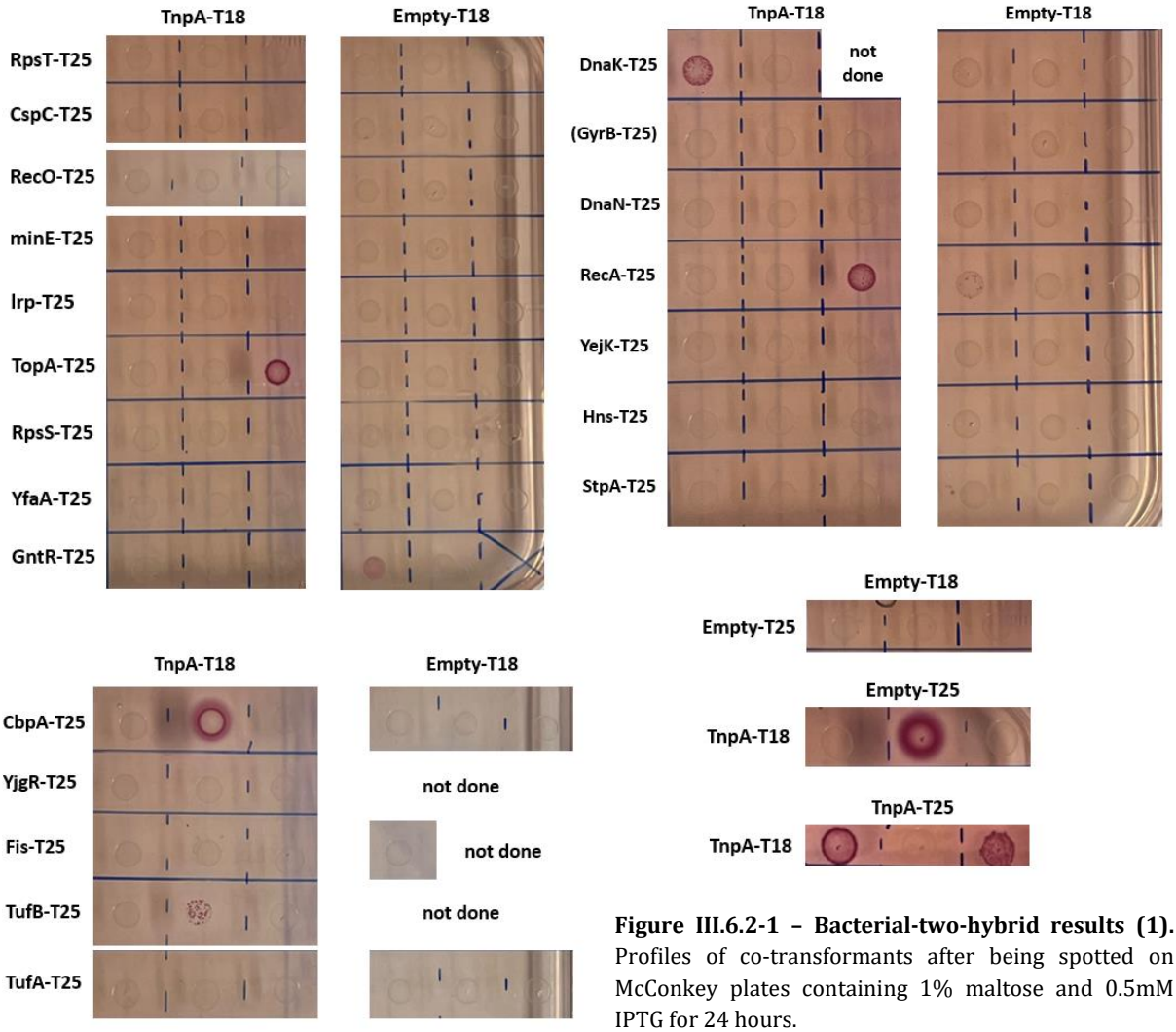


Figure III.6.2-1 – Bacterial-two-hybrid results (1). Profiles of co-transformants after being spotted on McConkey plates containing 1% maltose and 0.5mM IPTG for 24 hours.

In the third experiment we decided to spot cells on two types of plates in parallel: on McConkey plates and on X-Gal-containing plates (Supplementary Figure **III.6.2-2**). Intriguingly, in this experiment, every tested prey protein showed an activation of the reporter gene when co-expressed with the TnpA-T18 bait. In addition, some negative controls like TnpA-T18 and T25 alone were also giving signal, meaning that positive interactions with tested proteins were not significant. These observations on McConkey plates were coherent with results obtained on X-Gal plates suggesting that results were independent from the reporter gene.

IV. Discussion

IV.1 Reduced background biotinylation during proximity labelling

The proximity labelling approach has become instrumental in investigating protein-protein interactions across various organisms, including *Arabidopsis* (Khan et al. 2018; Mair et al. 2019), *C. elegans* (Branon et al. 2018b; Reinke et al. 2017), *Drosophila* (Branon et al. 2018b; Chen et al. 2015; Li et al. 2020; Shinoda et al. 2019), and mouse (Brudvig et al. 2018; Dingar et al. 2015; Uezu et al. 2016). To effectively employ this method, it is imperative to control endogenous biotinylation and fine-tune experimental conditions to minimize background noise while amplifying specific biotinylation signals. Given that proximity labelling experiments utilize different biotin ligases in diverse organisms and organelles, the translation of experimental conditions from one case to another poses a significant challenge. The inherent diversity among organisms necessitates a thorough examination of background biotinylation to ensure the reliability and specificity of the proximity labelling results.

In this study, we first assessed endogenous biotinylation in *E. coli* by BirA in (**III.3.1-1**, see recipient cells only). In the absence of exogenous biotin and BirA*, western blot analysis revealed a single intense band corresponding to the biotin carboxyl carrier protein (BCCP), in agreement with the fact that only a specific lysine residue of BCCP is efficiently recognized and biotinylated by BirA (Chapman-Smith and Cronan 1999). The addition of exogenous biotin directly to the culture media resulted in a discernible increase in global biotinylation, indicating successful biotin entry into the cells, likely facilitated by active biotin transporters in *E. coli* (Prakash and Eisenberg 1974). Furthermore, exogenous biotin did not seem to cause specific biotinylation of some proteins, showing that *E. coli* is suitable to carry out our experiment.

The suicide conjugation assay played a pivotal role in controlling and maximizing the specificity of biotinylation, ensuring that TnpA was in conditions to transpose and interacts with its partners. We systematically tested and analysed biotinylation profiles of cell lysates during and after the conjugation assay. We systematically examined biotinylation profiles of cell lysates during and after the conjugation assay, optimizing various conditions such as time, donor and recipient cell ratios, and the addition of ampicillin. This optimization aimed to mitigate biotinylation background resulting from the overnight activity of BirA* enzymes, known to vary depending on the specific enzyme. Notably, TurboID which is known for utilizing endogenous biotin(Branon et al. 2018b), exhibited the highest activity among all enzymes tested in our study. Despite some residual background biotinylation from the overnight incubation, the fine-tuning of these conditions substantially reduced background signals.

IV.2 TnpA fusion proteins are active in transposition

Corrected transposition frequencies of TnpA-biotin ligases fusion proteins (highlighted in **III.3.3**) differs depending on the biotin ligase. The TnpA-TurboID fusion has a reduced activity (compared to miniTurbo and AirID fusion) which is at similar level than the negative control lacking TnpA. Therefore, the transposition activity observed for TnpA-TurboID is probably an artefact making the use of this fusion protein inappropriate for our experiment. Notably, the corrected transposition frequencies also show a significant difference in the transposition background level of negative controls. In fact, TurboID and mini-Turbo alone have a similar frequency of "false" transposition events that sits around twice more than AirID alone. This observation raises questions about the reasons for this difference. TurboID and mini-Turbo biotin ligases were created by directed evolution and have been developed by the same team, with mini-Turbo being a shorter version of TurboID with a deleted N-terminal domain (Branon et al. 2018b). Following the hypothesis of DNA recombination between the donor plasmid containing the biotin ligase and chromosomal DNA containing the wild-type BirA protein we investigated sequence similarity by local pairwise sequence alignment (see Supplementary Data). This local alignment analysis revealed a high sequence similarity for both TurboID and miniTurbo with BirA (75.4% and 75.7%) while AirID sequence similarity was lower (67.2%). Unfortunately, even if the requirement of a 20bp complete homologous fragment for significant recombination is reached it has also been shown that mismatches dramatically decrease recombination (Watt et al. 1985). Thus, results do not allow any conclusions whether the sequence similarity and recombination events could be an explanation for "false" transposition events.

IV.3 Characterization of transposition events

Measurements of transposition frequencies showed that the TnpA-AirID and TnpA-miniTurbo fusion proteins were active in transposition. This step was important to validate the proximity labelling approach. We then investigated transposition products resulting from the mating out assay with TnpA-AirID fusion as well as donor strains containing the BirA* enzyme without TnpA. These transconjugants were unexpected at first but were consistent with findings reported by d'Udekem, O. in her PhD thesis, prompting the formulation of two hypotheses to elucidate these observations. Firstly, it was hypothesized that the suicide donor plasmid might not be immediately lost due to the presence of its replication origin, which prevents replication in the recipient strain. Secondly, considering that MG1655 recipient strain possess all components of homologous recombination machinery, it was plausible that recombination events between the donor plasmid and the recipient plasmid or chromosome occurred.

In the subsequent experiments performed with AirID, the above-mentioned hypotheses were explored by extracting plasmids from cells growing on kanamycin after conjugation. Among the colonies growing on kanamycin and ampicillin containing the TnpA-AirID fusion, three out of five displayed an expected transposition profile. In one case, the mini-Tn*Km* might have inserted into the chromosome, resulting in the loss of resistance after back-transformation and the disappearance of the band corresponding to the mini-Tn*Km*.

Colonies containing TnpA-AirID fusions but growing on chloramphenicol and ampicillin revealed two distinct profiles, one suggesting a recombination event with the chromosome, and the second profile suggesting the conservation of the donor plasmid. Although theoretically unlikely due to its replication impossibility in the recipient, recombination between plasmids might explain the conservation of the donor plasmid, rescued by the recipient plasmid's origin of replication.

Negative controls containing only the AirID enzyme exhibited colonies resistant to both kanamycin and chloramphenicol in all investigated cases, suggesting plasmid conservation or recombination with recipient DNA molecules. While the results obtained from this experiment did not provide sufficient information to validate the initial hypotheses, it should be noted that a detailed characterization of these "transposition events" was beyond the scope of the study's objectives, and thus, no further experiments were conducted.

IV.4 Stability of fusion proteins *in vivo*

Concurrent with the conjugation assays, we conducted stability tests for both TnpA-BirA* and BirA* alone, utilizing previously optimized conjugation conditions designed to minimize background biotinylation. To assess the stability of fusion proteins, we performed western blot analysis on cell lysates collected at various time points during conjugation. The detection of FLAG sequences (see **III.3.5**) revealed the presence of both TnpA-AirID and AirID alone in cell lysates after 30 minutes of conjugation. The protein abundance of both fusions increased until 1 hour of conjugation, and while it is challenging to attribute this increase specifically to expression in donor or recipient cells, it is noteworthy. Subsequently, degradation became evident, particularly as donor cells lysed in the presence of ampicillin. Even in experiments without ampicillin, where proteins exhibited a slightly higher stability, degradation persisted, suggesting that ampicillin alone is not the sole cause of protein degradation. Expression levels of both TnpA-AirID and AirID alone decreased after 1 hour. These observations could be due to both the elimination of donor cells by ampicillin treatment and transient expression resulting from the loss of the expression plasmids after conjugation. In absence of ampicillin, the transient expression of TnpA-AirID could be the result of the plasmid tranfer into the recipient cell followed by the loss of the plasmid due to replication failure. The low expression background could then result from the presence of donor cells, in low quantity compared to recipient cells. Collectively, these results suggest a conjugation-dependent degradation process.

Notably, fusion proteins containing the miniTurbo biotin ligase exhibited high instability, with only one band corresponding to degradation product being seen under the conditions tested, suggesting that the stability of fusion protein is dependent on the type of biotin ligase used. AirID fusion proteins demonstrated expression and stability for 2 hours and 30 minutes under conjugation conditions, making it the preferred choice for further mass spectrometry analysis. Conversely, miniTurbo suggested protein instability and/or cell toxicity, leading to its exclusion from further study. The instability of the miniTurbo fusion protein has been observed in other organisms (May et al. 2020).

IV.5 Protein enrichment and analysis of mass spectrometry data

To enrich biotinylated proteins, we employed cell lysis followed by purification on magnetic beads conjugated to streptavidin. Verification through western blot analysis demonstrated that both the non-soluble and soluble fractions of cells exhibited similar biotinylation profiles. This crucial experiment ensured that we did not inadvertently exclude biotinylated proteins present in the non-soluble fraction but not in the soluble fraction. While the profiles after lysis resembled the elution profiles, we observed variations in the propensity of certain bands to bind to the beads. This indicates that the binding efficiency may differ depending on the protein, and given our bead saturation strategy, the elution might contain proteins in a different ratio than in the original cell lysates. This approach was adopted to maximize the sample quantity sent for mass spectrometry analysis, as the amount is constrained by the quantity of beads. Further refinement in mass spectrometry analysis may benefit from using beads in excess.

Both test and negative control were sent for analysis without normalization by protein quantity, and the absence of replicates added complexity to result quantification. To address this challenge, we utilized riBAQ to determine which proteins were enriched in the positive sample (Krey et al. 2014; Lu et al. 2007; Schwanhäusser et al. 2011). Additionally, we analyzed mass spectrometry data from a qualitative perspective, selecting proteins based on their function and conservation within bacteria. Notably, some proteins enriched in the test sample are known to play roles in the transposition mechanisms of other transposable elements. For instance, the Insertion Host Factor (IHF) is recognized for its involvement in the transposition of Mu (Harshey 2014) and Tn4652 (Ilves et al. 2004), a Tn3-family member.

IV.6 Testing protein-protein interactions by bacterial-two-hybrid

Given that one of the primary attributes of the proximity labelling approach is its capacity to identify the proximity of a protein rather than merely detecting its interactants (Qin et al. 2021a; Varnaitė and MacNeill 2016), it was imperative for us to distinguish genuine interactants of TnpA from other proteins. To address this, we employed the bacterial two-hybrid assay, a method offering the advantage of in vivo testing, particularly within the *E. coli* context. This approach enabled the examination of TnpA interactions within its cellular environment.

The current results, unfortunately, were inconclusive primarily due to issues related to reproducibility. Throughout the experiment, which involved the transformation of two constructs containing a protein and the T25 or T18 fragment into a BTH101 strain, significant variability was observed. A factor contributing to variability was the inconsistent time required for cells to exhibit signals after being spotted on McConkey plates. The duration before obtaining a positive signal was notably erratic, dependent on the competent cells used where basal levels of reporter gene expression could be variable. Furthermore, over time, even negative controls began to show signals. Such challenges, including high rates of false positives, are well-known within two-hybrid methods (Mehla et al. 2017) and necessitate additional adjustments to enhance the reliability and reproducibility of results.

V. Material and Methods

Plasmids and strains

The genotype of the various strains used is shown in Table **V.1**, along with their antibiotic-resistances. Table **V.2** summarises all the plasmids that were used.

Table V.1 – Strains used in this thesis, with antibiotic resistances and genotype.

Strain	Antibiotic resistance	Genotype
MG1655	/	E. coli K12 wild-type F- λ- rph-1
MFD*pir*	ApraR ZeoR ErmR	MG1655 RP4-2-Tc::[ΔMu1::aac(3)IV-ΔaphA-Δnic35-ΔMu2::zeo] ΔdapA::(erm-pir) ΔrecA
TOP 10	SmR	F–mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK λ–rpsL(StrR) endA1 nupG
BTH101	SmR	F cya-99 araD139 galE15 galK16 rpsL1 (Strr) hsdR2 mcrA1 mcrB1

Table V.2 – Plasmids used in this thesis, with antibiotic resistances and genotype.

Plasmid	Antibiotic resistance	Origin of replication	Features
pRS415	AmpR	ColE1	TI4-promoterless lacZYA operon, pBR322-based
pCS05	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, tnpA::TurboID under *pLac-Oc* and TnpI gene under *pLac*
pCS06	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, TurboID under *pLac-Oc* and TnpI gene under *pLac*
pCS07	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, tnpA::miniTurbo under *pLac-Oc* and TnpI gene under *pLac*
pCS08	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, miniTurbo under *pLac-Oc* and TnpI gene under *pLac*
pCS09	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, tnpA::AirID under *pLac-Oc* and TnpI gene under *pLac*
pCS10	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, AirID under *pLac-Oc* and TnpI gene under *pLac*
pCS09FLAG	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, tnpA::AirID::FLAG under *pLac-Oc* and TnpI gene under *pLac*
PCS10FLAG	CmR KanR	R6Kγ	pSW23T derivative with mini-Tn*Km* cassette, tnpA::AirID::FLAG under *pLac-Oc* and TnpI gene under *pLac*

Cell growth conditions

The only media used for cell growth is lysogeny broth (LB). It is composed of 10 g/L of tryptone, 5 g/L of yeast extract and 5 g/L of sodium chloride. Volume is then adjusted with demineralised water (dH_2O) and sterilised. Solid LB plates are prepared using the same recipe, with the addition of 20 g/L of agar-agar. Once sterile, the medium is cooled down at room temperature and then can be poured into petri dishes. Once they have solidified and inoculated, the plates can be incubated at the desired temperature.

Antibiotics

Antibiotic resistances genes were used as selection marker for cell cultures and experiments. Stock solutions were prepared with a concentration a thousand times superior to the final concentration. These solutions were sterilized by filtration through a 0,45 μm filter before utilization and were stocked at -20°C. Antibiotics, solvents and final concentrations are described in Table **V.3**.

Table V.3 – Final concentration of antibiotics and their appropriate solvent.

Antibiotic	Solvent	Final concentration
Ampicillin	Demineralized water	250μg/mL
Kanamycin	Demineralized water	25μg/mL
Chloramphenicol	Ethanol	25μg/mL

Supplements to growth medium

Since MFDpir is a diaminopimelic acid (DAP) auxotroph, cells were cultured with the addition of 57 μg/mL of DAP. Stock solution was concentrated hundred times by dissolving 57 mg of DAP in 10mL of demineralized water and sterilized by filtration on a 0,45 μm filter. 1000X Concentrated stock solution of biotin was prepared by adding 49 mg/mL of biotin in DMSO. Bacterial-two-hybrid experiment required addition of IPTG (final concentration of 0.5 mM), X-Gal (40 μg/mL) and/or maltose (1% final concentration). These components were added either directly in liquid cultures or solid medium (melted) after sterilization.

Genetic constructions with TnpA and biotin ligases

Constructions containing fusion proteins of AirID were generated by amplifying the backbone of existing plasmids. We amplified the AirID fragment from a plasmid given by Yoann Santin from Geraldine Laloux's lab at de Duve Institute. Polymerase chain reactions (PCR) were realized using Q5 high fidelity DNA polymerase and primers described in Table **V.4**. PCR products were cleaned up with the Monarch PCR & DNA cleanup kit and ligated by Gibson assembly.

Table V.4 – List of the primers used for proximity labelling constructions.

Name	Sequence 5'-3'
CS-AirIDnoTnpA-F	ATTTCACACAGGAAACAGCCATGAAGGACAACACCGTTCC
CS-AirIDTnpA-F	TAAAAGAGCCGTTTTATTCTATGAAGGACAACACCGTTCC
CS-GibF-pOD18(2)	GAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGT
CS-GibR-pOD18(1)	ACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTC
CS-pCS09(1)-F	TCAGCCTCAGATCTGCTTAATATAACGGCTCTTTTTATAGAAAAAAT
CS-pCS09(2)-R	GGAACGGTGTTGTCCTTCATAGAATAAAACGGCTCTTTTATACG
CS-pCS10(2)-R	GGAACGGTGTTGTCCTTCATGGCTGTTTCCTGTGTGAAAT
CS-AirID-R	CTATAAAAAGAGCCGTTATATTAAGCAGATCTGAGGCTGA

Electro-competent cells

A single colony from the desired strain to be made electro-competent is inoculated into 10 mL of LB medium, with the appropriate antibiotic, and incubated overnight at 37°C. 1 L of LB broth is inoculated with this pre-culture and shaken at 37°C until the culture reaches an optical density measured at a wavelength of 600 nm (OD_{600}) of 0.5-0.7. Immediately, the culture is placed on ice for at least 15 min. From this point on, the cells need to be kept ice cold for the remainder of the procedure. The cells are then harvested by centrifugation at 4000 rpm during 10 min, at 4°C. The supernatant is removed, and the pellet is resuspended in a total volume of 1L of cold and sterile water. Again, the cells are centrifuged at 4000 rpm during 10 min, at 4°C. The supernatant is decanted, and 500 mL of cold and sterile water is added to completely suspend the pellet. After another centrifugation of 10 min at 4000 rpm and at 4°C, the pellet is resuspended in 20 mL of cold 10% glycerol. The cells are finally harvested one last time by centrifugation at 4000 rpm during 10 min at 4°C. The supernatant is poured off and 2 mL of cold 10% glycerol is added to the pellet. After resuspension, small aliquots can be made (volume depends on further use) and the microtubes are stored at -80°C.

Electroporation

The electro-competent cells are thawed on ice for at least 15 min (or freshly prepared electro-competent may be used). 1 or 2 µL of plasmid, depending on the conditions, are added along the wall of the cold electroporation cuvette. 50 µL are then transferred on top of the plasmid. The cuvette is chilled for 15 min on ice. The excess moisture present at the exterior of the cuvette is wiped and the cuvette is gently tapped on the countertop a couple of times to move cells to the bottom. The cuvette is placed in the electroporator, and pulse is pressed (program *E. coli* 2 mm - 2.5V). Immediately, 450 µL of LB are added to the cuvette to prevent cells from dying off. The mixture is transferred to an Eppendorf tube and incubated at 37°C with agitation for an hour, to allow expression of antibiotic resistance gene. Transformation product is plated on agar plates supplemented with the appropriate antibiotic. These are then incubated overnight at 37°C.

Conjugation assay

Donor and recipient strain were inoculated overnight in LB at 37°C with according antibiotic. The next day, OD_{600} of cultures diluted 5X were measured after and quantities engaged for the conjugation were calculated by using the following formula:

$$\text{Cell culture volume (mL)} = \text{Cell quantity*} / OD_{600}\ \text{Value}$$

*Cell quantity correspond either to 0.5 for recipient cells or 0.05 for donor cells.

Both donor and recipient cells were pulled together in a 1.5 mL Eppendorf tube and centrifugated at 8000 rpm during 3'. Supernatant was removed, and cells were resuspended in 1 mL of fresh LB. Cells were centrifugated a second time at 8000 rpm during 3'. Supernatant was discarded to leave a volume of 50 µL to resuspend the pellet. Controls at 0' of conjugation were added 1 mL of LB, 750 µL were used for WB analysis and remaining volume was used to streak on plates and assess transposition frequency. For other timepoints, 50 µL of resuspended pellet are spotted on Merck Millipore 0.45µm filters in petri dishes containing LB agar, 50 µM of biotin and 250µg/mL of ampicillin. Cells were then incubated at 37°C for required time. Filters were then removed from petri dishes and cells were resuspended by putting the filter in a 50mL falcon tube containing 1 mL of fresh LB and vortexing for 30 seconds. 750 µL were used for WB analysis and remaining volume for transposition frequency assessment.

Western blot and SDS gel for cell lysates after conjugation

A volume of 750 µL of cell resuspension after or before (0' controls) conjugation underwent centrifugation at 12000 rpm for 1 minute and was then resuspended in 50µL of sterile H_2O after discarding the supernatant. The samples were stored at -20°C until utilized for Western blot (WB) or SDS. Before loading onto gels, 12 µL of a crack solution was added to the initial volume of 50 µL. For western blot analysis, gels were transferred onto a nitrocellulose membrane, washed with TBS twice, and blocked with a solution of 3% BSA for 1 hour. Prior to incubation with streptavidin AP, the membrane was quickly washed with TBS and TBS-Tween-Triton. After washing, the membrane was incubated with streptavidin AP (diluted 1:1250 in 3% BSA) away from light for 90 minutes. Finally, the membrane was washed three times with TBS-Tween-Triton and rinsed with H_2O before being revealed with 5 mL of MB purple.

Transposition frequency

To calculate transposition frequency, quantity of donor cells and recipient cells which gained kanamycin resistance was determined. To count donor cells, we used the same protocol as in conjugation assay and when cells were pulled and centrifugated we resuspended them in 1 mL of LB. We diluted this solution 1000X and used 50µ L of it to streak on a plate containing 57 µg/mL of DAP and 25 µg/mL of chloramphenicol. Recipient cells in which transposition occurred were counted by following the conjugation assay until cells present on the filter are resuspended in 1mL of LB. This solution was diluted 10X before being streaked on ampicillin and kanamycin plate. Transposition frequency is the calculated by suing the following formula:

$$\text{Transposition frequency} = \frac{\#\text{ Recipient cells with miniTn (LB + Amp + Kan)}}{\#\text{ Donor cells (LB + DAP + Chl)}}$$

Characterisation of transposition events

Colonies that grew on kanamycin and ampicillin after conjugation were subcultured onto two separate plates, one containing ampicillin and kanamycin, and the other containing ampicillin and chloramphenicol. Twenty colonies were cultured overnight in fresh LB for plasmid extraction using the Eurogentec Smartpure DNA purification kit. Plasmids extracted from these cultures underwent restriction analysis with NcoI and KpnI, along with back transformation into the TOP10 strain.

Western blot of fusion proteins

To test the stability of the constructs, 750 µL of cell resuspension are centrifugated and resuspended in 50 µL of sterile H_2O and 12 µL of sample buffer are added before loading 25µL on gel. Gel is transferred on a nitrocellulose membrane and rinsed twice with TBS for 5 minutes. Membrane was blocked for 3 hours with a solution of 1% casein in TBS before rinsing with TBS. We then incubated the membrane with anti-FLAG antibody in 1% casein overnight. Primary antibody was washed twice with TBS-Tween-Triton and once with TBS. Secondary antibody anti-mouse is incubated for 90 minutes and washed rigorously with TBS-Tween-Triton before reveal.

Protein enrichment

5 filters were taken after conjugation and cells were resuspended in 1 mL of fresh LB in a 50 mL falcon tube. 750 μL of it was measured and centrifugated during 1 min at 12000 rpm. Remaining volume was used to control transposition frequency. The pellet was then resuspended in 1 mL of lysis buffer (10 mM Tris pH 7.4, 100 mM NaCl, 10% glycerol). Cells are then lysed in a 2 mL tube with 300 μL of glass beads, on ice. Tubes were then shacked in fast prep at 6 m/s during 30'' twice with 5' of rest time. Cell lysates were centrifugated (13000 rpm, 15' at 4°C). Supernatant was centrifugated a second time with same settings. After centrifugation, lysates were dialysed using cassettes of 3000-4500 MWCO during two hours in 1 L of lysis buffer. Buffer was then changed before overnight dialysis. Magnetic beads were prepared by resuspending them and transferring 20 μL of beads to a tube with 1mL of binding buffer (10 mM Tris pH 7.4, 100 mM NaCl, 0.01% tween 20) and vortexing 5 seconds. Supernatant is then discarded by placing the tube on a magnet and beads are resuspended in 20 μL of binding buffer. 50 μL of cell lysates were put aside as pre-binding control and protein concentration was measured. Remaining volume of cell lysate is incubated with the washed beads during 90 minutes at room temperature. 50 μL of supernatant were put aside as post-binding control, and the protein concentration was also measured. Tubes were placed on the magnet to remove supernatant. Beads were then washed 3 times in 1mL of washing buffer (20 mM Tris pH 8, 2 mM CaCl2) and 50 μL of each wash was kept for control.

Mass spectrometry

Mass spectrometry analysis was carried at VIB-KU Leuven Center for Brain and Disease research.

Bacterial-two-hybrid constructions

Plasmids for *in vivo* expression of fusion proteins between TnpA and adenylate cyclase fragments (T18 and T25) were constructed using standard molecular biology techniques (see above) and with primers presented in Table **V.5**.

Table V.5 – List of the primers used for proximity labelling constructions.

Name	Sequence 5'-3'
CS-HU-A pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGAACAAGACTCAACTGATTGATG
CS-HU-A pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTAACTGCGTCTTTCAGTGCCTT
CS-HU-B pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGGTGAATAAATCTCAATTGATCGACAAG
CS-HU-B pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTAGTTTACCGCGTCTTTCAG
CS-allR pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGACGGAAGTTAGACGGC
CS-allR pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTATGGATGTGCTTTCAGTCCC
CS-YjhU pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGGTGGATAGAGATCAAACAAACAGC
CS-YjhU pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTAGTGTGGTAATTCACACGC
CS-rplE pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGCGAAACTGCATGATTAC
CS-rplE pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTACTTGCGGAACGGGAA
CS-himA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGCGCTTACAAAAGCTG
CS-himA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTACTCGTCTTTGGGCGAA
CS-YdiZ pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGCCAGTGGCGATC
CS-YdiZ pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTATTTTCTGTTGCTATTCCATTCCT
CS-tufA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGGTGTCTAAAGAAAAATTTGAACGTACAA
CS-tufA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTAGCCCAGAACTTTAGCAAC
CS-yfaA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGAGTGGTGAAAAAAAGGCG
CS-yfaA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTCATAGCGGCTGCCAG
CS-cspC pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGCAAAGATTAAAGGTCAGG
CS-cspC pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTCAGATAGCTGTTACGTTAACAGC
CS-gntR pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGAAAAAGAAAAGACCCGTACTT
CS-gntR pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTAAATAGATCCGCCCGGTG
CS-tufB pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGTCTAAAGAAAAGTTTGAACGTACA
CS-tufB pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTAGCTCAGAACTTTTGCTACAAC
CS-ftsA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGATCAAGGCGACGGAC
CS-ftsA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTAAAACTCTTTTCGCAGCCA
CS-minE pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGCATTACTCGATTTCTTTCTC
CS-minE pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTATTTCAGCTCTTCTGCTTCC
CS-recO pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGAAGGCTGGCAGC
CS-recO pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTCATTCATAATGTGTTTTCACCGT
CS-rpsS pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGCCACGTTCTCTCAAGAA
CS-rpsS pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTATTTCTTCTTCGCTTTTTTATCAGC
CS-yjgR pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGAGTGAACCCCTGTTAATTG
CS-yjgR pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTACCTTCTTCTCCCCCCC
CS-rpsT pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGTTGGCTAATATCAAATCAGCTAAGAAG
CS-rpsT pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTACTTTTAAGCCAGTTTGTTGATCTGTG
CS-helD pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGAACTGAAAGCGACAAC
CS-helD pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTACGGTTTTCTCGCCAC
CS-yejK pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGAGTCTGGATATCAACCAGATT
CS-yejK pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTAATTGCCGCCAGACG

CS-fis pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGTTCGAACAACGCGTAAA
CS-fis pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTAGTTCATGCCGTATTTTTTCAAT
CS-lrp pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGTAGATAGCAAGAAGCGC
CS-lrp pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTAGCGCGTCTTAATAACCAG
CS-recA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGCTATCGACGAAAACAAA
CS-recA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTAAAAATCTTCGTTAGTTTCTGCTAC
CS-hns pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGAGCGAAGCACTTAAAATTCT
CS-hns pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTATTGCTTGATCAGGAAATCGT
CS-topA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGGTAAAGCTCTTGTCATC
CS-topA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTATTTTTTTCCTTCAACCCATTTGC
CS-stpA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGTCCGTAATGTTACAAAGTTTAAA
CS-stpA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTAGATCAGGAAATCGTCGAGA
CS-cbpA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGAATTAAAGGATTATTACGCCA
CS-cbpA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTCAAAAGACGACTGGGGC
CS-polA pKT25 F	GCAGGGTCGACTCTAGAGGATCCCCGGATGGTTCAGATCCCCCAAAA
CS-polA pKT25 R	CGAATTCTTAGTTACTTAGGTACCCTTTTAGTGCGCCTGATCCC

Chemically competent cells

A single colony from the BTH101 strain was inoculated into 10 mL of LB medium and incubated overnight at 37°C. 1 L of LB is inoculated with this pre-culture and shaken at 37°C until the culture reaches mid-log phase. Cells are then harvested (2500 rpm during 20 min at 4°C) and resuspended in 100 mL of ice-cold 100 mM $MgCl_2$. Once resuspended cells are incubated on ice for 30 minutes. Once again, cells are harvested (2500 rpm during 20 min at 4°C) and resuspended in 20 mL of $CaCl_2$ 100 mM, 15% glycerol. After resuspension, small aliquots are made microtubes are stored at -80°C.

Chemical transformation for bacterial-two-hybrid

Chemically competent cells were thawed on ice. 200 µL of cells were mixed with one plasmid encoding a T25 fusion and the other coding for the T18 fusion. In parallel, positive, and negative controls were also co-transformed. Addition of plasmids containing T25 and T18 fragment fused to TnpA served as positive control (dimerization of TnpA) and addition of plasmids containing both T25 and T18 unfused served as a negative control. Cells were incubated in ice for 30 minutes. Cells (mixed with DNA) are then heat-shocked at 42°C during 90 seconds before returning them to ice and adding 200 µL of fresh LB. After 1 hour of incubation at 37°C, 200 µL of cells are plated on petri dishes containing the appropriate selection markers.

Screening of interacting partners

Multiple independent colonies from each set of transformation were used to inoculate 500 μL of sterile LB containing 250 μg/mL ampicillin, 25 μg/mL kanamycin and 0.5 mM IPTG in 1.5 mL tubes. Cells were incubated for 24 hours at 30°C with strong agitation. Triplicates for each pair of plasmids were made. After overnight incubation, 2μL of each culture was dropped on either MacConkey plates (250 μg/mL ampicillin, 25 μg/mL kanamycin, 0.5 mM IPTG, and 1% maltose) or LB + X-Gal plates (250 μg/mL ampicillin, 25 μg/mL kanamycin, 0.5 mM IPTG, and 40 μg/mL X-Gal). Plates were incubated at 30°C for several days (until apparition of signal).

References

Ayarpadikannan, Selvam, and Heui-Soo Kim. 2014. 'The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases'. *Genomics & Informatics* 12(3):98. doi: 10.5808/GI.2014.12.3.98.

Battesti, Aurélia, and Emmanuelle Bouveret. 2012. 'The Bacterial Two-Hybrid System Based on Adenylate Cyclase Reconstitution in Escherichia Coli'. *Methods* 58(4):325–34. doi: 10.1016/j.ymeth.2012.07.018.

Branon, Tess C., Justin A. Bosch, Ariana D. Sanchez, Namrata D. Udeshi, Tanya Svinkina, Steven A. Carr, Jessica L. Feldman, Norbert Perrimon, and Alice Y. Ting. 2018a. 'Efficient Proximity Labeling in Living Cells and Organisms with TurboID'. *Nature Biotechnology* 36(9):880–87. doi: 10.1038/nbt.4201.

Branon, Tess C., Justin A. Bosch, Ariana D. Sanchez, Namrata D. Udeshi, Tanya Svinkina, Steven A. Carr, Jessica L. Feldman, Norbert Perrimon, and Alice Y. Ting. 2018b. 'Efficient Proximity Labeling in Living Cells and Organisms with TurboID'. *Nature Biotechnology* 36(9):880–87. doi: 10.1038/nbt.4201.

Brosius, J. 1999. 'RNAs from All Categories Generate Retrosequences That May Be Exapted as Novel Genes or Regulatory Elements'. *Gene* 238(1):115–34. doi: 10.1016/S0378-1119(99)00227-9.

Brudvig, J. J., J. T. Cain, G. G. Schmidt-Grimminger, D. J. Stumpo, K. J. Roux, P. J. Blackshear, and J. M. Weimer. 2018. 'MARCKS Is Necessary for Netrin-DCC Signaling and Corpus Callosum Formation'. *Molecular Neurobiology* 55(11):8388–8402. doi: 10.1007/s12035-018-0990-3.

Cervantes, C., K. Chávez, and S. Vaca. 1991. '[Mechanisms of Bacterial Resistance to Heavy Metals].' *Revista Latinoamericana de Microbiologia* 33(1):61–70.

Chapman-Smith, Anne, and John E. Cronan. 1999. 'The Enzymatic Biotinylation of Proteins: A Post-Translational Modification of Exceptional Specificity'. *Trends in Biochemical Sciences* 24(9):359–63. doi: 10.1016/S0968-0004(99)01438-3.

Chen, Chiao-Lin, Yanhui Hu, Namrata D. Udeshi, Thomas Y. Lau, Frederik Wirtz-Peitz, Li He, Alice Y. Ting, Steven A. Carr, and Norbert Perrimon. 2015. 'Proteomic Mapping in Live *Drosophila* Tissues Using an Engineered Ascorbate Peroxidase'. *Proceedings of the National Academy of Sciences* 112(39):12093–98. doi: 10.1073/pnas.1515623112.

Curcio, M. Joan, and Keith M. Derbyshire. 2003. 'The Outs and Ins of Transposition: From Mu to Kangaroo'. *Nature Reviews Molecular Cell Biology* 4(11):865–77. doi: 10.1038/nrm1241.

Deininger, Prescott. 2011. 'Alu Elements: Know the SINEs'. *Genome Biology* 12(12):236. doi: 10.1186/gb-2011-12-12-236.

Demarre, Gaëlle, Anne-Marie Guérout, Chiho Matsumoto-Mashimo, Dean A. Rowe-Magnus, Philippe Marlière, and Didier Mazel. 2005. 'A New Family of Mobilizable Suicide Plasmids Based on Broad Host Range R388 Plasmid (IncW) and RP4 Plasmid (IncPα) Conjugative Machineries and Their Cognate Escherichia Coli Host Strains'. *Research in Microbiology* 156(2):245–55. doi: 10.1016/j.resmic.2004.09.007.

Dingar, Dharmendra, Manpreet Kalkat, Pak-Kei Chan, Tharan Srikumar, Swneke D. Bailey, William B. Tu, Etienne Coyaud, Romina Ponzielli, Max Kolyar, Igor Jurisica, Annie Huang, Mathieu Lupien, Linda Z. Penn, and Brian Raught. 2015. 'BioID Identifies Novel C-MYC Interacting Partners in Cultured Cells and Xenograft Tumors'. *Journal of Proteomics* 118:95–111. doi: 10.1016/j.jprot.2014.09.029.

Dubin, Manu J., Ortrun Mittelsten Scheid, and Claude Becker. 2018. 'Transposons: A Blessing Curse'. *Current Opinion in Plant Biology* 42:23–29. doi: 10.1016/j.pbi.2018.01.003.
Ellabaan, Mostafa M. H., Christian Munck, Andreas Porse, Lejla Imamovic, and Morten O. A. Sommer. 2021. 'Forecasting the Dissemination of Antibiotic Resistance Genes across Bacterial Genomes'. *Nature Communications* 12(1):2435. doi: 10.1038/s41467-021-22757-1.
Fernandez, Maricruz, Alexander V Shkumatov, Yun Liu, Claire Stulemeijer, Sylvie Derclaye, Rouslan G. Efremov, Bernard Hallet, and David Alsteens. 2023. 'AFM-Based Force Spectroscopy Unravels Stepwise Formation of the DNA Transposition Complex in the Widespread Tn3 Family Mobile Genetic Elements'. *Nucleic Acids Research* 51(10):4929–41. doi: 10.1093/nar/gkad241.
Ferrières, Lionel, Gaëlle Hémery, Toan Nham, Anne-Marie Guérout, Didier Mazel, Christophe Beloin, and Jean-Marc Ghigo. 2010. 'Silent Mischief: Bacteriophage Mu Insertions Contaminate Products of *Escherichia Coli* Random Mutagenesis Performed Using Suicidal Transposon Delivery Plasmids Mobilized by Broad-Host-Range RP4 Conjugative Machinery'. *Journal of Bacteriology* 192(24):6418–27. doi: 10.1128/JB.00621-10.
Feschotte, Cédric, and Ellen J. Pritham. 2007. 'DNA Transposons and the Evolution of Eukaryotic Genomes'. *Annual Review of Genetics* 41(1):331–68. doi: 10.1146/annurev.genet.40.110405.090448.
Harshey, Rasika M. 2014. 'Transposable Phage Mu'. *Microbiology Spectrum* 2(5). doi: 10.1128/microbiolspec.MDNA3-0007-2014.
Häsler, Julien, and Katharina Strub. 2006. 'Alu Elements as Regulators of Gene Expression'. *Nucleic Acids Research* 34(19):5491–97. doi: 10.1093/nar/gkl706.
Hedges, R. W., and A. E. Jacob. 1974. 'Transposition of Ampicillin Resistance from RP4 to Other Replicons'. *Molecular and General Genetics MGG* 132(1):31–40. doi: 10.1007/BF00268228.
Hickman, Alison B., and Fred Dyda. 2015a. 'Mechanisms of DNA Transposition'. *Microbiology Spectrum* 3(2). doi: 10.1128/microbiolspec.MDNA3-0034-2014.
Hickman, Alison B., and Fred Dyda. 2015b. 'Mechanisms of DNA Transposition'. *Microbiology Spectrum* 3(2). doi: 10.1128/microbiolspec.MDNA3-0034-2014.
Hof, Arjen E. van't, Pascal Campagne, Daniel J. Rigden, Carl J. Yung, Jessica Lingley, Michael A. Quail, Neil Hall, Alistair C. Darby, and Ilik J. Saccheri. 2016. 'The Industrial Melanism Mutation in British Peppered Moths Is a Transposable Element'. *Nature* 534(7605):102–5. doi: 10.1038/nature17951.
Ilves, Heili, Rita Hõrak, Riho Teras, and Maia Kivisaar. 2004. 'IHF Is the Limiting Host Factor in Transposition of *Pseudomonas Putida* Transposon Tn *4652* in Stationary Phase'. *Molecular Microbiology* 51(6):1773–85. doi: 10.1111/j.1365-2958.2003.03948.x.
Kalocsay, Marian. 2019. 'APEX Peroxidase-Catalyzed Proximity Labeling and Multiplexed Quantitative Proteomics'. Pp. 41–55 in.
Kamruzzaman, Muhammad, Alma Y. Wu, and Jonathan R. Iredell. 2021. 'Biological Functions of Type II Toxin-Antitoxin Systems in Bacteria'. *Microorganisms* 9(6):1276. doi: 10.3390/microorganisms9061276.
Karimova, Gouzel, Josette Pidoux, Agnes Ullmann, and Daniel Ladant. 1998. 'A Bacterial Two-Hybrid System Based on a Reconstituted Signal Transduction Pathway'. *Proceedings of the National Academy of Sciences* 95(10):5752–56. doi: 10.1073/pnas.95.10.5752.

Keenholtz, Ross A., Kent W. Mouw, Martin R. Boocock, Nan-Sheng Li, Joseph A. Piccirilli, and Phoebe A. Rice. 2013. 'Arginine as a General Acid Catalyst in Serine Recombinase-Mediated DNA Cleavage'. *Journal of Biological Chemistry* 288(40):29206–14. doi: 10.1074/jbc.M113.508028.

Keenholtz, Ross A., Sally-J. Rowland, Martin R. Boocock, W. Marshall Stark, and Phoebe A. Rice. 2011. 'Structural Basis for Catalytic Activation of a Serine Recombinase'. *Structure* 19(6):799–809. doi: 10.1016/j.str.2011.03.017.

Khan, Madiha, Ji-Young Youn, Anne-Claude Gingras, Rajagopal Subramaniam, and Darrell Desveaux. 2018. 'In Planta Proximity Dependent Biotin Identification (BioID)'. *Scientific Reports* 8(1):9212. doi: 10.1038/s41598-018-27500-3.

Kido, Kohki, Satoshi Yamanaka, Shogo Nakano, Kou Motani, Souta Shinohara, Akira Nozawa, Hidetaka Kosako, Sohei Ito, and Tatsuya Sawasaki. 2020a. 'AirID, a Novel Proximity Biotinylation Enzyme, for Analysis of Protein–Protein Interactions'. *ELife* 9. doi: 10.7554/eLife.54983.

Kido, Kohki, Satoshi Yamanaka, Shogo Nakano, Kou Motani, Souta Shinohara, Akira Nozawa, Hidetaka Kosako, Sohei Ito, and Tatsuya Sawasaki. 2020b. 'AirID, a Novel Proximity Biotinylation Enzyme, for Analysis of Protein–Protein Interactions'. *ELife* 9. doi: 10.7554/eLife.54983.

Krey, Jocelyn F., Phillip A. Wilmarth, Jung-Bum Shin, John Klimek, Nicholas E. Sherman, Erin D. Jeffery, Dongseok Choi, Larry L. David, and Peter G. Barr-Gillespie. 2014. 'Accurate Label-Free Protein Quantitation with High- and Low-Resolution Mass Spectrometers'. *Journal of Proteome Research* 13(2):1034–44. doi: 10.1021/pr401017h.

Lambin, Michaël, Emilien Nicolas, Cédric A. Oger, Nathan Nguyen, Deborah Prozzi, and Bernard Hallet. 2012. 'Separate Structural and Functional Domains of Tn4430 Transposase Contribute to Target Immunity'. *Molecular Microbiology* 83(4):805–20. doi: 10.1111/j.1365-2958.2012.07967.x.

Lereclus, Didier, Jacques Mahillon, Ghislaine Menou, and Marguerite-M. Lecadet. 1986. 'Identification of Tn4430, a Transposon of Bacillus Thuringiensis Functional in Escherichia Coli'. *Molecular and General Genetics MGG* 204(1):52–57. doi: 10.1007/BF00330186.

Li, Jiefu, Shuo Han, Hongjie Li, Namrata D. Udeshi, Tanya Svinkina, D. R. Mani, Chuanyun Xu, Ricardo Guajardo, Qijing Xie, Tongchao Li, David J. Luginbuhl, Bing Wu, Colleen N. McLaughlin, Anthony Xie, Pornchai Kaewsapsak, Stephen R. Quake, Steven A. Carr, Alice Y. Ting, and Liqun Luo. 2020. 'Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators'. *Cell* 180(2):373-386.e15. doi: 10.1016/j.cell.2019.12.029.

Liu, Xiaonan, Kari Salokas, Rigbe G. Weldatsadik, Lisa Gawriyski, and Markku Varjosalo. 2020. 'Combined Proximity Labeling and Affinity Purification–mass Spectrometry Workflow for Mapping and Visualizing Protein Interaction Networks'. *Nature Protocols* 15(10):3182–3211. doi: 10.1038/s41596-020-0365-x.

Lu, Peng, Christine Vogel, Rong Wang, Xin Yao, and Edward M. Marcotte. 2007. 'Absolute Protein Expression Profiling Estimates the Relative Contributions of Transcriptional and Translational Regulation'. *Nature Biotechnology* 25(1):117–24. doi: 10.1038/nbt1270.

Mahillon, J., and D. Lereclus. 1988. 'Structural and Functional Analysis of Tn4430: Identification of an Integrase-like Protein Involved in the Co-Integrate-Resolution Process.' *The EMBO Journal* 7(5):1515–26. doi: 10.1002/j.1460-2075.1988.tb02971.x.

Mair, Andrea, Shou-Ling Xu, Tess C. Branon, Alice Y. Ting, and Dominique C. Bergmann. 2019. 'Proximity Labeling of Protein Complexes and Cell-Type-Specific Organellar Proteomes in Arabidopsis Enabled by TurboID'. *ELife* 8. doi: 10.7554/eLife.47864.
May, Danielle G., Kelsey L. Scott, Alexandre R. Campos, and Kyle J. Roux. 2020. 'Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation'. *Cells* 9(5):1070. doi: 10.3390/cells9051070.
Mehla, J., J. H. Caufield, N. Sakhawalkar, and P. Uetz. 2017. 'A Comparison of Two-Hybrid Approaches for Detecting Protein–Protein Interactions'. Pp. 333–58 in.
Morales, Grace, Benjamin Abelson, Seth Reasoner, Jordan Miller, Ashlee M. Earl, Maria Hadjifrangiskou, and Jonathan Schmitz. 2023. 'The Role of Mobile Genetic Elements in Virulence Factor Carriage from Symptomatic and Asymptomatic Cases of Escherichia Coli Bacteriuria'. *Microbiology Spectrum* 11(3). doi: 10.1128/spectrum.04710-22.
Nesmelova, Irina V., and Perry B. Hackett. 2010. 'DDE Transposases: Structural Similarity and Diversity'. *Advanced Drug Delivery Reviews* 62(12):1187–95. doi: 10.1016/j.addr.2010.06.006.
Nicolas, Emilien, Michael Lambin, Damien Dandoy, Christine Galloy, Nathan Nguyen, Cédric A. Oger, and Bernard Hallet. 2015. 'The Tn *3* -Family of Replicative Transposons'. *Microbiology Spectrum* 3(4). doi: 10.1128/microbiolspec.MDNA3-0060-2014.
Nicolas, Emilien, Cédric A. Oger, Nathan Nguyen, Michaël Lambin, Amandine Draime, Sébastien C. Leterme, Michael Chandler, and Bernard F. J. Hallet. 2017a. 'Unlocking Tn3-Family Transposase Activity in Vitro Unveils an Asymetric Pathway for Transposome Assembly'. *Proceedings of the National Academy of Sciences* 114(5). doi: 10.1073/pnas.1611701114.
Nicolas, Emilien, Cédric A. Oger, Nathan Nguyen, Michaël Lambin, Amandine Draime, Sébastien C. Leterme, Michael Chandler, and Bernard F. J. Hallet. 2017b. 'Unlocking Tn3-Family Transposase Activity in Vitro Unveils an Asymetric Pathway for Transposome Assembly'. *Proceedings of the National Academy of Sciences* 114(5). doi: 10.1073/pnas.1611701114.
Noel, Hannah R., Jessica R. Petrey, and Lauren D. Palmer. 2022. 'Mobile Genetic Elements in *Acinetobacter* Antibiotic-resistance Acquisition and Dissemination'. *Annals of the New York Academy of Sciences* 1518(1):166–82. doi: 10.1111/nyas.14918.
Pasquali, Frédérique, Corinna Kehrenberg, Gerardo Manfreda, and Stefan Schwarz. 2005. 'Physical Linkage of Tn3 and Part of Tn1721 in a Tetracycline and Ampicillin Resistance Plasmid from Salmonella Typhimurium'. *Journal of Antimicrobial Chemotherapy* 55(4):562–65. doi: 10.1093/jac/dkh553.
Peters, Joseph E., and Nancy L. Craig. 2000. 'Tn7 Transposes Proximal to DNA Double-Strand Breaks and into Regions Where Chromosomal DNA Replication Terminates'. *Molecular Cell* 6(3):573–82. doi: 10.1016/S1097-2765(00)00056-3.
Peters, Joseph E., and Nancy L. Craig. 2001. 'Tn7 Recognizes Transposition Target Structures Associated with DNA Replication Using the DNA-Binding Protein TnsE'. *Genes & Development* 15(6):737–47. doi: 10.1101/gad.870201.
Pourrajab, Fatemeh, and Seyedhossein Hekmatimoghaddam. 2021. 'Transposable Elements, Contributors in the Evolution of Organisms (from an Arms Race to a Source of Raw Materials)'. *Heliyon* 7(1):e06029. doi: 10.1016/j.heliyon.2021.e06029.

Prakash, Om, and Max A. Eisenberg. 1974. 'Active Transport of Biotin in *Escherichia Coli* K-12'. *Journal of Bacteriology* 120(2):785–91. doi: 10.1128/jb.120.2.785-791.1974.
Qin, Wei, Kelvin F. Cho, Peter E. Cavanagh, and Alice Y. Ting. 2021a. 'Deciphering Molecular Interactions by Proximity Labeling'. *Nature Methods* 18(2):133–43. doi: 10.1038/s41592-020-01010-5.
Qin, Wei, Kelvin F. Cho, Peter E. Cavanagh, and Alice Y. Ting. 2021b. 'Deciphering Molecular Interactions by Proximity Labeling'. *Nature Methods* 18(2):133–43. doi: 10.1038/s41592-020-01010-5.
Reinke, Aaron W., Raymond Mak, Emily R. Troemel, and Eric J. Bennett. 2017. 'In Vivo Mapping of Tissue- and Subcellular-Specific Proteomes in *Caenorhabditis Elegans*'. *Science Advances* 3(5). doi: 10.1126/sciadv.1602426.
Rosta, Edina, Wei Yang, and Gerhard Hummer. 2014. 'Calcium Inhibition of Ribonuclease H1 Two-Metal Ion Catalysis'. *Journal of the American Chemical Society* 136(8):3137–44. doi: 10.1021/ja411408x.
Roux, Kyle J., Dae In Kim, Manfred Raida, and Brian Burke. 2012. 'A Promiscuous Biotin Ligase Fusion Protein Identifies Proximal and Interacting Proteins in Mammalian Cells'. *Journal of Cell Biology* 196(6):801–10. doi: 10.1083/jcb.201112098.
Samavarchi-Tehrani, Payman, Reuben Samson, and Anne-Claude Gingras. 2020. 'Proximity Dependent Biotinylation: Key Enzymes and Adaptation to Proteomics Approaches'. *Molecular & Cellular Proteomics* 19(5):757–73. doi: 10.1074/mcp.R120.001941.
Santin, Yoann G., Thierry Doan, Régine Lebrun, Leon Espinosa, Laure Journet, and Eric Cascales. 2018. 'In Vivo TssA Proximity Labelling during Type VI Secretion Biogenesis Reveals TagA as a Protein That Stops and Holds the Sheath'. *Nature Microbiology* 3(11):1304–13. doi: 10.1038/s41564-018-0234-3.
Schwanhäusser, Björn, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. 2011. 'Global Quantification of Mammalian Gene Expression Control'. *Nature* 473(7347):337–42. doi: 10.1038/nature10098.
Shapiro, James A. 1979. 'Molecular Model for the Transposition and Replication of Bacteriophage Mu and Other Transposable Elements'. *Proceedings of the National Academy of Sciences of the United States of America* 76(4):1933–37.
Shinoda, Natsuki, Nozomi Hanawa, Takahiro Chihara, Akiko Koto, and Masayuki Miura. 2019. 'Dronc-Independent Basal Executioner Caspase Activity Sustains *Drosophila* Imaginal Tissue Growth'. *Proceedings of the National Academy of Sciences* 116(41):20539–44. doi: 10.1073/pnas.1904647116.
Shkumatov, Alexander V., Nicolas Aryanpour, Cédric A. Oger, Gérôme Goossens, Bernard F. Hallet, and Rouslan G. Efremov. 2022. 'Structural Insight into Tn3 Family Transposition Mechanism'. *Nature Communications* 13(1):6155. doi: 10.1038/s41467-022-33871-z.
Tian, Dongxing, Mingqi Zhao, Sasa Zheng, Xiaofei Jiang, and Bin Zhang. 2023. 'Involvement of Tn3 Transposon in Formation and Transmission of Hypervirulent and Carbapenem-Resistant *Klebsiella Pneumoniae*'. *Microbiology Spectrum* 11(6). doi: 10.1128/spectrum.03038-23.
Tomilin, Nikolai V. 1998. 'Control of Genes by Mammalian Retroposons'. Pp. 1–48 in.
Ton-Hoang, Bao, Cécile Pasternak, Patricia Siguier, Catherine Guynet, Alison Burgess Hickman, Fred Dyda, Suzanne Sommer, and Michael Chandler. 2010. 'Single-Stranded DNA Transposition Is Coupled to Host Replication'. *Cell* 142(3):398–408. doi: 10.1016/j.cell.2010.06.034.

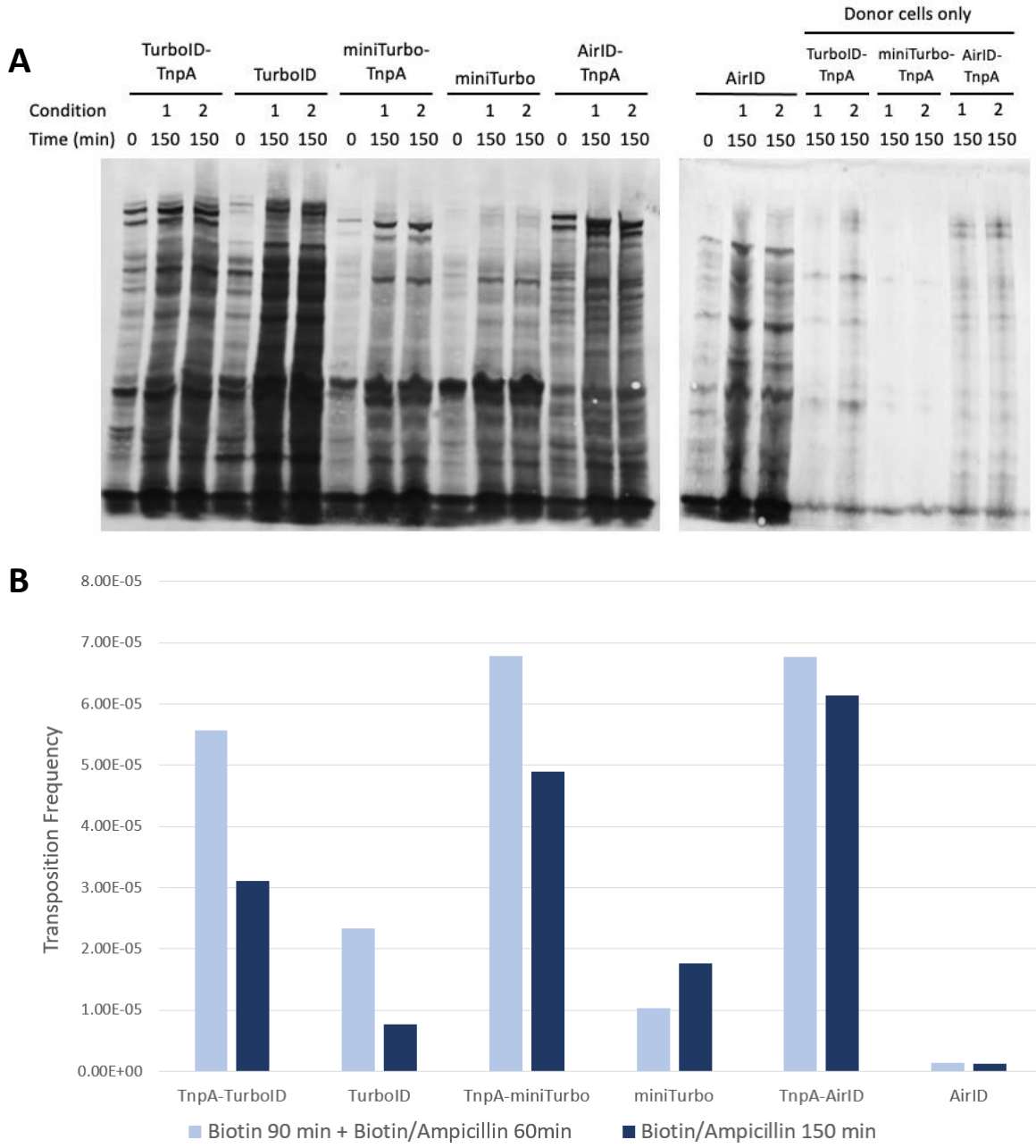
Uezu, Akiyoshi, Daniel J. Kanak, Tyler W. A. Bradshaw, Erik J. Soderblom, Christina M. Catavero, Alain C. Burette, Richard J. Weinberg, and Scott H. Soderling. 2016. 'Identification of an Elaborate Complex Mediating Postsynaptic Inhibition'. *Science* 353(6304):1123–29. doi: 10.1126/science.aag0821.

Varnaitė, Renata, and Stuart A. MacNeill. 2016. 'Meet the Neighbors: Mapping Local Protein Interactomes by Proximity-dependent Labeling with BioID'. *PROTEOMICS* 16(19):2503–18. doi: 10.1002/pmic.201600123.

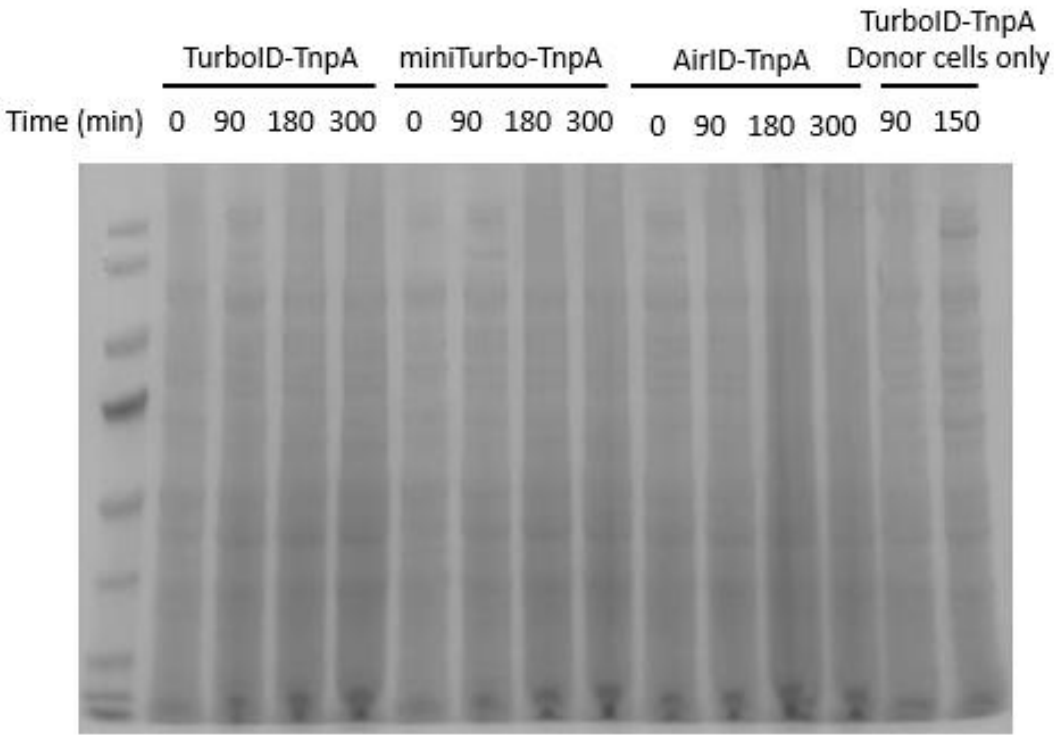
Watt, V. M., C. J. Ingles, M. S. Urdea, and W. J. Rutter. 1985. 'Homology Requirements for Recombination in Escherichia Coli.' *Proceedings of the National Academy of Sciences* 82(14):4768–72. doi: 10.1073/pnas.82.14.4768.

Xu, Yangfan, Xianqun Fan, and Yang Hu. 2021. 'In Vivo Interactome Profiling by Enzyme-catalyzed Proximity Labeling'. *Cell & Bioscience* 11(1):27. doi: 10.1186/s13578-021-00542-3.

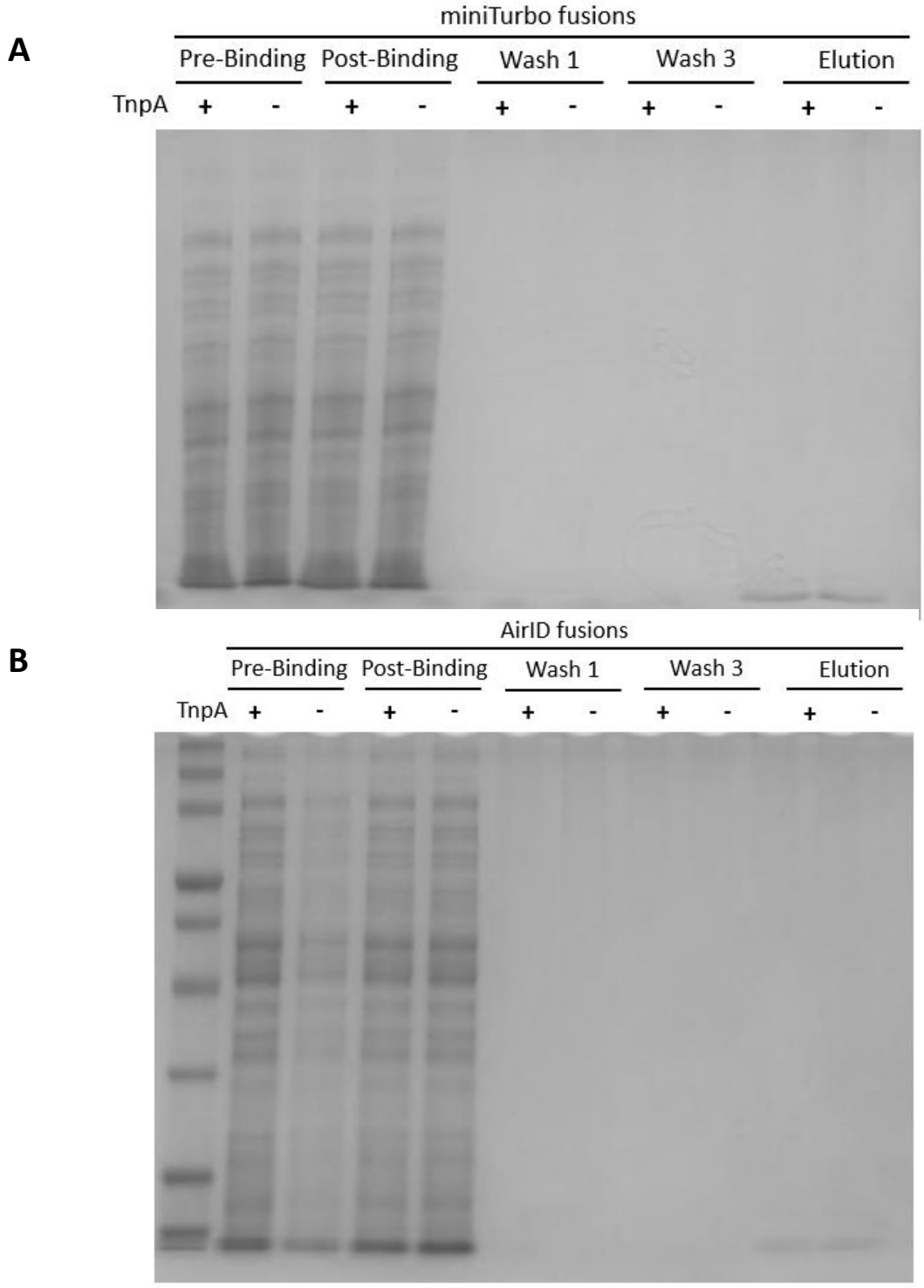
Supplementary figures



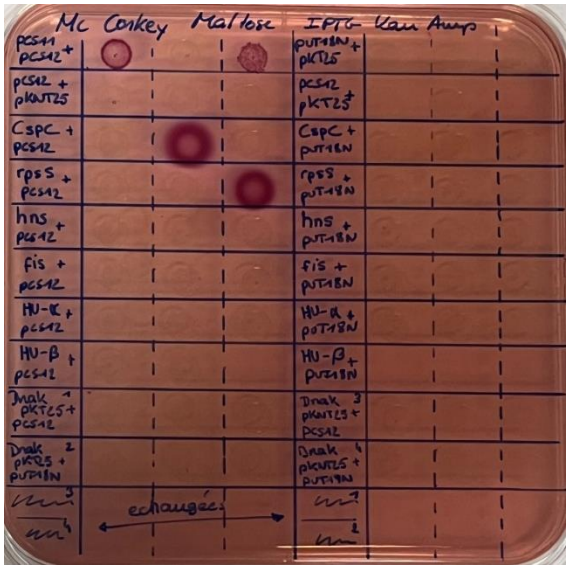
Supplementary Figure III.3.3-1 - Ampicillin does not affect transposition. (A) Western blot analysis was performed to examine biotinylation profiles after 0 and 150 minutes under two different conjugation conditions. In the first condition, conjugation was carried out for 150 minutes on plates containing both ampicillin and biotin. In the second condition, cells were initially conjugated for 90 minutes on a biotin plate and then transferred to a plate containing both biotin and ampicillin for the remaining 60 minutes. Lysates were run on SDS-PAGE and revealed using AP Streptavidin. The experiment was conducted with three different biotin ligases: TurboID, miniTurbo, and AirID. Test conditions, consisting of TnpA fused to the respective labeling enzyme, and negative controls, consisting of the labeling enzyme alone, were included. Six additional lanes were included as test conditions without conjugation (no recipient cells) to follow the decrease of the donor cell signal. **(B)** Transposition frequencies were calculated and plotted for every construction (positive and negative controls) in the two conditions. Transposition frequency was obtained by dividing the number of colonies on overnight incubated kanamycin ampicillin LB plates by colonies on overnight incubated ampicillin plates.

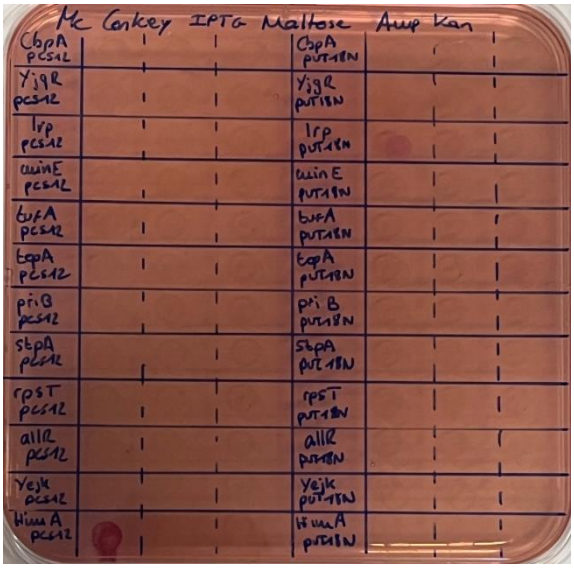


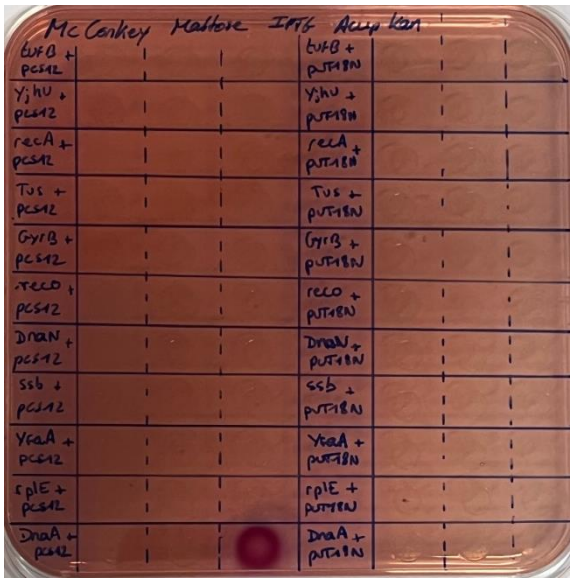
Supplementary Figure III.4.1-1 - Cell ratio and ampicillin reduce overnight signal. SDS gel was done to control for total protein in each sample. Lanes 1-12 corresponds to cell lysates after different time point of conjugation and for the three biotin ligases. Donor/recipient cell ratio used for the conjugation was lowered from 1:2 to 1:10.



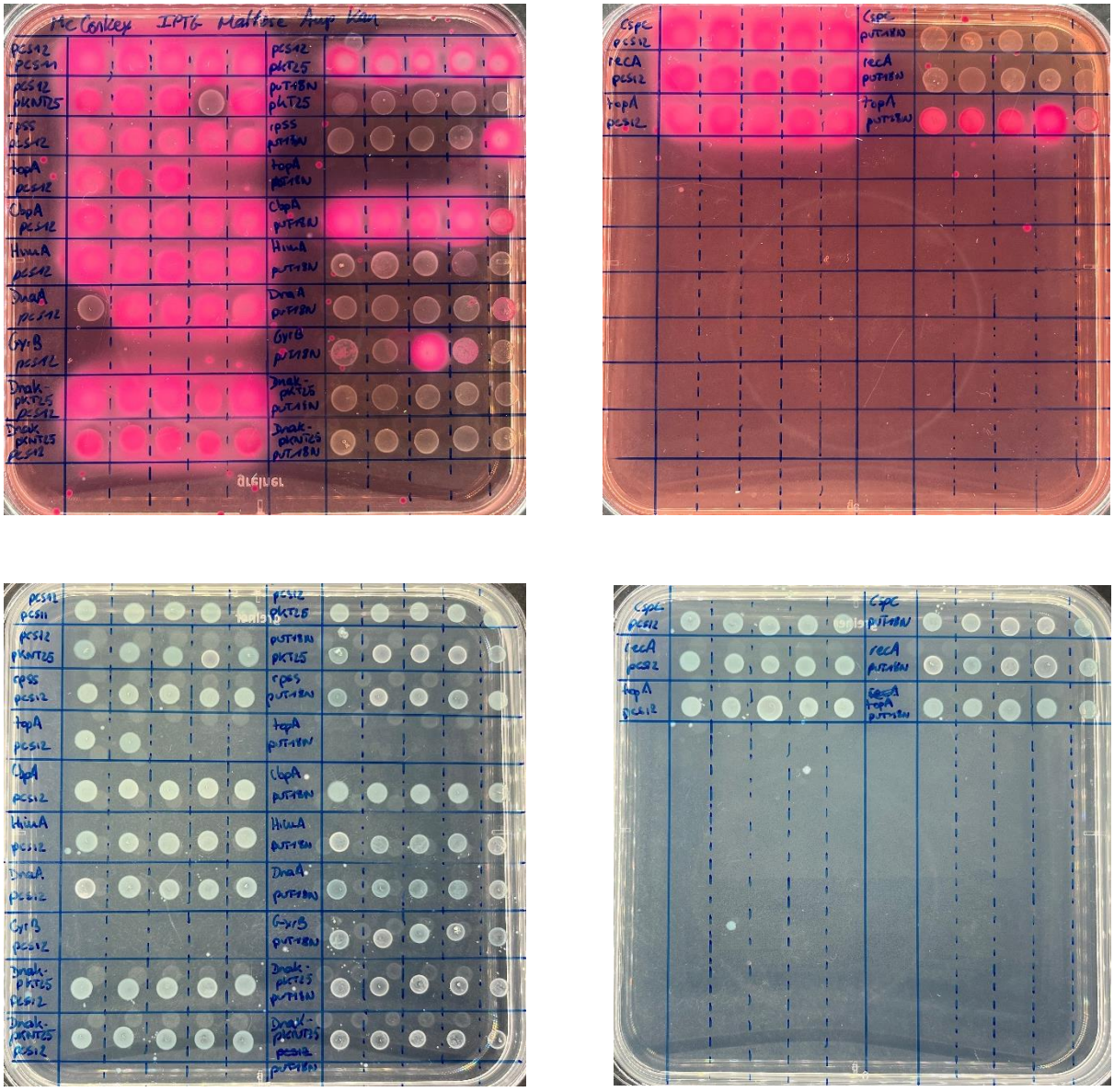
Supplementary Figure III.4.2-1 - Biotinylated proteins enrichment. (A) Coomassie blue gels of miniTurbo samples at different steps of the enrichment process. Samples analysed include "Pre-binding represents the cell lysate after dialysis and prior to incubation with streptavidin beads. The "post-binding" samples correspond to the supernatant after incubation with the beads. "Wash 1 and 3" samples correspond to the first and last washes of streptavidin beads before the elution step. Elution was performed by adding 5% SDS to the beads and heating samples at 95°C for 10 minutes. Control (+) lanes correspond to TnpA-BirA* and (-) corresponds to BirA* alone. **(B)** Coomassie blue gels of AirID samples at different steps of the enrichment process.







Supplementary Figure III.6.2-1 – Bacterial-two-hybrid results (2). Profiles of co-transformants after being spotted on McConkey plates containing 1% maltose and 0.5mM IPTG for 24 hours. (Protein + pCS12) is the protein fused to T25 fragment with TnpA fused to T18. (Protein + pUT18N) is the protein fused to T25 with an empty T18 fragment. (pCS11 + pCS12) is T25 and T28 fused to interacting proteins. (pUT18 + pKT25), (pCS12 + pKT25), and (pCS12 + pKNT25) are negative controls.



Supplementary Figure III.6.2-2 – Bacterial-two-hybrid results (3). Profiles of co-transformants after being spotted on McConkey plates containing 1% maltose, 0.5mM IPTG and X-Gal-containing plates (40µg/mL) for 24 hours. (Protein + pUT18N) is the protein fused to T25 with an empty T18 fragment. (pCS11 + pCS12) is T25 and T28 fused to interacting proteins. (pUT18 + pKT25), (pCS12 + pKT25), and (pCS12 + pKNT25) are negative controls.

Supplementary Data

Mass spectrometry data

Protein IDs	Gene names	iBAQ TnpA-	iBAQ TnpA+	riBAQ TnpA-	rIBAQ TnpA+	riBAQ TnpA+/ riBAQ TnpA-
P0A8C1	ybjQ	2.72E+06	0.00E+00	3.12E-02	0.00E+00	Not present in TnpA+ sample
P0A862	tpx	9.51E+05	0.00E+00	1.09E-02	0.00E+00	Not present in TnpA+ sample
P69831	gatC	9.37E+05	0.00E+00	1.07E-02	0.00E+00	Not present in TnpA+ sample
Q46802	ygeV	8.94E+05	0.00E+00	1.03E-02	0.00E+00	Not present in TnpA+ sample
P0A9S3	gatD	7.25E+05	0.00E+00	8.32E-03	0.00E+00	Not present in TnpA+ sample
P0A9F3	cysB	4.10E+05	0.00E+00	4.70E-03	0.00E+00	Not present in TnpA+ sample
P37685	aldB	2.68E+05	0.00E+00	3.07E-03	0.00E+00	Not present in TnpA+ sample
P0AGJ2	trmH	2.37E+05	0.00E+00	2.71E-03	0.00E+00	Not present in TnpA+ sample
P0A7Y8	rnpA	2.36E+05	0.00E+00	2.71E-03	0.00E+00	Not present in TnpA+ sample
P39199	prmB	2.08E+05	0.00E+00	2.39E-03	0.00E+00	Not present in TnpA+ sample
P0ACP1	cra	2.07E+05	0.00E+00	2.38E-03	0.00E+00	Not present in TnpA+ sample
P0AEI4	rimO	1.91E+05	0.00E+00	2.19E-03	0.00E+00	Not present in TnpA+ sample
P33913	yejA	1.81E+05	0.00E+00	2.07E-03	0.00E+00	Not present in TnpA+ sample
P22333	add	1.67E+05	0.00E+00	1.91E-03	0.00E+00	Not present in TnpA+ sample
P09372	grpE	1.49E+05	0.00E+00	1.71E-03	0.00E+00	Not present in TnpA+ sample
P75783	ybiO	1.16E+05	0.00E+00	1.33E-03	0.00E+00	Not present in TnpA+ sample
P76658	hldE	1.09E+05	0.00E+00	1.26E-03	0.00E+00	Not present in TnpA+ sample
P0ABJ9	cydA	1.01E+05	0.00E+00	1.16E-03	0.00E+00	Not present in TnpA+ sample
P03004	dnaA	9.50E+04	0.00E+00	1.09E-03	0.00E+00	Not present in TnpA+ sample
P06709	birA	9.42E+04	0.00E+00	1.08E-03	0.00E+00	Not present in TnpA+ sample
P0A8J8	rhlB	9.10E+04	0.00E+00	1.04E-03	0.00E+00	Not present in TnpA+ sample
P38683	torT	0.00E+00	1.21E+06	0.00E+00	2.08E-02	Not present in TnpA- sample
P77398	arnA	0.00E+00	1.03E+06	0.00E+00	1.77E-02	Not present in TnpA- sample
P0AGE9	sucD	0.00E+00	6.59E+05	0.00E+00	1.13E-02	Not present in TnpA- sample
P0A972	cspE	0.00E+00	6.34E+05	0.00E+00	1.09E-02	Not present in TnpA- sample
P0AB71	fbaA	0.00E+00	4.48E+05	0.00E+00	7.68E-03	Not present in TnpA- sample
P0C0L2	osmC	0.00E+00	4.12E+05	0.00E+00	7.07E-03	Not present in TnpA- sample
P07014	sdhB	0.00E+00	3.92E+05	0.00E+00	6.71E-03	Not present in TnpA- sample

P0A7H3	recO	0.00E+00	3.90E+05	0.00E+00	6.68E-03	Not present in TnpA- sample
P00448	sodA	0.00E+00	2.88E+05	0.00E+00	4.94E-03	Not present in TnpA- sample
P0AEE5	mglB	0.00E+00	2.39E+05	0.00E+00	4.10E-03	Not present in TnpA- sample
P0AC69	grxD	0.00E+00	2.33E+05	0.00E+00	3.99E-03	Not present in TnpA- sample
P0ADH5	fimB	0.00E+00	2.22E+05	0.00E+00	3.81E-03	Not present in TnpA- sample
P0AFR4	yciO	0.00E+00	2.10E+05	0.00E+00	3.60E-03	Not present in TnpA- sample
P0AGL2	tdcF	0.00E+00	1.89E+05	0.00E+00	3.24E-03	Not present in TnpA- sample
P0A6E4	argG	0.00E+00	1.50E+05	0.00E+00	2.58E-03	Not present in TnpA- sample
P31658	hchA	0.00E+00	1.29E+05	0.00E+00	2.20E-03	Not present in TnpA- sample
P0A7B5	proB	0.00E+00	1.22E+05	0.00E+00	2.10E-03	Not present in TnpA- sample
P06610	btuE	0.00E+00	1.18E+05	0.00E+00	2.02E-03	Not present in TnpA- sample
P39831	ydfG	0.00E+00	1.11E+05	0.00E+00	1.90E-03	Not present in TnpA- sample
P76290	cmoA	0.00E+00	1.04E+05	0.00E+00	1.78E-03	Not present in TnpA- sample
P02925	rbsB	0.00E+00	9.66E+04	0.00E+00	1.65E-03	Not present in TnpA- sample
P25516	acnA	0.00E+00	5.56E+04	0.00E+00	9.53E-04	Not present in TnpA- sample
P29012	dadX	0.00E+00	5.17E+04	0.00E+00	8.86E-04	Not present in TnpA- sample
P22259	pckA	0.00E+00	4.98E+04	0.00E+00	8.53E-04	Not present in TnpA- sample
P07813	leuS	0.00E+00	2.52E+04	0.00E+00	4.33E-04	Not present in TnpA- sample
P0AAE2	proY	0.00E+00	5.28E+03	0.00E+00	9.04E-05	Not present in TnpA- sample
P0A7U3	rpsS	3.01E+06	3.23E+07	3.45E-02	5.54E-01	1.61E+01
P08200	icd	3.00E+04	2.93E+05	3.44E-04	5.02E-03	1.46E+01
P0ABD8	accB	1.88E+07	1.32E+08	2.16E-01	2.27E+00	1.05E+01
P0A799	pgk	1.63E+05	1.02E+06	1.87E-03	1.75E-02	9.33E+00
P0ABK5	cysK	3.93E+04	2.26E+05	4.51E-04	3.88E-03	8.60E+00
P08997	aceB	4.88E+04	2.62E+05	5.60E-04	4.48E-03	8.00E+00
P0ACP5	gntR	8.47E+06	4.26E+07	9.71E-02	7.30E-01	7.52E+00
P39342	yjgR	2.89E+05	1.29E+06	3.32E-03	2.20E-02	6.65E+00
P23538	ppsA	4.97E+04	2.10E+05	5.70E-04	3.61E-03	6.33E+00
P0A7U7	rpsT	1.06E+08	4.12E+08	1.21E+00	7.06E+00	5.82E+00
P0AES9	hdeA	2.38E+06	8.47E+06	2.73E-02	1.45E-01	5.31E+00
P0A7D4	purA	1.58E+05	5.51E+05	1.81E-03	9.44E-03	5.22E+00
P0A9L8	proC	2.28E+05	7.83E+05	2.61E-03	1.34E-02	5.13E+00

P0A9B2	gapA	2.63E+05	8.24E+05	3.02E-03	1.41E-02	4.68E+00
P68919	rplY	5.32E+05	1.59E+06	6.10E-03	2.72E-02	4.46E+00
P0AET2	hdeB	5.54E+05	1.47E+06	6.36E-03	2.52E-02	3.96E+00
P10121	ftsY	2.97E+04	7.22E+04	3.41E-04	1.24E-03	3.63E+00
P0A9G6	aceA	8.29E+05	1.99E+06	9.51E-03	3.41E-02	3.58E+00
P0ADU5	ygiW	2.40E+05	5.61E+05	2.75E-03	9.62E-03	3.50E+00
P0ACY3	yeaG	1.06E+05	2.45E+05	1.22E-03	4.20E-03	3.45E+00
Q46851	gpr	1.82E+05	4.11E+05	2.08E-03	7.05E-03	3.38E+00
P0A6F5	groL	7.02E+05	1.49E+06	8.05E-03	2.55E-02	3.17E+00
P37744;P61887	rmlA1	1.51E+05	3.19E+05	1.73E-03	5.47E-03	3.16E+00
P08178	purM	1.26E+05	2.54E+05	1.45E-03	4.35E-03	3.00E+00
P0ABH0	ftsA	6.41E+04	1.28E+05	7.36E-04	2.19E-03	2.98E+00
P0A910	ompA	3.34E+05	6.41E+05	3.83E-03	1.10E-02	2.86E+00
P0ADB1	osmE	2.16E+05	4.10E+05	2.48E-03	7.03E-03	2.84E+00
P0A6P9	eno	9.40E+05	1.77E+06	1.08E-02	3.04E-02	2.82E+00
P0ADZ4	rpsO	1.24E+07	2.25E+07	1.42E-01	3.85E-01	2.71E+00
P23845	cysN	4.97E+04	8.72E+04	5.70E-04	1.49E-03	2.62E+00
P0A9Y6	cspC	1.61E+06	2.80E+06	1.85E-02	4.80E-02	2.60E+00
P37188	gatB	1.15E+07	1.95E+07	1.32E-01	3.35E-01	2.54E+00
P23865	prc	4.15E+04	7.01E+04	4.76E-04	1.20E-03	2.53E+00
P25553	aldA	8.61E+04	1.37E+05	9.88E-04	2.34E-03	2.37E+00
P17994	yfaA	5.52E+05	8.71E+05	6.33E-03	1.49E-02	2.36E+00
P0A7X3	rpsI	2.95E+07	4.58E+07	3.39E-01	7.85E-01	2.32E+00
P37128	nudK	1.58E+06	2.39E+06	1.82E-02	4.09E-02	2.25E+00
P0C8J6	gatY	3.75E+06	5.63E+06	4.30E-02	9.64E-02	2.24E+00
P0AEK4	fabI	1.46E+05	2.03E+05	1.68E-03	3.48E-03	2.08E+00
P60723	rplD	2.29E+07	3.15E+07	2.62E-01	5.40E-01	2.06E+00
P0CE48;P0CE47	tufB;tufA	3.27E+07	4.36E+07	3.75E-01	7.47E-01	1.99E+00
P0A836	sucC	1.27E+05	1.67E+05	1.45E-03	2.86E-03	1.97E+00
P0A7G2	rbfA	7.46E+05	9.80E+05	8.56E-03	1.68E-02	1.96E+00
P0DTT0	bipA	1.93E+05	2.49E+05	2.21E-03	4.26E-03	1.93E+00
P0A734	minE	6.56E+05	7.87E+05	7.53E-03	1.35E-02	1.79E+00

P0ABH7	gltA	2.72E+05	3.21E+05	3.12E-03	5.50E-03	1.76E+00
P75931	yceM	1.96E+05	2.28E+05	2.25E-03	3.91E-03	1.74E+00
P68679	rpsU	7.45E+08	8.59E+08	8.55E+00	1.47E+01	1.72E+00
P0A6Y8	dnaK	1.59E+06	1.81E+06	1.83E-02	3.11E-02	1.70E+00
P0AED0	uspA	1.81E+06	2.06E+06	2.08E-02	3.52E-02	1.70E+00
P0A7R9	rpsK	1.37E+08	1.53E+08	1.57E+00	2.61E+00	1.67E+00
P16456	selD	3.77E+05	4.19E+05	4.33E-03	7.19E-03	1.66E+00
P0A6H1	clpX	1.83E+05	2.03E+05	2.10E-03	3.47E-03	1.65E+00
P23893	hemL	1.39E+06	1.52E+06	1.59E-02	2.61E-02	1.64E+00
P0A7A9	ppa	1.42E+05	1.55E+05	1.62E-03	2.65E-03	1.63E+00
P0A9A9	fur	4.86E+05	5.29E+05	5.57E-03	9.06E-03	1.63E+00
P61889	mdh	2.30E+06	2.49E+06	2.64E-02	4.27E-02	1.61E+00
P0A9L3	fklB	1.56E+05	1.68E+05	1.79E-03	2.88E-03	1.61E+00
P0A817	metK	9.80E+05	1.05E+06	1.12E-02	1.80E-02	1.60E+00
P33920	yejK	9.55E+04	1.01E+05	1.10E-03	1.73E-03	1.58E+00
P0A8M6	yeeX	4.34E+07	4.57E+07	4.98E-01	7.83E-01	1.57E+00
P00452	nrdA	9.06E+04	9.49E+04	1.04E-03	1.63E-03	1.56E+00
P0A6P1	tsf	3.09E+06	3.17E+06	3.54E-02	5.43E-02	1.53E+00
P21170	speA	1.57E+05	1.59E+05	1.80E-03	2.72E-03	1.51E+00
P00961	glyS	7.66E+05	7.70E+05	8.79E-03	1.32E-02	1.50E+00
P0ABU2	ychF	1.44E+05	1.37E+05	1.65E-03	2.36E-03	1.43E+00
P24182	accC	9.26E+05	8.81E+05	1.06E-02	1.51E-02	1.42E+00
P60716	lipA	1.60E+05	1.52E+05	1.84E-03	2.61E-03	1.42E+00
P0A8F0	upp	2.40E+06	2.26E+06	2.75E-02	3.86E-02	1.41E+00
P77611	rsxC	6.41E+05	5.88E+05	7.35E-03	1.01E-02	1.37E+00
P0A6M8	fusA	1.96E+06	1.79E+06	2.25E-02	3.07E-02	1.36E+00
P00954	trpS	3.40E+05	3.09E+05	3.89E-03	5.29E-03	1.36E+00
P0C0V0	degP	1.59E+05	1.44E+05	1.83E-03	2.47E-03	1.35E+00
P69797	manX	7.00E+04	6.28E+04	8.03E-04	1.08E-03	1.34E+00
P21599	pykA	4.45E+04	3.98E+04	5.10E-04	6.82E-04	1.34E+00
P69910;P69908	gadB/B	2.61E+06	2.32E+06	2.99E-02	3.98E-02	1.33E+00
P62620	ispG	2.18E+05	1.92E+05	2.50E-03	3.30E-03	1.32E+00

P69924	nrdB	9.10E+04	7.96E+04	1.04E-03	1.36E-03	1.31E+00
P0A763	ndk	6.60E+05	5.70E+05	7.56E-03	9.76E-03	1.29E+00
P62768	yaeH	3.08E+08	2.57E+08	3.53E+00	4.41E+00	1.25E+00
P0AGE0	ssb	9.52E+05	7.93E+05	1.09E-02	1.36E-02	1.25E+00
P0A9K9	slyD	2.77E+05	2.29E+05	3.18E-03	3.92E-03	1.23E+00
P0A7K2	rplL	7.99E+06	6.59E+06	9.16E-02	1.13E-01	1.23E+00
P37903	uspF	1.41E+07	1.15E+07	1.62E-01	1.98E-01	1.23E+00
P39177	uspG	3.28E+06	2.68E+06	3.76E-02	4.58E-02	1.22E+00
P17846	cysI	3.45E+04	2.80E+04	3.95E-04	4.80E-04	1.21E+00
P76539	ypeA	2.41E+06	1.93E+06	2.77E-02	3.31E-02	1.20E+00
P0A7M9	rpmE	1.09E+06	8.65E+05	1.25E-02	1.48E-02	1.18E+00
P0A6H5	hslU	2.94E+06	2.32E+06	3.37E-02	3.97E-02	1.18E+00
P0ABB4	atpD	1.98E+06	1.55E+06	2.27E-02	2.66E-02	1.17E+00
P14081	selB	2.33E+04	1.82E+04	2.68E-04	3.12E-04	1.17E+00
P0AFG8	aceE	5.77E+05	4.50E+05	6.61E-03	7.70E-03	1.16E+00
P00561	thrA	2.81E+05	2.18E+05	3.22E-03	3.74E-03	1.16E+00
P13029	katG	4.23E+04	3.21E+04	4.85E-04	5.50E-04	1.13E+00
P0A7S3	rpsL	9.81E+08	7.38E+08	1.13E+01	1.26E+01	1.12E+00
P62399	rplE	2.02E+07	1.49E+07	2.31E-01	2.55E-01	1.10E+00
P0C0R7	rlmE	2.46E+06	1.82E+06	2.83E-02	3.11E-02	1.10E+00
P0ACJ0	lrp	2.29E+06	1.68E+06	2.63E-02	2.88E-02	1.10E+00
P0A850	tig	5.00E+06	3.65E+06	5.74E-02	6.25E-02	1.09E+00
P0AEP3	galU	9.13E+05	6.62E+05	1.05E-02	1.13E-02	1.08E+00
P0A6X7	ihfA	2.57E+07	1.84E+07	2.94E-01	3.16E-01	1.07E+00
P0C018	rplR	2.69E+06	1.91E+06	3.08E-02	3.27E-02	1.06E+00
P0A8S1	argP	2.17E+05	1.54E+05	2.49E-03	2.63E-03	1.06E+00
P0A7L0	rplA	7.46E+06	5.25E+06	8.56E-02	8.99E-02	1.05E+00
P09373;P68066	pflB	1.31E+06	9.03E+05	1.50E-02	1.55E-02	1.03E+00
P0A7M6	rpmC	3.59E+06	2.45E+06	4.12E-02	4.19E-02	1.02E+00
P0A9K3	ybeZ	4.91E+05	3.34E+05	5.63E-03	5.72E-03	1.02E+00
P00968	carB	3.64E+05	2.47E+05	4.17E-03	4.24E-03	1.02E+00
P0A6Z3	htpG	6.19E+05	4.18E+05	7.10E-03	7.15E-03	1.01E+00

P37440	ucpA	4.00E+05	2.69E+05	4.58E-03	4.60E-03	1.00E+00
P0A7G6	recA	1.91E+06	1.28E+06	2.20E-02	2.20E-02	1.00E+00
P36683	acnB	6.71E+05	4.48E+05	7.69E-03	7.67E-03	9.97E-01
P06966	dicA	1.10E+06	7.27E+05	1.27E-02	1.25E-02	9.83E-01
P0AEK2	fabG	3.55E+06	2.31E+06	4.07E-02	3.96E-02	9.75E-01
P0C8J8	gatZ	7.05E+06	4.59E+06	8.08E-02	7.87E-02	9.74E-01
P0A6X3	hfq	5.68E+06	3.68E+06	6.52E-02	6.31E-02	9.69E-01
P0ABH9	clpA	1.62E+05	1.02E+05	1.86E-03	1.74E-03	9.38E-01
P0A7V8	rpsD	1.08E+07	6.78E+06	1.24E-01	1.16E-01	9.36E-01
P0A7L3	rplT	4.71E+06	2.92E+06	5.41E-02	5.00E-02	9.25E-01
P23830	pssA	6.57E+06	4.06E+06	7.53E-02	6.96E-02	9.23E-01
P0A9A6	ftsZ	1.08E+06	6.60E+05	1.24E-02	1.13E-02	9.10E-01
P0A7R5	rpsJ	9.81E+06	5.96E+06	1.12E-01	1.02E-01	9.09E-01
P0A705	infB	9.31E+05	5.64E+05	1.07E-02	9.66E-03	9.05E-01
P0A7T7	rpsR	2.96E+07	1.78E+07	3.39E-01	3.06E-01	9.01E-01
P08839	ptsI	2.14E+05	1.28E+05	2.45E-03	2.20E-03	8.95E-01
P0A7P5	rpmH	2.69E+09	1.60E+09	3.09E+01	2.74E+01	8.87E-01
P27306	sthA	8.68E+05	5.12E+05	9.96E-03	8.77E-03	8.81E-01
P0ACF0	hupA	2.35E+07	1.38E+07	2.69E-01	2.36E-01	8.76E-01
P0A7Q6	rpmJ	2.70E+07	1.57E+07	3.10E-01	2.69E-01	8.68E-01
P17169	glmS	6.51E+04	3.76E+04	7.46E-04	6.45E-04	8.64E-01
P0AG63	rpsQ	4.97E+06	2.87E+06	5.70E-02	4.92E-02	8.63E-01
P0AFF6	nusA	1.34E+05	7.43E+04	1.53E-03	1.27E-03	8.30E-01
P68767	pepA	6.19E+06	3.40E+06	7.10E-02	5.82E-02	8.20E-01
P24251	crl	4.85E+05	2.64E+05	5.56E-03	4.52E-03	8.12E-01
P0A7D7	purC	5.42E+05	2.89E+05	6.22E-03	4.96E-03	7.97E-01
P0AFY8	seqA	1.50E+06	7.96E+05	1.72E-02	1.36E-02	7.93E-01
P06959	aceF	2.55E+05	1.35E+05	2.93E-03	2.32E-03	7.91E-01
P08622	dnaJ	1.64E+06	8.62E+05	1.88E-02	1.48E-02	7.85E-01
P02413	rplO	1.13E+07	5.89E+06	1.30E-01	1.01E-01	7.76E-01
P60785	lepA	1.62E+05	8.39E+04	1.85E-03	1.44E-03	7.76E-01
P0A7B1	ppk	8.85E+05	4.59E+05	1.01E-02	7.87E-03	7.75E-01

P0ABT2	dps	3.40E+07	1.76E+07	3.90E-01	3.02E-01	7.74E-01
P63284	clpB	1.44E+06	7.40E+05	1.65E-02	1.27E-02	7.69E-01
P0AGJ5	yfiF	3.30E+05	1.69E+05	3.79E-03	2.89E-03	7.64E-01
P06996;P21420	ompC	1.02E+06	5.21E+05	1.17E-02	8.92E-03	7.59E-01
P0AA16	ompR	2.15E+05	1.09E+05	2.46E-03	1.87E-03	7.58E-01
P33599	nuoC	2.08E+05	1.04E+05	2.38E-03	1.78E-03	7.49E-01
P0A8V2	rpoB	1.69E+06	8.42E+05	1.94E-02	1.44E-02	7.45E-01
P60624	rplX	2.04E+07	1.01E+07	2.34E-01	1.74E-01	7.42E-01
P16095;P42630;P30744	sdaA	1.91E+06	9.48E+05	2.19E-02	1.62E-02	7.42E-01
P0A9J0	rng	5.24E+04	2.56E+04	6.01E-04	4.39E-04	7.30E-01
P0AEH5	elaB	1.35E+06	6.60E+05	1.55E-02	1.13E-02	7.29E-01
P0ADY3	rplN	3.99E+08	1.92E+08	4.57E+00	3.29E+00	7.21E-01
P0ACF8	hns	8.88E+06	4.27E+06	1.02E-01	7.32E-02	7.19E-01
P0ADG4	suhB	3.29E+05	1.57E+05	3.78E-03	2.69E-03	7.13E-01
P0A7E5	pyrG	8.28E+05	3.95E+05	9.50E-03	6.77E-03	7.13E-01
P36767	rdgC	7.68E+05	3.65E+05	8.81E-03	6.26E-03	7.11E-01
P00582	polA	2.71E+05	1.28E+05	3.10E-03	2.19E-03	7.07E-01
P0A7T3	rpsP	2.28E+07	1.07E+07	2.62E-01	1.84E-01	7.04E-01
P0ACJ8	crp	1.86E+06	8.72E+05	2.13E-02	1.49E-02	7.00E-01
P0AB89	purB	2.90E+05	1.35E+05	3.33E-03	2.32E-03	6.96E-01
P00960	glyQ	8.03E+05	3.73E+05	9.21E-03	6.40E-03	6.94E-01
P00490	malP	1.32E+06	6.07E+05	1.51E-02	1.04E-02	6.89E-01
P0A7M2	rpmB	5.82E+08	2.68E+08	6.67E+00	4.59E+00	6.88E-01
P60422	rplB	2.32E+07	1.07E+07	2.66E-01	1.83E-01	6.87E-01
P64581	yqjD	6.13E+06	2.80E+06	7.03E-02	4.80E-02	6.82E-01
P0A9P0	lpdA	1.52E+06	6.94E+05	1.74E-02	1.19E-02	6.82E-01
P76402	yegP	6.28E+06	2.86E+06	7.20E-02	4.90E-02	6.80E-01
P0A6Q3	fabA	1.85E+06	8.40E+05	2.12E-02	1.44E-02	6.79E-01
P0AFG6	sucB	1.49E+06	6.72E+05	1.71E-02	1.15E-02	6.74E-01
P0A7V3	rpsC	3.33E+08	1.46E+08	3.82E+00	2.50E+00	6.55E-01
P0A7Q1	rpmI	2.55E+07	1.11E+07	2.92E-01	1.91E-01	6.53E-01
P08142	ilvB	2.45E+05	1.07E+05	2.81E-03	1.84E-03	6.52E-01

P0AC59	grxB	7.34E+05	3.20E+05	8.42E-03	5.48E-03	6.51E-01
P0A6R3	fis	3.25E+06	1.40E+06	3.72E-02	2.40E-02	6.45E-01
P0A7V0	rpsB	1.71E+07	7.33E+06	1.96E-01	1.26E-01	6.41E-01
P0AEZ3	minD	6.03E+06	2.55E+06	6.91E-02	4.37E-02	6.32E-01
P0AG30	rho	1.81E+06	7.61E+05	2.08E-02	1.30E-02	6.27E-01
P0A993	fbp	2.46E+05	1.03E+05	2.82E-03	1.77E-03	6.26E-01
P76550	yffS	1.83E+05	7.62E+04	2.10E-03	1.31E-03	6.22E-01
P69783	crr	5.01E+05	2.08E+05	5.74E-03	3.56E-03	6.21E-01
P77735	yajO	1.66E+05	6.90E+04	1.91E-03	1.18E-03	6.21E-01
P0AE08	ahpC	3.89E+06	1.60E+06	4.47E-02	2.75E-02	6.15E-01
P0ABB0	atpA	5.77E+05	2.38E+05	6.62E-03	4.07E-03	6.15E-01
P0AC41	sdhA	1.54E+05	6.20E+04	1.77E-03	1.06E-03	6.01E-01
P31979	nuoF	5.70E+05	2.29E+05	6.54E-03	3.92E-03	5.99E-01
P75864	rlmL	4.42E+04	1.75E+04	5.07E-04	3.01E-04	5.93E-01
P0AG55	rplF	1.50E+07	5.92E+06	1.72E-01	1.02E-01	5.91E-01
P0AES6	gyrB	4.57E+05	1.80E+05	5.24E-03	3.08E-03	5.88E-01
P0AG67	rpsA	3.97E+06	1.56E+06	4.55E-02	2.68E-02	5.87E-01
P0AEU7	skp	4.45E+06	1.73E+06	5.11E-02	2.97E-02	5.81E-01
P23843	oppA	2.04E+06	7.80E+05	2.34E-02	1.34E-02	5.71E-01
P0A853	tnaA	2.06E+07	7.80E+06	2.37E-01	1.34E-01	5.65E-01
P60438	rplC	1.76E+07	6.66E+06	2.02E-01	1.14E-01	5.65E-01
P0A6J5	dadA	1.95E+05	7.38E+04	2.24E-03	1.26E-03	5.64E-01
P0A6B7	iscS	1.97E+05	7.41E+04	2.26E-03	1.27E-03	5.62E-01
P0AG59	rpsN	3.42E+06	1.26E+06	3.93E-02	2.15E-02	5.48E-01
P02358	rpsF	1.01E+07	3.68E+06	1.16E-01	6.30E-02	5.41E-01
P0A7R1	rplI	3.65E+06	1.30E+06	4.19E-02	2.23E-02	5.32E-01
P0A717	prs	4.03E+06	1.41E+06	4.62E-02	2.42E-02	5.23E-01
P06612	topA	1.86E+05	6.47E+04	2.13E-03	1.11E-03	5.20E-01
P0A6P5	der	4.42E+05	1.53E+05	5.06E-03	2.63E-03	5.19E-01
P0AAB6	galF	6.92E+05	2.38E+05	7.93E-03	4.08E-03	5.14E-01
P0A8N5;P0A8N3	lysU;lysS	8.30E+04	2.72E+04	9.52E-04	4.67E-04	4.90E-01
P0A7S9	rpsM	4.75E+07	1.56E+07	5.45E-01	2.67E-01	4.89E-01

P0AES4	gyrA	8.65E+05	2.83E+05	9.92E-03	4.85E-03	4.89E-01
P0ADZ0	rplW	1.05E+08	3.42E+07	1.20E+00	5.86E-01	4.88E-01
P02359	rpsG	4.23E+07	1.36E+07	4.85E-01	2.33E-01	4.80E-01
P0A9X4	mreB	3.63E+06	1.16E+06	4.16E-02	1.99E-02	4.79E-01
P37095	pepB	1.15E+05	3.64E+04	1.31E-03	6.24E-04	4.75E-01
P0AFG3	sucA	3.32E+05	1.04E+05	3.81E-03	1.78E-03	4.68E-01
P61175	rplV	5.33E+08	1.66E+08	6.11E+00	2.85E+00	4.66E-01
P39356	yjhU	3.03E+06	9.31E+05	3.48E-02	1.59E-02	4.58E-01
P0ADG7	guaB	6.42E+05	1.94E+05	7.37E-03	3.32E-03	4.50E-01
P0A9Q5	accD	4.36E+05	1.30E+05	5.00E-03	2.22E-03	4.44E-01
P0A7W7	rpsH	2.73E+07	7.96E+06	3.13E-01	1.36E-01	4.36E-01
P09546	putA	4.10E+05	1.18E+05	4.70E-03	2.03E-03	4.32E-01
P0A9E5	fnr	7.60E+05	2.19E+05	8.71E-03	3.76E-03	4.31E-01
P0A948	rimJ	8.77E+04	2.51E+04	1.01E-03	4.30E-04	4.28E-01
P0ACY1	ydjA	6.58E+06	1.88E+06	7.55E-02	3.22E-02	4.26E-01
P0AB91	aroG	3.60E+05	1.03E+05	4.13E-03	1.76E-03	4.25E-01
P39902	sprT	5.46E+05	1.52E+05	6.26E-03	2.60E-03	4.16E-01
P0A6Q6	fabZ	3.91E+07	1.07E+07	4.48E-01	1.84E-01	4.11E-01
P0AGD7	ffh	1.25E+06	3.40E+05	1.43E-02	5.82E-03	4.08E-01
P0A6F9	groS	1.89E+07	5.16E+06	2.17E-01	8.84E-02	4.07E-01
P0A832	smpB	2.87E+06	7.63E+05	3.29E-02	1.31E-02	3.97E-01
P0A707	infC	6.68E+06	1.75E+06	7.66E-02	3.00E-02	3.92E-01
P0AA10	rplM	3.48E+07	9.10E+06	3.99E-01	1.56E-01	3.91E-01
P0AG51	rpmD	9.96E+07	2.60E+07	1.14E+00	4.46E-01	3.90E-01
P0A9M0	lon	1.79E+06	4.59E+05	2.05E-02	7.86E-03	3.83E-01
P0A7W1	rpsE	7.87E+07	2.01E+07	9.03E-01	3.44E-01	3.81E-01
P30177	ybiB	2.38E+06	5.88E+05	2.73E-02	1.01E-02	3.69E-01
P0A9Q1	arcA	2.78E+06	6.85E+05	3.19E-02	1.17E-02	3.67E-01
P0A698	uvrA	5.90E+05	1.43E+05	6.77E-03	2.44E-03	3.61E-01
P0A8M0	asnS	1.68E+05	3.96E+04	1.93E-03	6.79E-04	3.52E-01
P0AFW2	rmf	2.95E+06	6.86E+05	3.39E-02	1.18E-02	3.47E-01
P36659	cbpA	2.74E+06	6.34E+05	3.14E-02	1.09E-02	3.46E-01

P33224	aidB	6.46E+06	1.47E+06	7.41E-02	2.52E-02	3.40E-01
P0A7L8	rpmA	2.76E+08	6.26E+07	3.16E+00	1.07E+00	3.39E-01
P0AGK4	yhbY	7.68E+06	1.74E+06	8.81E-02	2.98E-02	3.38E-01
P12008	aroC	1.85E+05	4.09E+04	2.12E-03	7.01E-04	3.31E-01
P06987	hisB	3.02E+06	6.68E+05	3.46E-02	1.14E-02	3.30E-01
P0A9W3	ettA	7.10E+05	1.54E+05	8.15E-03	2.65E-03	3.25E-01
P36548	amiA	1.16E+06	2.39E+05	1.34E-02	4.10E-03	3.07E-01
P0A9P6	deaD	6.72E+04	1.35E+04	7.71E-04	2.32E-04	3.00E-01
P0A7I4	prfC	4.32E+05	8.66E+04	4.95E-03	1.48E-03	3.00E-01
P39406	rsmC	1.11E+06	2.09E+05	1.27E-02	3.59E-03	2.83E-01
P0AG44	rplQ	5.10E+07	9.56E+06	5.85E-01	1.64E-01	2.80E-01
P0AE18	map	4.26E+05	7.99E+04	4.89E-03	1.37E-03	2.80E-01
P0ACG1	stpA	5.22E+06	9.46E+05	5.98E-02	1.62E-02	2.71E-01
P21507	srmB	1.45E+06	2.62E+05	1.66E-02	4.49E-03	2.70E-01
P0AFI2	parC	5.30E+05	9.32E+04	6.08E-03	1.60E-03	2.62E-01
P0A8T7	rpoC	3.16E+06	5.53E+05	3.62E-02	9.48E-03	2.62E-01
P0AFL6	ppx	8.39E+06	1.36E+06	9.62E-02	2.34E-02	2.43E-01
P37671	yiaJ	3.00E+06	4.61E+05	3.44E-02	7.90E-03	2.30E-01
P0A9Q7	adhE	3.84E+07	5.82E+06	4.40E-01	9.97E-02	2.27E-01
P0A6P7	engB	8.06E+05	1.15E+05	9.24E-03	1.96E-03	2.13E-01
P0A7J3	rplJ	1.32E+07	1.76E+06	1.51E-01	3.02E-02	2.00E-01
P69828	gatA	2.24E+06	2.92E+05	2.57E-02	5.01E-03	1.95E-01
P0A7J7	rplK	1.61E+07	2.09E+06	1.84E-01	3.59E-02	1.95E-01
P00579	rpoD	7.38E+05	9.46E+04	8.46E-03	1.62E-03	1.92E-01
P77488	dxs	2.50E+05	3.16E+04	2.86E-03	5.41E-04	1.89E-01
P00957	alaS	4.31E+05	5.21E+04	4.94E-03	8.93E-04	1.81E-01
P0A8M3	thrS	2.19E+05	2.50E+04	2.51E-03	4.28E-04	1.70E-01
P0A7Z4	rpoA	5.87E+06	6.66E+05	6.73E-02	1.14E-02	1.70E-01
P37765	rluB	1.38E+06	1.45E+05	1.58E-02	2.49E-03	1.58E-01
P08312	pheS	3.13E+05	3.20E+04	3.59E-03	5.48E-04	1.52E-01
P15038	helD	1.94E+05	1.71E+04	2.23E-03	2.93E-04	1.31E-01
P0ACF4	hupB	9.99E+06	8.66E+05	1.15E-01	1.48E-02	1.29E-01

P0A7K6	rplS	7.46E+06	5.68E+05	8.55E-02	9.73E-03	1.14E-01
P0AA39	rluC	6.75E+05	4.94E+04	7.74E-03	8.47E-04	1.09E-01
P21499	rnr	6.61E+05	4.72E+04	7.58E-03	8.08E-04	1.07E-01
P0A6Y1	ihfB	3.88E+07	2.74E+06	4.44E-01	4.69E-02	1.05E-01
P37339	lhgO	1.02E+06	6.63E+04	1.17E-02	1.14E-03	9.69E-02
P0ADY7	rplP	3.05E+07	1.89E+06	3.50E-01	3.24E-02	9.24E-02
P07395	pheT	5.01E+05	2.54E+04	5.75E-03	4.35E-04	7.57E-02
P0AG48	rplU	3.35E+07	9.74E+05	3.84E-01	1.67E-02	4.35E-02

Local pairwise alignments mentioned in section IV.2

```
#=======================================
#
# Aligned_sequences: 2
# 1: BirA
# 2: miniTurbo
# Matrix: EBLOSUM62
# Gap penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 773
# Identity:     585/773 (75.7%)
# Similarity:   585/773 (75.7%)
# Gaps:           3/773 ( 0.4%)
# Score: 3272.0
#
#
#=======================================

BirA             190 ATCCAGTTACTTAATGCTAAACAGATATTGGGTCAGCTGGATGGCGGTAG    239
                     ||||.|.|.||.||.||||||||||..||||.||||||||.||||||||
miniTurbo          1 ATCCCGCTGCTGAACGCTAAACAGATTCTGGGACAGCTGGACGGCGGGAG     50

BirA             240 TGTAGCCGTGCTGCCAGTGATTGACTCCACGAATCAGTACCTTCTTGATC    289
                     .||.||.||.|||||.|||.|.||||||||.||||||||||.||.||||
miniTurbo         51 CGTGGCAGTCCTGCCTGTGGTCGACTCCACCAATCAGTACCTGCTGGATC    100

BirA             290 GTATCGGAGAGCTTAAATCGGGCGATGCTTGCATTGCAGAATACCAGCAG    339
                     |.||||||||||.||....||.||||||||||||||||||||.||||||
miniTurbo        101 GAATCGGCGAGCTGAAGAGTGGGGATGCTTGCATTGCAGAATATCAGCAG    150

BirA             340 GCTGGCCGTGGTCGCCGGGGTCGGAAATGGTTTTCGCCTTTTGGCGCAAA    389
                     ||.||..|.||..||.|.||..||||||||||.||.||||||||.||.||
miniTurbo        151 GCAGGGAGAGGAAGCAGAGGGAGGAAATGGTTCTCTCCTTTTGGAGCTAA    200

BirA             390 CTTATATTTGTCGATGTTCTGGCGTCTGGAACAAGGCCCGGCGGCGGCGA    439
                     |.|.||..||...||||||||||.|||.|.|..||.|   |.||.||.|
miniTurbo        201 CCTGTACCTGAGTATGTTTTGGCGCCTGAAGCGGGGAC---CAGCAGCAA    247

BirA             440 TTGGTTTAAGTCTGGTTATCGGTATCGTGATGGCGGAAGTATTACGCAAG    489
                     |.||..|..|.|.|||.||||||.||.||.||||||.||||...|.||.|||
miniTurbo        248 TCGGCCTGGGCCCGGTCATCGGAATTGTCATGGCAGAAGCGCTGCGAAAG    297

BirA             490 CTGGGTGCAGATAAAGTTCGTGTTAAATGGCCTAATGACCTCTATCTGCA    539
                     |||||.||||||.||.||.||.||||||||.||||||||.||||||||
miniTurbo        298 CTGGGAGCAGACAAGGTGCGAGTCAAATGGCCCAATGACCTGTATCTGCA    347

BirA             540 GGATCGCAAGCTGGCAGGCATTCTGGTGGAGCTGACTGGCAAAACTGGCG    589
                     ||||.|.||||||||||||||.||||||||||||.|.||.|.||||.||||
miniTurbo        348 GGATAGAAAGCTGGCAGGCATCCTGGTGGAGCTGGCCGGAATAACAGGCG    397

BirA             590 ATGCGGCGCAAATAGTCATTGGAGCCGGGATCAACATGGCAATGCGCCGT    639
```

```
                     ||||.||.||.||.||||||||.||||||||.|||.||||.|||.|.||.
miniTurbo        398 ATGCTGCACAGATCGTCATTGGCGCCGGGATTAACGTGGCTATGAGGCGC    447

BirA             640 GTTGAAGAGAGTGTCGTTAATCAGGGGTGGATCACGCTGCAGGAAGCGGG    689
                     ||.||.||.||.||.||.||||||||.||||||||.||||||||||.||
miniTurbo        448 GTGGAGGAAAGCGTGGTCAATCAGGGCTGGATCACACTGCAGGAAGCAGG    497

BirA             690 GATCAATCTCGATCGTAATACGTTGGCGGCCATGCTAATACGTGAATTAC    739
                     |||.||.||.||..|.||||||..||||.||.|||||.||.||.||..|.|
miniTurbo        498 GATTAACCTGGACAGGAATACTCTGGCCGCTATGCTGATCCGAGAGCTGC    547

BirA             740 GTGCTGCGTTGGAACTCTTCGAACAAGAAGGATTGGCACCTTATCTGTCG    789
                     |.||.||..|||||||.|||||.||.||||||..||||.||.||||||||.
miniTurbo        548 GGGCAGCCCTGGAACTGTTCGAGCAGGAAGGCCTGGCTCCATATCTGTCA    597

BirA             790 CGCTGGGAAAAGCTGGATAATTTTATTAATCGCCCAGTGAAACTTATCAT    839
                     ||.|||||.||||||||||.||.||.|||.|.||.|||||.||.|||||
miniTurbo        598 CGGTGGGAGAAGCTGGATAACTTCATCAATAGACCCGTGAAGCTGATCAT    647

BirA             840 TGGTGATAAAGAAATATTTGGCATTTCACGCGGAATAGACAAACAGGGGG    889
                     |||.||.|||||.||.||.||.|||...||.||.||.||.||||||||.|
miniTurbo        648 TGGGGACAAAGAGATTTTCGGGATTAGCCGGGGGATTGATAAACAGGGAG    697

BirA             890 CTTTATTACTTGAGCAGGATGGAATAATAAAACCCTGGATGGGCGGTGAA    939
                     |..|..|.||.||.||||||.|||.|.||.||||||||||||||||.|||
miniTurbo        698 CCCTGCTGCTGGAACAGGACGGAGTTATCAAACCCTGGATGGGCGGAGAA    747

BirA             940 ATATCCCTGCGTAGTGCAGAAAA    962
                     ||....||||||...|||.|||||
miniTurbo        748 ATCAGTCTGCGGTCTGCCGAAAA    770

#=======================================
#
# Aligned_sequences: 2
# 1: BirA
# 2: AirID
# Matrix: EBLOSUM62
# Gap penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 963
# Identity:     647/963 (67.2%)
# Similarity:   647/963 (67.2%)
# Gaps:          14/963 ( 1.5%)
# Score: 3475.0
#
#
#=======================================

BirA               1 ATGAAGGATAACACCGTGCCACTGAAATTGATTGCCCTGTTAGCGAACGG     50
                     ||||||||.||||||||.||.||.|...|.||..|..|..|.||..|.||
AirID              1 atgaaggacaacaccgttccgctcacgcttatctctatccttgctgatgg     50

BirA              51 TGAATTTCACTCTGGCGAGCAGTTGGGTGAAACGCTGGGAATGAGCCGGG    100
                     |||.||.||||||||.||.||..|.||.||...|||.|||||....|||
AirID             51 tgagttccactctggtgaacaacttggagagcagctcggaatgtctaggg    100

BirA             101 CGGCTATTAATAAACACATTCAGACACTGCGTGACTGGGGCGTTGATGTC    150
                     |.||||||||.||.||||||..||||.||.||||||||||.||||||||.
AirID            101 ctgctattaacaagcacatcaagaccctccgtgactggggagttgatgtg    150

BirA             151 TTTACCGTTCCGGGTAAAGGATACAGCCTGCCTGAGCCTATCCAGTTACT    200
                     ||.|..||||..|||||.||.|||.||||.||||||||||||||..|.||
AirID            151 ttcagagttcaaggtaaggggtactgccttcctgagcctatccaacttct    200

BirA             201 TAATGCTAAACAGATATTGGGTCAGCTGGATGGCGGTAGTGTAGCCGTGC    250
                     ..|.|...|..||||...|...|||||.|||||..||...||||..|.||.|
AirID            201 cgacgaagagaagatcaggcagcagcttgatgagggatctgttactgttc    250

BirA             251 TGCCAGTGATTGACTCCACGAATCAGTACCTTCTTGATCGTATCGGAGAG    300
                     |.||.|||||.||.||.||.||.||||||||||.|||.|..||||..|||
AirID            251 tcccggtgatcgattcgaccaaccagtaccttctcgataggctcgatgag    300

BirA             301 CTTAAATCGGGCGATGCTTGCATTGCAGAATACCAGCAGGCTGGCCGTGG    350
                     ||||..||.||.||||..||.||.||.|||.|||||.|||||||..|.||
AirID            301 cttacctctggtgatgtgtgtatcgctgagtaccaacaggctggaagagg    350
```

```
BirA            351 TCGCCGGGGTCGGAAATGGTTTTCGCCTTTTGGCGCAAACTTATATTTGT    400
                    ..||.|.|||.||||.||||||||.||.||.||.||.|||.|.||..|..
AirID           351 aagcagaggtaggaagtggttttctccgttcggagctaacctctacctca    400

BirA            401 CGATGTTCTGGCGTCTGGAACAAGGCCCGGCGGCGGCGATTGGTTTAAGT    450
                    ..||||..|||.|.||.||.|||||.||.||.||.||.||.||..|...|
AirID           401 gcatgtattggagacttgagcaaggacctgctgctgctatgggactttct    450

BirA            451 CTGGTTATCGGTATCGTGATGGCGGAAGTATTACGCAAGCTGGGTGCAGA    500
                    ||.||||||||.||||||||||.||.....|.|..|||||.||.||.||
AirID           451 ctcgttatcggaatcgtgatggctgagactctccaaaagcttggagctga    500

BirA            501 TAAAGTTCGTGTTAAATGGCCTAATGACCTCTATCTGCAGGATCGCAAGC    550
                    ....|||.|.||.||.||||||||.|||||.||.||..|.|||.|.||||
AirID           501 cggtgttagagtgaagtggcctaacgacctttacctcaacgataggaagc    550

BirA            551 TGGCAGGCATTCTGGTGGAGCTGACTGGCAAAACTGGCGATGCGGCGCAA    600
                    |.||.||.||.||.||.|||.|||||||.||.|||||.||.||.||.||.
AirID           551 tcgctggaatcctcgttgagatgactggaaagactggtgacgctgctcat    600

BirA            601 ATAGTCATTGGAGCCGGGATCAACATGGCAATGC----GCC-GTGT-TGA    644
                    ||.||.||||||||.||.||||||.|..|.||||    ||| |.|. |||
AirID           601 atcgtgattggagctggaatcaacctctctatgcgtgagcctgagactga    650

BirA            645 AGAGAGTGTCGTTAATCAGGGGTGGATCACGCTGCAGGAAGCGGGGATCA    694
                    .|||      |||.|.|||...|||||||..||.||.||.||.||.||||
AirID           651 tgag------gttgaccagtcttggatcaacctccaagaggctggtatca    694

BirA            695 ATCTCGATCGTAATACGTTGGCGGCCATGCTAATACGTGAATTACGTGCT    744
                    ...||||||.|.||...|.|.||.||||.|||.||....||..|..|..||
AirID           695 ccatcgatagaaaccagcttgctgccaggctcatcaaggatcttagatct    744

BirA            745 GCGTTGGAACTCTTCGAACAAGAAGGATTGGCACCTTATCTGTCGC-GCT    793
                    ||..|....|..|||||.|||.|||||.|.||.||.|..|| ..|| |.|
AirID           745 gctctcaggcagttcgagcaacaaggacttgctccattcct-cagcagat    793

BirA            794 GGGAAAAGCTGGATAATTTTATTAATCGCCCAGTGAAACTTATCATTGGT    843
                    |||||...||.||.||.||.||.||..|.|||||||.||.|||||.|||
AirID           794 gggaagctctcgacaacttcatcaacaggccagtgaagctcatcatcggt    843

BirA            844 GATAAAGAAATATTTGGCATTTCACGCGGAATAGACAAACAGGGGGCTTT    893
                    ||||.|||.||....||.||..|..|.||||..||.||||.|||||||.|
AirID           844 gatagagagatccacggaatcgctaggggaatcaacgaacaaggggctct    893

BirA            894 ATTACTTGAGCAGGATGGAATAATAAAACCCTGGATGGGCGGTGAAATAT    943
                    .||.||||||||||.||..|.||.||.||.|||||.||.|||||.||..
AirID           894 tttgcttgagcaggacggtgtgattaagccttggattggaggtgagatca    943

BirA            944 CCCTGCGTAGTGC    956
                    .|||..|...|||
AirID           944 gcctcagatctgc    956
```

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