



# When mRNA folding rules gene expression: lessons from type I toxin-antitoxin systems

Sara Masachis Gelo

## ► To cite this version:

Sara Masachis Gelo. When mRNA folding rules gene expression: lessons from type I toxin-antitoxin systems. Human health and pathology. Université de Bordeaux, 2018. English. NNT: 2018BORD0191 . tel-02275785

**HAL Id: tel-02275785**

**<https://theses.hal.science/tel-02275785>**

Submitted on 2 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **When mRNA folding rules gene expression: lessons from type I toxin-antitoxin systems**

Thèse pour obtenir le grade de docteur délivré par l'Université de Bordeaux,

Ecole doctorale Sciences de la Vie et de la Santé, Spécialité Biochimie

Présentée et soutenue publiquement par

**Sara MASACHIS GELO**

*Née le 13 Juillet 1992 à Séville, Espagne*

le 18 Octobre 2018

## **Membres du jury:**

Dr. Harald PUTZER	Université Paris Diderot Paris 7	Président
Prof. Dr. Cynthia SHARMA	University of Würzburg (Germany)	Rapporteur
Prof. Dr. Gerhart WAGNER	University of Uppsala (Sweden)	Rapporteur
Prof. Dr. Christine VARON	Université de Bordeaux	Examineur
Dr. Jean-Jacques TOULME	Université de Bordeaux	Invité
Dr. Fabien DARFEUILLE	Université de Bordeaux	Directeur



*This thesis is dedicated to my father, Ricardo Masachis Rodriguez,*

*and to the memory of my mother, Yolanda Gelo Uruburu, whose  
energy keeps running until the last atom of my cells*



# Abstract/Résumé

## When mRNA folding rules gene expression: lessons from type I toxin-antitoxin systems

Toxin-antitoxin (TA) systems are small genetic modules widely present in bacterial genomes. They usually code for a small toxic protein and its cognate antitoxin and can be classified into six types depending on the nature and mode of action of the antitoxin. This work focuses on the study of type I, for which the antitoxin is an antisense RNA that targets the toxin mRNA to inhibit its expression. We characterized the *aapA3*/IsoA3 system, encoded on the chromosome of the human gastric pathogen *Helicobacter pylori*. To date, most TAs have been studied using artificial expression systems, which do not allow the characterization of transcriptional or post-transcriptional regulation. Taking advantage of the lethality induced by the toxin chromosomal expression in the absence of antitoxin, we developed a high-throughput genetic selection of suppressor mutations revealed by Next-Generation Sequencing. This approach, named FASTBAC-Seq, allowed us to map a myriad of toxicity determinants located in both, coding and non-coding regions, of the *aapA3* toxic gene. More precisely, some suppressor mutations revealed the existence of transient RNA hairpins that act co-transcriptionally to prevent translation initiation while the toxin-encoding mRNA is being made. Such functional RNA metastable structures are essential to uncouple the transcription and translation processes and allow the presence of these toxic genes on bacterial chromosomes. Although untranslated mRNAs become rapidly unstable, our work also revealed the presence of two protective stem-loops located at both mRNA ends that prevent from both, 5' and 3' exonucleolytic activity. Altogether, our work evidenced the consequences of the strong selection pressure to silence toxin expression under which the TAs evolve, and highlighted the key role of mRNA folding in the co- and post-transcriptional regulation of this family of genes. These RNA-based regulatory mechanisms may be exploited in the future for biotechnological (e.g., increased protein production through mRNA stabilization) or biomedical (e.g., development of alternative antimicrobial strategies aiming at the activation of toxin synthesis) applications.

Toxin-antitoxin systems, *Helicobacter pylori*, gene expression, Next-Generation Sequencing, genetic suppressor mutations, RNA folding, RNA stability, transcription/translation coupling

## Lorsque le repliement de l'ARNm gouverne l'expression des gènes : leçons tirées des systèmes toxine-antitoxine de type I

Les systèmes toxine-antitoxine (TA) sont de petits modules génétiques largement présents dans les génomes bactériens. Ils codent pour une petite protéine toxique et une antitoxine. Ils sont classés en six types en fonction de la nature et du mode d'action de l'antitoxine. Ce travail a porté sur l'étude du type I, pour lequel l'antitoxine est un ARN antisens qui cible l'ARNm de la toxine afin de réprimer son expression. Au cours de cette thèse, nous avons étudié le système *aapA3*/IsoA3, codé sur le chromosome du pathogène gastrique humain *Helicobacter pylori*. À ce jour, la plupart des systèmes TA ont été étudiés à l'aide de systèmes d'expression artificiels, qui ne permettent pas de caractériser la régulation transcriptionnelle ou post-transcriptionnelle. En utilisant la létalité induite par l'expression chromosomique de la toxine obtenue en absence d'antitoxine, nous avons développé une sélection génétique de mutants suppresseurs révélés par séquençage haut-débit. Cette approche, appelée FASTBAC-Seq, nous a permis de cartographier une myriade de déterminants de toxicité localisés dans les régions codantes et non codantes du gène de la toxine AapA3. En particulier, certaines de ces mutations ont révélé l'existence de tige-boucles ARN transitoires qui agissent de manière co-transcriptionnelle pour empêcher l'initiation de la traduction pendant la synthèse de l'ARNm codant pour la toxine. Ces structures ARN métastables fonctionnelles sont nécessaires pour découpler les processus de transcription et de traduction et permettent la présence de ces gènes toxiques sur le chromosome bactérien. Bien que les ARNm non traduits deviennent rapidement instables, nos travaux ont également révélé l'existence de deux tige-boucles protectrices situées aux deux extrémités de l'ARNm. Ces structures secondaires empêchent des activités exonucléolytiques agissant en 5' et 3'. Dans l'ensemble, notre travail met en évidence les conséquences de la forte pression de sélection pour limiter l'expression des toxines sous laquelle évoluent les systèmes TA. Cela nous a permis de mieux comprendre l'influence du repliement secondaire des ARNm, non seulement lors de la régulation post-transcriptionnelle, mais aussi co-transcriptionnelle de l'expression de cette famille particulière de gènes. Ces caractéristiques de régulation basées sur l'ARN peuvent être exploitées à l'avenir pour des applications biotechnologiques (p. ex., production accrue de protéines par stabilisation d'ARNm) ou biomédicales (p. ex., développement de stratégies antimicrobiennes alternatives pour l'activation de la synthèse de toxines).

Systèmes toxine-antitoxine, *Helicobacter pylori*, expression génique, séquençage haut débit, mutations suppresseurs, repliement de l'ARN, stabilité de l'ARN, couplage transcription/traduction

This thesis has been performed in the INSERM Unit 1212  
146 Rue Léo Saignat, 33000, Bordeaux, France

# Résumé détaillé

## Contexte de la recherche

*Helicobacter pylori* est une bactérie Gram-négative qui représente un des principaux agents pathogènes gastrique humain. Le génome de cette bactérie fut l'un des premiers séquencés (Tomb et al., 1997a) et seul un petit nombre de facteurs de transcription a pu être prédit sur ce petit génome de 1,67 Mb (Thompson and de Reuse, 2002). En 2010, une combinaison d'approches bio-informatiques et d'analyse ARN par séquençage à grande échelle (dRNA-seq) a permis la caractérisation du transcriptome de *H. pylori*. Ces études ont révélé l'existence de plus de 60 petits ARN non codants (sRNAs) (Sharma et al., 2010). Jusqu'à présent, seul quatre d'entre eux ont été caractérisés. Parmi les plus abondants, il existe une famille intrigante de six ARN antisens homologues (nommés IsoA1 à IsoA6, inhibitor of small QRF family A, locus 1 à 6). Ces ARNs antisens sont exprimés dans le brin d'ADN opposé à une nouvelle classe de petits ARNm (nommés AapA1 à AapA6, représentant : antisense associated peptide family A) (Sharma et al., 2010). En effet des tests de traduction *in vitro* ont montré que chaque petit ARNm AapA code pour un peptide de 30 acides aminés dont l'expression est réprimée spécifiquement par le sARN IsoA. De plus, l'ARNm AapA et le sRNA IsoA sont transcrits de manière constitutive (Sharma et al., 2010). De fait, ces modules d'expression, répétés à six loci différents sur le chromosome d'*H. pylori*, pourraient appartenir à une nouvelle famille de systèmes toxine antitoxine (TA) bactériens.

Les systèmes TA sont de petits modules génétiques composés d'une toxine et d'une antitoxine. Cette dernière est plus instable et elle permet d'empêcher l'activité de la toxine. En fonction de la nature et du mode d'action de l'antitoxine, les systèmes TA peuvent être organisés en six classes (I-VI). Il est important de noter que, les toxines sont toujours des protéines tandis que les antitoxines peuvent être des protéines (dans les systèmes TA de type II, IV, V et VI) ou des ARNs (dans les systèmes TA de type I et III) (pour une revue, voir Goeders et Van Melder, 2014). Les antitoxines de type I sont des ARNs qui peuvent s'apparier avec l'ARNm de leur toxine cible pour inhiber l'initiation de la traduction (en bloquant l'accessibilité SD) et/ou favoriser la dégradation de l'ARNm de la toxine ou de l'hétéroduplex d'ARN (pour une revue, voir Fozo et al., 2010). Une caractérisation plus poussée de la famille *aapA*/IsoA chez *H. pylori* a confirmé son identification comme étant la première famille de systèmes TA de type I identifiée jusqu'à présent chez les epsilonprotéobactéries (Arnion et al., 2017; Sharma et al., 2010).

Dans nos travaux récents (Arnion et al., 2017, et Korkut et al., *in preparation*), le module présent au locus I du chromosome d' *H. pylori* (*aapA*/IsoA1) a été caractérisé. Nous avons montré que le gène *aapA* code pour une petite protéine membranaire dont l'expression est toxique pour *H. pylori*. D'ailleurs, la synthèse de la protéine est réprimée par plusieurs mécanismes de régulation post-transcriptionnelle. La transcription du gène *aapA* génère un ARNm de 250 nucléotides appelé *aapA*-FL (pour « plein taille »), qui est inactif en traduction. Ce même ARNm est raccourci par une coupure à son extrémité 3' pour générer une version tronquée de 225 nucléotides qui elle peut être traduite. Ce transcrit actif, noté *aapA*-Tr (pour « tronqué »), s'apparie avec l'ARN antisens IsoA1 pour former un hétéro-duplex d'ARN étendu qui est rapidement dégradé par la ribonucléase III (RNase III). En l'absence d'IsoA1, la toxine AapA est synthétisée, entraînant un arrêt de croissance de la bactérie et la mort cellulaire.

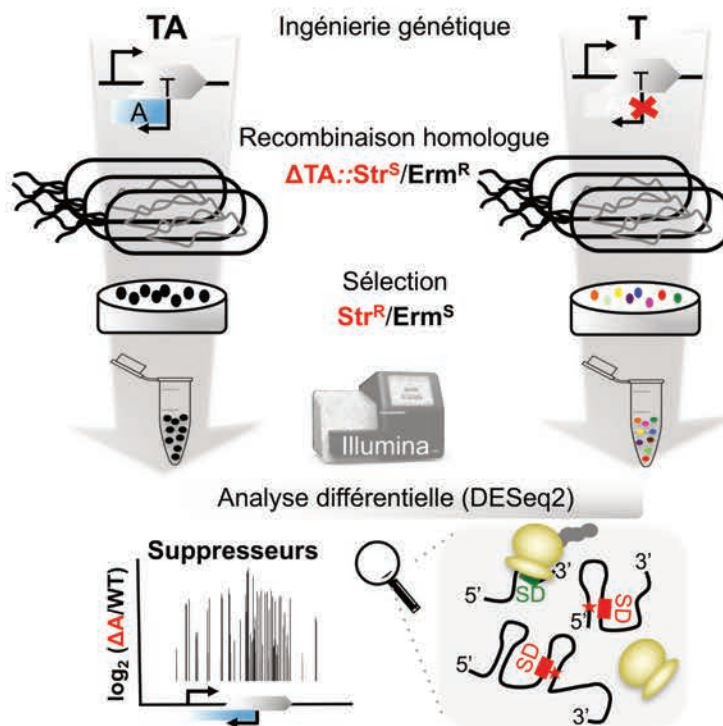
À ce jour, la plupart des systèmes TA ont été étudiés à l'aide de vecteurs d'expression synthétiques ou chez des organismes hétérologues (Fozo et al., 2008; Jahn et Brantl, 2013; Unoson et Wagner, 2008), ce qui peut entraîner des erreurs d'interprétation des effets physiologiques de l'expression des toxines, qui ne se retrouveraient jamais à des niveaux aussi élevés que dans les conditions endogènes. Une autre limite de l'utilisation de tels systèmes d'expression est le fait qu'ils ne

permettent pas d'étudier la régulation transcriptionnelle ou post-transcriptionnelle. En effet la majeure partie de la régulation des systèmes TA de type I se produit après la transcription, il s'agit d'un facteur limitant important pour la compréhension de la régulation de nouveaux loci identifiés.

Dans un premier temps, l'objectif de mon travail de thèse a été de caractériser un nouveau module *aapA*/IsoA récemment identifié au locus III du chromosome de *H. pylori*. Pour cela nous avons développé une nouvelle approche permettant la caractérisation à haut débit d'éléments régulateurs codés en *cis* (c'est-à-dire dans le locus TA) impliqués dans le contrôle de l'expression des systèmes TA de type I au niveau chromosomique.

### Démarche expérimentale

Nous avons utilisé une approche systématique à haut débit appelée FASTBAC-Seq (Functional Analysis of Toxin antitoxin systems in Bacteria by deep-Sequencing) qui combine une sélection génétique de mutants suppresseurs couplée à une technologie de séquençage de nouvelle génération (NGS). En effet, nous avons utilisé la létalité induite par l'expression endogène de la toxine lorsque l'antitoxine est inactivée, pour chercher des mutations supprimant la toxicité. À cette fin, nous avons tout d'abord introduit des mutations ponctuelles inactivant le promoteur de l'antitoxine dans le chromosome de la bactérie, puis sélectionné les clones survivants. L'ADN génomique de ces clones a été isolé afin de cartographier les mutations sur le locus TA par séquençage haut débit.



### Résumé de l'approche FASTBAC-Seq

Au Chapitre 4 section 4.1 (article II), nous avons présenté un protocole détaillé de FASTBAC-Seq. Cette méthode a été appliquée pour étudier le système *aapA3*/IsoA3, appartenant à la famille des systèmes TA de type I *aapA*/IsoA. Cette approche a permis d'identifier une myriade d'éléments régulateurs intégrés dans l'ARNm, qui agissent de concert avec l'ARN antitoxine IsoA3, et qui sont essentiels pour contrôler efficacement l'expression de la toxine AapA3 (Chapitre 4, sections 4.2 et 4.3 (articles III et IV)).

## Résultats

### 1. Le repliement co-transcriptionnel contrôle l'efficacité d'initiation de la traduction

Les résultats obtenus avec l'approche FASTBAC-Seq sont présentés dans la première partie du [Chapitre 4 section 4.2 \(article III\)](#). Dans le premier article nous avons montré que l'expression chromosomique de la toxine AapA3 est létale pour *H. pylori*. En effet nous n'avons pas pu construire une souche inactivée pour le promoteur de l'antitoxine en absence de mutations non-sens dans le gène de la toxine. Nous avons utilisé le phénotype létal pour sélectionner des mutants suppresseurs. Environ 60000 clones bactériens contenant au moins une mutation ont été sélectionnés et séquencés par séquençage haut-débit. Nous avons obtenu majoritairement des mutations ponctuelles qui correspondent soit à des substitutions, des délétions ou des insertions. Les mutations responsables de la suppression de la toxicité chez *H. pylori* ont été cartographiées sur le gène de la toxine. Si la plupart des mutations ponctuelles sont localisés dans le cadre ouvert de la toxine AapA3 (et sont décrits dans le [Chapitre 4 section 4.1 \(article II\)](#)), un tiers correspond à des régions non-codantes du gène de la toxine. Nous avons analysé l'effet des mutations des régions non codantes sur la transcription, la traduction et la stabilité de l'ARNm de la toxine. Dans l'article III, nous nous sommes focalisés sur l'effet de mutations sur l'efficacité de traduction de la toxine.

La traduction est un processus complexe où, globalement, toutes les étapes contribuent au niveau final de production des protéines. Cependant, chez les bactéries, l'initiation de la traduction reste l'étape limitante (Duval et al., 2015; Kozak, 2005; Nakagawa et al., 2017; Simonetti et al., 2009). Le mécanisme d'initiation de la traduction canonique implique l'interaction par appariement de base entre la séquence Shine-Dalgarno (SD, de séquence consensus AGGAG) dans l'ARNm et la séquence anti-SD de l'ARNr 16S incorporé dans la sous-unité ribosomale 30S (Shine et Dalgarno, 1974; Steitz et Jakes, 1975). Ensuite le facteur d'initiation 2 (IF-2) va reconnaître un codon initiateur (c.-à-d. AUG ou UUG / GUG) situé à une distance optimale de la séquence SD (Chen et al., 1994). Ceci forme ce que l'on appelle « le complexe de d'initiation de la traduction ». Plusieurs études ont déjà souligné l'importance de l'accessibilité de la séquence SD pour l'initiation de la traduction (Bhattacharyya et al., 2018; Espah Borujeni et al., 2014; Kozak, 2005). Dans le [Chapitre 4 section 4.2 \(article III\)](#), nous avons présenté une partie des résultats obtenus de l'analyse du locus *aapA3/IsoA3* avec FASTBAC-Seq en se focalisant sur des mutations suppresseurs localisées dans des régions non-codantes. Nous avons montré que si certaines mutations agissent au niveau de la séquence primaire de l'ARNm (c.-à-d. mutation dans la séquence SD), certaines mutations agissent au niveau du repliement de l'ARNm en permettant la séquestration de la séquence SD dans une structure tige-boucle très stable. Ces mutations agissent comme des pièges énergétiques piégeant l'ARNm de la toxine AapA3 dans des conformations qui ne sont pas traductibles. De manière très intéressante, nos résultats suggèrent que ces tige-boucles ARN sont en fait des reliques de structures transitoires formées lors de la transcription de l'ARNm de la toxine AapA3.

En effet, nous avons montré que ces mutations ponctuelles agissent dans un contexte de séquence préexistant, en permettant de stabiliser des structures secondaires d'ARN qui masquent le site de fixation du ribosome. Habituellement ces structures sont formées de manière transitoire afin d'empêcher le démarrage de la traduction durant la synthèse de l'ARNm de la toxine. En effet ces structures métastables permettent de masquer la séquence SD temporairement en attendant qu'une structure plus stable soit formée entre les régions 5' et 3' de l'ARNm. Ce dernier adopte ainsi une structure dite « en trèfle à quatre feuilles » qui est très stable et non traduite, mais qui est lentement maturée en forme plus courte pour devenir traductible. C'est cette forme mature de l'ARNm qui est régulée exclusivement par l'ARN antitoxine. Une étude d'un système TA de type I très similaire (c.-à-d. *hok/Sok* chez *E. coli*) avait décrit l'existence d'une tige-boucle métastable dans la voie de repliement



de l'ARNm de la toxine Hok (Repsilber et al., 1999). Ces études ont été réalisées *in vitro* en utilisant des versions d'ARNm tronqué en analysant par empreinte enzymatique les différentes structures formées. Toutefois, contrairement à notre étude qui a révélé l'existence de structures métastables par des mutations stabilisantes, l'étude *in vivo* sur l'ARNm Hok a révélé l'existence d'une tige-boucle métastable à l'aide d'une mutation déstabilisante (Møller-Jensen et al., 2001). De plus, il est important de noter que dans le cas de l'ARNm de Hok, la structure métastable ne joue pas de rôle direct dans la séquestration de la séquence SD pendant la transcription. Au lieu de cela, elle est requise pour l'interaction à longue distance entre les extrémités des ARNm 5' et 3' (élément *fold-back-inhibition*, fbi) qui obstruent la séquence SD du peptide leader Mok dans le transcrit de pleine taille. Dans notre cas, les structures transitoires ne sont pas de simples structures intermédiaires dans le processus de repliement de l'ARNm, mais ces structures participent au contrôle traductionnel de la toxine. Par conséquent, notre étude représente le premier exemple de structures secondaires transitoires qui agissent directement au niveau co-transcriptionnel sur le démarrage de la traduction, quel que soit le devenir de l'ARNm.

## **2. Deux tige-boucle présentes aux extrémités 5' et 3' non traduites sont responsables de la très forte stabilité de l'ARNm codant pour la toxine AapA3**

Dans l'article III, nous avons décrit comment le repliement de l'ARNm codant pour la toxine AapA3 permet de prévenir la traduction tout au long de sa synthèse. Toutefois, il a été montré que la stabilité des ARNm est fortement diminuée en absence de ribosomes qui les protègent de l'action des ribonucléases. De fait, que se passe-t-il dans le cas des ARNm codant pour des toxines de type I et qui échappent au couplage transcription/traduction. L'approche FASTBAC-Seq a identifié un très grand nombre de mutations suppresseurs capable d'accélérer la dégradation de l'ARNm codant pour la toxine. Nos travaux ont montré que ces mutations suppresseurs sont localisées dans des structures tige-boucle très stables. Ces mutations permettent d'induire une très forte déstabilisation de ces structures par la création d'un seul mésappariement dans la tige. Cette approche a donc mis en évidence l'existence de deux tige-boucles qui permettent de stabiliser l'ARNm codant pour la toxine, de manière à ce qu'il puisse être traduit de manière efficace.

Par ailleurs, sachant que les structures secondaires des ARN influent fortement sur la plupart des activités exonucléases (et certaines endonucléases), nous avons cherché les exonucléases dont l'activité est bloquée par ces structures. Par exemple, il a été démontré chez *E. coli* que la Polynucéotide Phosphorylase (PNPase), la principale exonucléase 3'-5' de *H. pylori*, dégrade l'extrémité 3' de l'ARN jusqu'à ce qu'elle rencontre une structure tige-boucle énergétiquement suffisamment stable (Spickler et Mackie, 2000). De même, une forte sensibilité à la structure secondaire a été démontrée pour la seule exonucléase 5'-3' identifiée chez *H. pylori*, la RNase J ((Redko et al., 2013), montrée aussi chez *E. coli*, Even et al., 2005)). Le rôle clé des structures de l'ARN en 5' dans la stabilité des ARNm a été démontré il y a longtemps (Chen et al., 1991; Emory et al., 1992). Avec le développement récent de l'approche Term-Seq (Dar et al., 2016), un très grand nombre de structures protectrices localisées dans les extrémités 5' et 3' des transcrits dans le transcriptome d'*E. coli* a été rapporté (Dar et Sorek, 2018). L'ARNm d'AapA3 mature possède deux structures en boucle-tige hautement stables aux deux extrémités de l'ARNm, ce qui garantit sa grande stabilité malgré l'absence de traduction. En effet, nous avons montré que la structure tige-boucle présente à l'extrémité 3' fonctionne comme un barrage thermodynamique pour l'activité 3'-exonucléase de la PNPase, et elle est essentielle pour la maturation de l'ARNm de AapA3. En effet, nos travaux suggèrent que l'activation de l'ARNm codant pour la toxine AapA3 requiert une dégradation 3'-5' réalisé par la PNPase. De manière surprenante, des résultats similaires ont été obtenus avec l'antitoxine IsoA3, qui est lui aussi raccourci par une dégradation exonucléotidique 3'-5' (réalisé par la PNPase) (section 4.4 du Chapitre 4). Dans l'ensemble, ce travail a mis en évidence le rôle clé de la structure de

l'ARN secondaire dans la stabilité et l'activation de l'ARNm codant pour la toxine AapA3, ainsi que pour le sARN IsoA3.

Les ARN CRISPR (ARNcr) en sont un exemple intéressant. Ils sont généralement transcrits sous forme de molécules inactives de pleine longueur qui nécessitent d'être maturées pour générer les ARNcr actifs plus courts. La maturation des ARNcr représente un événement clé dans l'activation de CRISPR (Deltcheva et al., 2011), de même que la maturation des ARNm codant des toxines est essentielle pour l'activation des systèmes TA de type I (Masachis et Darfeuille, 2018). Cependant, contrairement aux ARNcr, pour lesquels les endoribonucléases responsables de son activation sont souvent codées dans le cadre du système (CasE, Cas6, Csy4) (Deltcheva et al., 2011), l'identification des activités impliquées dans l'activation de type I ARNm codant la toxine, ne faisant pas partie des modules TA, est difficile. À ce jour, un tel mécanisme d'activation n'a été signalé que dans quelques cas, y compris l'activation exonucléolytique en 3' de l'ARNm de Hok (Franch et al., 1997) et l'activation en 5' de la TisB (Darfeuille et al., 2007), DinQ (Weel-Sneve et al., 2013), ShoB (Kawano et al., 2005) et ZorO (Wen et al., 2017). Cependant, l'identité des enzymes responsables est dans la plupart des cas inconnue. Dans cet article, nous avons cherché à explorer les déterminants de la stabilité de l'ARNm codant pour la toxine AapA3 et à comprendre comment ils influencent sa voie d'activation ainsi que la régulation du système.

Les résultats de mes travaux de thèse ont mis en évidence la forte pression de sélection à laquelle les systèmes TA sont soumis afin d'éviter l'expression des toxines. Ainsi, les systèmes TA ont développé des stratégies de régulation de l'expression des gènes hautement sophistiquées. Pour cela, l'étude des systèmes TA représente un outil permettant la découverte des mécanismes de régulation de l'expression des gènes inexplorés et imprévisibles chez les bactéries. Notamment, cette étude nous a permis de mieux comprendre l'influence du repliement secondaire des ARNm, pas seulement lors de la régulation post-transcriptionnelle, mais aussi, co-transcriptionnelle de l'expression de cette famille particulière de gènes. À l'avenir, la compréhension de ces mécanismes régulateurs basés sur la structure de l'ARN pourrait être exploitée pour le développement des applications biotechnologiques (*p. ex.*, production accrue de protéines par stabilisation d'ARNm) ou biomédicales (*p.ex.*, développement de stratégies antimicrobiennes alternatives pour l'activation de la synthèse de toxines).

# Acknowledgments

The Ph.D. has been a life-changing journey. I want to thank everybody that, in one way or another, contributed to it:

First and foremost, I wish to thank my supervisor, **Dr. Fabien Darfeuille**, for giving me the opportunity to work with him. I've learned that, for a given situation, there are two kinds of persons: those who make things happen, and those who make things possible. Without the latter ones, the first ones are screwed. You made this possible. You trusted on me, despite a turbulent start, and I have no words to describe how grateful I am. You've been an astonishingly awesome mentor to me (yes, I know how much you like my exaggerated descriptions..!). You've taught me that 'there is no magic in science, only good protocols'. Thank you for your constant support, the freedom every time I needed it, and your infinite patience. It has been a pleasure to share my passion for science and to 'fly' through brain-storming ideas with you, thank you.

Next, I would like to thank **Dr. Jean-Jacques Toulmé** for giving me the opportunity to peruse this Ph.D., founded by a Marie Skłodowska-Curie Initial Training Fellowship. I am truly grateful for your trust, support, and respect, thank you. Also, I would like to thank **Dr. Jean-Louis Mergny**, INSERM U1212 director, for giving me the opportunity to work in such an inspiring and exciting research environment.

Thanks to the members of my mi-thèse committee, **Prof. Dr. Peter Redder**, **Prof. Dr. Carole Lartigue**, and **Prof. Dr. Richard Iggo**, for the useful advice and encouragement.

I thank all the members of my thesis jury, **Dr. Harald Putzer**, **Prof. Dr. Gerhart Wagner**, **Prof. Dr. Cynthia Sharma** and **Prof. Dr. Christine Varon**, for taking the time to read and evaluate my work.

Thanks to all the past and present members of our team. **Dr. Cathy Staedel**, you are an example of ambition and humility, thank you for inspiring me, I wish you all the best in your new life chapter. Special thanks to **Dr. Isabelle Iost** for sharing some of your time with me to discuss science, but not only! and for your knowledge of the polysome fractionation protocol. Thanks to **Sandrine Chabas** for your technical support and ability to maintain things in order. Thanks to **Dr. Fanny Boissier** for your positive energy, and for providing me with the purest RNase J! **Dr. Nicolas Tourasse**, thanks for your calmed energy, and for your help with the bioinformatics parts of the work! And thanks to our last newcomer, **Dr. Delphine Bronesky**, for sharing with me the 'Zen' side of life in the lab. Take care of my bench, and keep it zeeenn with the 'manips'!

Special thanks to **Prof. Dr. Cynthia Sharma** for a fruitful and inspiring short-stay in your lab. Thanks to **Dr. Denis Dupuy** for joyful moments in your lab learning how to inject DNA into worms to make them shine in the dark...who knows, one day it may be useful for me! I would also like to thank **Dr. Axel Innis**, **Dr. Martin Teichmann**, **Dr. Hélène Dumay**, **Dr. Stéphane Thore**, **Dr. Sébastien Fribourg** and **Dr. Lionel Minvielle-Sebastia**, for fruitful discussions about science and science as a career. Thanks to **Dr. Laurent Azema** for your (hidden) kindness and your ability to bring reality back. **Jacques Puyol**, thank you for your peaceful vibes and for keeping the lab material in place and clean. I have special thanks to my office-mate, **Eric Dausse**, you are one of the most sensitive persons I've met. Thank you for listening to me every time that I needed it, and thanks for your wise and thoughtful advises. **Kati Ba-Pierozzi**! You bring light and happiness to the lab. Thank you for your great energy, and for taking care of my (multiple) missions and administrative work, I wish you a new life chapter full of joy and dance!

Three years is time enough to make friends. Thanks to all the beautiful and inspiring humans that formed the **MetaRNA Ph.D. group**, it's been an awesomely enriching experience to share this journey with you guys! Now, from the lab...**Shujaat Ali**, we started together, such intense times! Thanks for being always there for me (even at times where I couldn't see further than my gels..hhh). **Finaritra Raoelijoana**, you are fun, I will never forget all the surrealist moments we lived together (and, of course, ...I'll keep using my favorite 'sac a main'!). **Dr. Arghya Sett**! Thanks a lot for your good vibes, your generosity and your sense of care, thanks to you I'm a fan of Indian food now! **Bianca Schnell**, ephemeral but intense time together, thanks for all the great moments! **Lorena Zara**, thanks for your kindness, you helped me feel better in hard moments (...your delicious Italian pasta and tiramisu helped, I'm not gonna hide it!). **Emilie Rousseau**, thanks for your positive energy, and for taking us out of the lab sometimes. **Alba Herrerro**, thanks for your support and for helping me not to forget all my Spanish! **Florian**, thanks for making my worm-injecting adventures so fun! **Dr. Caroline Seefeldt**, thanks for calming me down at rush moments with the thesis, wish you all the best in Uppsala! **Dr. Brita Seip** and **Dr. Aitor Manteca**, thanks for the fruitful discussions about 'the future'. Guys, you really made this journey something memorable, thank you. By the way, 11h30?!

Being here, I also had the opportunity to dance again. Dancing gives me life and thus, I would like to thank the people that make that possible. Thanks to **Michael Meyney** and **Dorotheé Touzanne** for tuning up my mind-body connection. Special thanks to my mentor and choreographer, **Marie Comandu**, you helped me discover and express myself in ways I would have never imagined. Intention, sensitivity, and limit-exploration, it is really all about that (and not only in dance), thank you.

Last but not least, I would like to thank my family for their patience and support. Special thanks to my father, **Ricardo Masachis Rodriguez**, for always helping me to put things into context, and encouraging me in the toughest moments. Thanks to **Cristina Vega Fernandez** for conveying your vital strength. Thanks to my uncles **Manolo Masachis Rodriguez**, **Jose Masachis Rodriguez**, my aunt **Maria Masachis Rodriguez**, and to **Mar Moreno**, for your unconditional love and support.

## Thank you !





# Contents

<b>Abstract/Résumé .....</b>	<b>ii</b>
<b>Résumé détaillé.....</b>	<b>iii</b>
<b>Acknowledgments .....</b>	<b>viii</b>
<b>Contents .....</b>	<b>x</b>
<b>List of Tables .....</b>	<b>xiii</b>
<b>List of Figures.....</b>	<b>xiv</b>
<b>Abbreviations .....</b>	<b>xvi</b>
<b>Units .....</b>	<b>xviii</b>
<b>Multiples .....</b>	<b>xviii</b>
<b>CHAPTER 1. Introduction .....</b>	<b>19</b>
1.1. <i>Helicobacter pylori</i> .....	19
1.1.1. Colonization and virulence .....	19
1.1.1.1. Adaptation to acidic pH: urease, taxis and cell shape.....	19
1.1.1.2. Secreted effectors.....	20
1.1.2. Natural competence and homologous recombination.....	21
1.2. Transcriptional regulators in <i>H. pylori</i> .....	23
1.3. Post-transcriptional regulation in <i>H. pylori</i> .....	24
1.3.1. mRNA stability .....	24
1.3.2. Translational regulation .....	28
1.3.3. Small regulatory RNA in bacteria .....	29
1.3.3.1. Biogenesis.....	29
1.3.3.2. Modes of action.....	30
1.3.3.3. Small regulatory RNAs in <i>H. pylori</i> .....	34
1.4. Type I Toxin-antitoxin systems .....	35
1.4.1. Paper I: Type I toxin-antitoxin systems: regulating toxin expression via Shine-Dalgarno sequence sequestration and small RNA binding.....	35
1.4.2. Biological role of type I TA systems.....	36
<b>CHAPTER 2. Aims of this thesis.....</b>	<b>37</b>
2.1. Research background: a new family of type I TA systems in <i>H. pylori</i> .....	37
2.2. Understanding the regulation and biological role of type I TAs in <i>H. pylori</i> .....	39
2.3. Our experimental strategy.....	40
<b>CHAPTER 3. Thesis map .....</b>	<b>42</b>
3.1. Organization.....	42
3.2. Main scientific outputs.....	42

3.3. Contribution by others .....	43
<b>CHAPTER 4. Results and discussion.....</b>	<b>45</b>
4.1. Paper II: FASTBAC-Seq: Functional Analysis of Toxin-antitoxin systems in Bacteria by deep-Sequencing.....	45
4.2. Paper III: A genetic selection reveals functional metastable structures embedded in a toxin-encoding mRNA.....	46
4.2.1. Results.....	46
4.2.2. Discussion.....	59
4.3. Paper IV: When mRNA folding rules decay: lessons from a type I toxin-antitoxin system 66	
4.3.1. Results.....	67
4.3.2. Discussion.....	75
4.4. On the decay of a sRNA antitoxin .....	80
4.4.1. Results.....	80
4.4.2. Discussion.....	84
<b>CHAPTER 5. Conclusion and outlook .....</b>	<b>88</b>
5.1. Conclusion .....	88
5.2. Open questions.....	90
5.2.1. Transcriptional regulation of the <i>IsoA3</i> sRNA? .....	90
5.2.2. Other <i>IsoA3</i> targets?.....	92
5.2.3. Missing actors in the <i>AapA3</i> mRNA activation? .....	96
5.2.4. Start codon accessibility, one more regulatory layer? .....	98
5.2.5. A 3' end standby site in a <i>AapA3</i> mRNA species? .....	102
5.2.6. The inner membrane, the <i>AapA3</i> target? .....	107
<b>CHAPTER 6. Material and Methods.....</b>	<b>113</b>
6.1. Material .....	113
6.2. General Methods.....	120
6.2.1. <i>E. coli</i> growth conditions and strain generation .....	120
6.2.1.1. Transformation of <i>E. coli</i> chemically competent cells .....	120
6.2.2. <i>H. pylori</i> growth conditions and strain generation .....	120
6.2.2.1. <i>H. pylori</i> chromosomal manipulation techniques .....	121
6.2.2.2. Generation of an <i>H. pylori</i> 26695 streptomycin resistant strain .....	121
6.2.2.3. Deletion of the <i>H. pylori</i> 26695 <i>aapA3</i> /IsoA3 locus using the <i>rpsL<sub>CJ</sub>-erm</i> cassette .....	121
6.2.2.4. <i>AapA3</i> /IsoA3 locus sub-cloning in <i>E. coli</i> .....	122
6.2.2.5. Mutant generation by Site-Directed mutagenesis PCR .....	122
6.2.2.6. Determining <i>H. pylori</i> transformation efficiency .....	122

6.2.2.7. <i>H. pylori</i> transformation assay to identify toxicity suppressors by Illumina sequencing.....	123
6.2.3. Nucleic acid techniques .....	123
6.2.3.1. Agarose gel electrophoresis .....	123
6.2.3.2. Polyacrylamide gel electrophoresis .....	123
6.2.3.3. Total RNA preparation with hot phenol .....	124
6.2.3.4. DNase I digestion.....	124
6.2.3.5. Radioactive labeling of DNA oligonucleotides for RNA detection .....	124
6.2.3.6. Northern Blot .....	124
6.2.3.7. Polysome fractionation in sucrose gradients.....	124
6.2.3.8. Cold and hot <i>in vitro</i> transcription.....	125
6.2.3.9. 5' end labeling of RNA.....	125
6.2.4. Protein techniques.....	126
6.2.4.1. One-dimensional SDS PAGE .....	126
6.2.4.2. <i>In vitro</i> translation assays.....	126
6.3. Methods Paper III .....	127
6.3.1. Bacterial strains and oligonucleotides .....	127
6.3.2. <i>In vitro</i> structure probing .....	127
6.3.3. RNase H1/oligonucleotide accessibility assays .....	127
6.6. Methods Paper IV .....	133
6.4.1. Bacterial strains and oligonucleotides .....	133
6.4.2. Rapid amplification of cDNA 3' ends (3'RACE) .....	134
6.4.3. RNA chemical synthesis and purification .....	134
6.4.4. UV melting experiments.....	135
6.4.5. RNase J overexpression and purification.....	135
6.4.6. RNase J <i>in vitro</i> cleavage assays .....	135
6.4.7. Mapping of RNase J and RNase III endonucleolytic cleavages .....	136
<b>CHAPTER 7. References .....</b>	<b>139</b>
<b>CHAPTER 8. Appendices .....</b>	<b>157</b>
8.1. Appendix Paper III.....	157
8.2. Appendix Paper IV .....	170
8.3. Related paper I: Mechanistic insights into type I toxin antitoxin systems in <i>Helicobacter pylori</i> : the importance of mRNA folding in controlling toxin expression...	185
<b><i>Curriculum vitae</i> .....</b>	<b>199</b>

## List of Tables

Table 6.3.1. Inserts of <i>rpsl<sub>C</sub>-erm</i> and <i>aapA3</i> /IsoA3 plasmids used in Paper III. ....	128
Table 6.3.2. Sequence of T7 RNAs used for in vitro studies in Paper III. ....	132
Table 6.4.1. Inserts of <i>aapA3</i> /IsoA3 mutant plasmids used in Paper IV. ....	136
Table 6.4.2. Sequence of T7 RNAs used for <i>in vitro</i> studies in Paper IV. ....	138
Table 6.4.3. Sequence of chemically synthesized RNAs used in Paper IV. ....	138
Table 8.1.1. <i>Helicobacter pylori</i> strains used in Paper III. ....	165
Table 8.1.2. Plasmids used in Paper III. ....	166
Table 8.1.3. <i>Escherichia coli</i> strains used in Paper III. ....	167
Table 8.1.4. Oligonucleotides used in Paper III. ....	168
Table 8.2.1. Minimum Free Energy and padj of the here studied single nucleotide substitutions in the <i>aapA3</i> 5'-terminal stem-loop. ....	176
Table 8.2.2. Minimum Free Energy and padj of the here studied tested single nucleotide substitutions in the <i>aapA3</i> 3'-terminal stem-loop. ....	176
Table 8.2.3. <i>Helicobacter pylori</i> strains used in Paper IV. ....	177
Table 8.2.4. <i>Escherichia coli</i> strains used in Paper IV. ....	179
Table 8.2.5. Plasmids used in Paper IV. ....	180
Table 8.2.6. Oligonucleotides used in Paper IV. ....	181

# List of Figures

Figure 1.1. Growth phase-dependent evolution of <i>H. pylori</i> cell morphology.....	20
Figure 1.2. <i>H. pylori</i> natural competence. ....	22
Figure 1.3. Possible pathways of mRNA decay in <i>H. pylori</i> .....	27
Figure 1.4. Mechanisms of action of <i>cis</i> -encoded sRNAs.....	31
Figure 1.5. Mechanisms of action of <i>trans</i> -encoded sRNAs.....	33
Figure 2.1. Location of the <i>aapA</i> /IsoA modules in the <i>H. pylori</i> chromosome.....	37
Figure 2.2. Model of the <i>aapA1</i> /IsoA1 system regulation.....	38
Figure 2.3. The AapA1 peptide is organized in three domains. ....	39
Figure 2.4. Pipeline of the FASTBAC-Seq approach.....	40
Figure 4.2.1. IsoA3 small RNA is essential to prevent AapA3 translation. ....	47
Figure 4.2.2. IsoA3 inhibits <i>aapA3</i> -Tr translation by binding to its SD region. ....	48
Figure 4.2.3. Unveiling local toxicity determinants with nucleotide resolution.....	50
Figure 4.2.4. Single-point mutations in the <i>aapA3</i> 5' UTR inhibit its translation in absence of IsoA3.....	53
Figure 4.2.5. Single nucleotide substitutions in the AapA3 5'UTR stabilize SD occlusion by an upstream to SD aSD sequence. ....	55
Figure 4.2.6. A synonymous single-point mutation within a downstream to SD aSD sequence motif stabilizes SD occlusion and hampers AapA3 translation. ....	58
Figure 4.2.7. Working model of the <i>aapA3</i> co- and post-transcriptional translational SD-accessibility regulation. ....	61
Figure 4.2.8. The two successive AapA3 mRNA metastable structures have increasing stability and are stabilized by the A40T and the T78C suppressors.....	62
Figure 4.2.9. Comparison of the co- and post-transcriptional translational regulation of the <i>hok</i> /Sok and <i>aapA3</i> /IsoA3 type I TA systems. ....	63
Figure 4.3.1. Rifampicin treatment reveals an exonucleolytic activation pathway of the <i>aapA3</i> mRNA led by PNPase. ....	67
Figure 4.3.2. 3' RACE reveals heterogeneous populations of the three main 3' end-processed <i>aapA3</i> mRNA species. ....	69
Figure 4.3.3. The active <i>aapA3</i> mRNA harbors a 3' end protective stem-loop structure. ....	72
Figure 4.3.4. The 5' stem-loop of the active <i>aapA3</i> mRNA avoids 5'-exonucleolytic degradation. ....	74
Figure 4.3.5. Functional AapA3 mRNA stability is ensured by two termini protective stem-loop structures.....	78
Figure 4.4.1. <i>E. coli</i> PNPase activity on the AapA3 and IsoA3 RNAs is strongly impaired. ....	81
Figure 4.4.2. IsoA3 sRNA processing by RNase J is sensitive to the 5' phosphorylation state.....	83
Figure 4.4.3. Model of IsoA3 maturation and decay.....	86
Figure 4.4.4. Putative single-stranded Rho-binding sites on the IsoA3-p sRNA. ....	87
Figure 5.1. The orphan response regulator HP1043 might act as a transcriptional activator of IsoA3. ....	91
Figure 5.2. IsoA3 sRNA is associated with translating ribosomes in a AapA3-independent manner. ....	92

Figure 5.3. IsoA3 <i>in silico</i> predicted interaction with the <i>aspA</i> and <i>acsA</i> mRNAs.....	93
Figure 5.4. IsoA3 <i>in silico</i> predicted interaction with the HP0893 mRNA.....	94
Figure 5.5. Regions of IsoA3 sRNA interaction with its <i>in silico</i> predicted targets.....	95
Figure 5.6. IsoA3 predicted interaction with the AapA3 and AapA5 mRNAs. ....	96
Figure 5.7. STRING <i>rnpA</i> functional association network.....	97
Figure 5.8. Start codon accessibility determines the visibility of AapA1 translation initiation from non-canonical initiation codons. ....	99
Figure 5.9. Consequences of the stability of the start-codon harboring step-loop in the AapA3 mRNA of two different <i>H. pylori</i> strains.....	100
Figure 5.10. Translation start codon accessibility determines the visibility of AapA1 translation stop codon misreading events.....	101
Figure 5.11. The <i>aapA3</i> -FFL species harbors a 3' end standby site but cannot be translated.....	104
Figure 5.12. Distribution of AU-CA-rich motifs in the <i>aapA3</i> mRNA. ....	106
Figure 5.13. AapA3 leaking chromosomal expression promotes the formation of coccoïd cells at early growth phase stages in <i>H. pylori</i> .....	108
Figure 5.14. Using T7 phage lysozyme as a reporter of inner-membrane damage in <i>E. coli</i> .....	110
Figure 5.15. In absence of LysS no coccoïd-like cells are formed upon toxin induction in <i>E. coli</i> . ...	111
Figure 8.1.1. Details on <i>aapA3</i> /IsoA3 locus deletion and deep-sequencing approaches. ....	157
Figure 8.1.2. Comparison of the distribution and relative frequency of single-nucleotide suppressors in the WT and pIsoaA3* <i>aapA3</i> /IsoA3 modules .....	158
Figure 8.1.3. Defining and validating AapA3 promoter with nucleotide resolution. ....	159
Figure 8.1.4. The A28C, A40T and T78C suppressors inhibit <i>aapA3</i> -Tr mRNA translation. ....	160
Figure 8.1.5. Gel analysis of RNase H/ oligonucleotide accessibility assays.....	161
Figure 8.1.6. Nucleotide alignment of AapA3 coding region of 49 <i>Helicobacter pylori</i> strains. ....	162
Figure 8.1.7. Quantification of the relative <i>aapA3</i> mRNA band intensity from polysome fractionation Northern Blots shown in Figure 4.....	163
Figure 8.1.8. Only three out of the thirteen potential aSD sequenced embedded in <i>aapA3</i> mRNA are functional. ....	164
Figure 8.2.1. PNPase and RNase Y are involved in <i>aapA3</i> mRNA and IsoA3 RNA processing. ....	170
Figure 8.2.2. 3'RACE on <i>aapA3</i> mRNA revealed a 3' end activation pathway lead 3'-5'-exonucleases and roadblocks. ....	171
Figure 8.2.3. Toxicity suppressor mutations cluster symmetrically in the termini stem-loops of the active <i>aapA3</i> mRNA.....	172
Figure 8.2.4. Denaturing PAA gel reveals dramatic stability differences of the 5'- and 3'-terminal stem-loops caused by the single nucleotide mutations G24A and G183A.....	173
Figure 8.2.5. The RhpA helicase activity is not required for the <i>aapA3</i> G24A-Tr 5'-exonucleolytic degradation.....	173
Figure 8.2.6. The 5'- exonucleolytic decay of <i>aapA3</i> mRNA is independent of the 5' phosphorylation state. ....	174
Figure 8.2.7. RNase J acts mainly as endonuclease on free <i>aapA3</i> -Tr RNA. ....	175

# Abbreviations

(v/v)	(volume/volume)
(w/v)	(weight/volume in g/ml)
A	adenine
aa	amino acid
AapA	Antisense Associated Peptide family A
Amp	ampicillin
aphA-3	kanamycin cassette
APS	ammonium persulfate
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	base pair
CDS	coding sequence
C	cytosine
catGC	chloramphenicol cassette
cag PAI	cag Pathogenicity Island
cag	Cytotoxin Associated genes
cag A	Cytotoxin Associated gene A
cDNA	complementary DNA
CTP	cytidine triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide
DTT	dithiothreitol
E. coli	Escherichia coli
EDTA	ethylenediamine tetraacetate
Fig.	figure
G	guanine
gDNA	genomic DNA
GFP	green fluorescent protein
GTP	guanosine triphosphate
Hfq	host factor for <i>Q<math>\beta</math></i> replication
H. pylori	Helicobacter pylori
IGR	intergenic region
IsoA	Inhibitor of Small ORF family A
Kan	kanamycin
kcal	kilocalorie
LB	Lennox broth
LPS	lipopolysaccharide
mfe structure	minimum-free-energy structure
mRNA	messenger RNA
miRNA	microRNA
NCBI	National Center for Biotechnology Information

ncRNA	noncoding RNA
nt	nucleotide
OD	optical density
OMP	outer membrane protein
ORF	open reading frame
PAA	polyacrylamide
PBS	phosphate buffered saline
PCR	polymerase chain reaction
RACE	rapid amplification of cDNA ends
RBS	ribosome binding site
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
rNTP	ribonucleotide
sRNA	small RNA
siRNA	short-interfering RNA
SB	structure buffer
SD	Shine-Dalgarno
SDS	sodiumdodecylsulfate
SSC	saline sodium citrate buffer
SST4	type IV Secretion System
SVF	sérum de veau foetal
T	thymine
TA	toxin-antitoxin
TBE	Tris/Borate/EDTA
TEMED	tetramethylethylenediamin
TIR	translation initiation region
Tris	tris-(hydroxymethyl)-aminomethan
TSS	transcription start site
tRNA	transfer RNA
U	uracil
UTP	uridine triphosphate
UTR	untranslated region
VacA	vacuolating toxin A



## Units

°C	degree Celsius
a.u	arbitrary unit
Da	Dalton
g	gram
h	hour
l	litre
M	molar
min	minute
molar	gram molecule
rpm	rounds per minute
s	second
u	unit
V	Volt
W	Watt

## Multiples

M	mega ( $10^6$ )
k	kilo ( $10^3$ )
m	mili ( $10^{-3}$ )
$\mu$	micro ( $10^{-6}$ )
n	nano ( $10^{-9}$ )
p	pico ( $10^{-12}$ )

## CHAPTER 1. Introduction

### 1.1. *Helicobacter pylori*

The human stomach, long believed to be a sterile organ, is, in reality, the niche for up to 200 different bacterial species. However, there is one species that, when present, tends to be numerically dominant (despite not changing the composition of the bacterial community), *Helicobacter pylori* (Bik et al., 2006). The *Helicobacter* genus is composed of Gram-negative bacteria that belong to the epsilon subdivision of Proteobacteria. So far, more than 20 different species have been described as belonging to the *Helicobacter* genus (Fox, 2002). They can be divided into gastric helicobacters, specialized in the colonization of the gastric mucosae and epithelial layers of their host (e.g., *H. pylori*, colonizing the human stomach, and the *H. himmannii* and *H. felis*, colonizing the stomach of animals but having zoonotic potential), and non-gastric helicobacter or enterohepatic helicobacters, that colonize the lower intestinal tract (i.e., intestine crypts) and the liver of animals -including humans- (e.g., *H. canis*, *H. pullorum*, *H. cinaedi* and *H. fennelliae*). Importantly, gastric species are not able to colonize the niche of enterohepatic species, and the same principle applies the other way around, which have important clinical consequences (Solnick and Schauer, 2001). Despite being a big family, among the various *Helicobacter* species, *H. pylori* colonizes the stomach of about 50% of the world's population. Therefore, *H. pylori* has the highest impact in human health and is the most studied and well-characterized species.

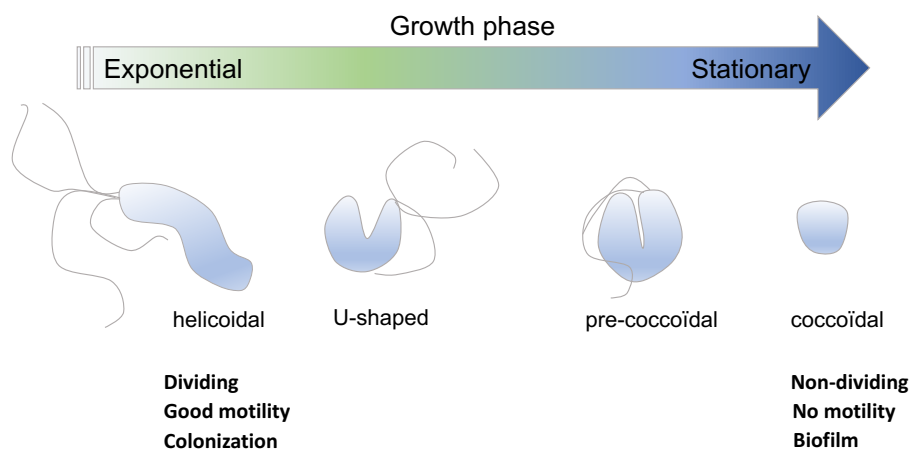
#### 1.1.1. Colonization and virulence

*H. pylori* is uniquely specialized to live in the human stomach. This specialization is thought to have been taking place for at least the last 100,000 years (Moodley et al., 2012). Modern *H. pylori* strains have been assigned to different populations based on their geographic origin (hpEurope, hpSahul, hpEastAsia1, hpAsia2, hpNEAfrica, hpAfrica1 and hpAfrica2), and taking advantage of its characteristic early childhood oral transmission (i.e., fast transmission within families), they have even been used as tracers of complex demographic events in the human prehistory (Falush et al., 2003; Maixner et al., 2016; Moodley and Linz, 2009; Moodley et al., 2009; Wirth et al., 2004). *H. pylori* infection is associated with chronic gastritis (usually asymptomatic until adulthood), peptic ulcers, and in fewer cases, gastric cancer (Kusters et al., 2006; Moss, 2017; Plummer et al., 2015; Wroblewski et al., 2010), being therefore the only bacterium species classified within the group I carcinogens by the World Health Organization (1994).

##### 1.1.1.1. Adaptation to acidic pH: urease, taxis and cell shape

The production of gastric acid results in a pH of 1-2 in the stomach lumen, which strongly limits luminal colonization. In order to survive, *H. pylori* needs to overcome the mucus layer in the stomach and rapidly migrate (within few minutes) to the gastric epithelial surface (Schreiber et al., 2005). To achieve this quick migration, *H. pylori* actively swims towards the mucous layer and the epithelial barrier guided by (a.k.a. taxis) pH (mediated by the TlpB chemotaxis receptor, Croxen et al., 2006), as well as several compounds, including mucin (mediated by CheY and CheA, Foyes et al., 2000), amino acids (Worku et al., 2004), sodium chloride (Mizote et al., 1997) and sodium bicarbonate, in a urease-dependent manner (Nakamura et al., 1998). Indeed, having a functional urease activity is essential for *H. pylori* acid resistance as it buffers the pH in the periplasm (Sachs et al., 2002) and has been demonstrated to be essential for *H. pylori* chronic infection (Debowski et al., 2017). Additionally, urease-catalyzed production of ammonium ions can greatly improve motility, as it raises the pH to near

neutrality and this leads to a mucus gel transition to a viscoelastic solution through which *H. pylori* can efficiently swim (Celli et al., 2009; Jonathan P. Celli et al., 2007). Additionally, the helical shape of *H. pylori* cells facilitates motility through the viscous mucus layer by a corkscrew mechanism, and shape mutants show impaired colonization (Bonis et al., 2010; Sycuro et al., 2012). Interestingly, as highlighted by Keshavarz et al., (1999) and Cellini, (2014), there is a relationship between *H. pylori* motility, cell-shape, and phase of growth, that have great consequences in the colonization and virulence processes. As illustrated in Figure 1.1, in the exponential growth phase, cells are dividing and have helical shape with multiple polar flagella. These cells have optimal motility and are therefore able to rapidly colonize the niche, being abundantly found in the mucus layer. Moving towards the stationary growth phase, cells progressively become coccoïdal and shed the flagella. These cells, no longer divide (they are Viable But Non Cultivable Cells or VBNC) and loss their motility. Nevertheless, coccoïds are abundantly found in *H. pylori* biofilms, suggesting a role in persistent colonization (Carron et al., 2006; García et al., 2014; Hathroubi et al., 2018).



**Figure 1.1. Growth phase-dependent evolution of *H. pylori* cell morphology.** During exponential growth, *H. pylori* cells are actively dividing and have typical helicoidal shape, which allows an optimum motility through the mucus layers of the human stomach to colonize and adhere to the epithelial cells. Approaching the stationary growth phase, cells progressively acquire coccoïdal shape, passing through the intermediates U-shape and pre-coccoïdal shape. At this stage, cells are no longer actively dividing (are what is called Viable But Non Cultivable Cells or VBNC), and they have lost their motility. However, it is well known that coccoïd cells, have increased tolerance of stress conditions and have important biological roles, as for instance, in *H. pylori* biofilms in the human stomach.

### 1.1.1.2. Secreted effectors

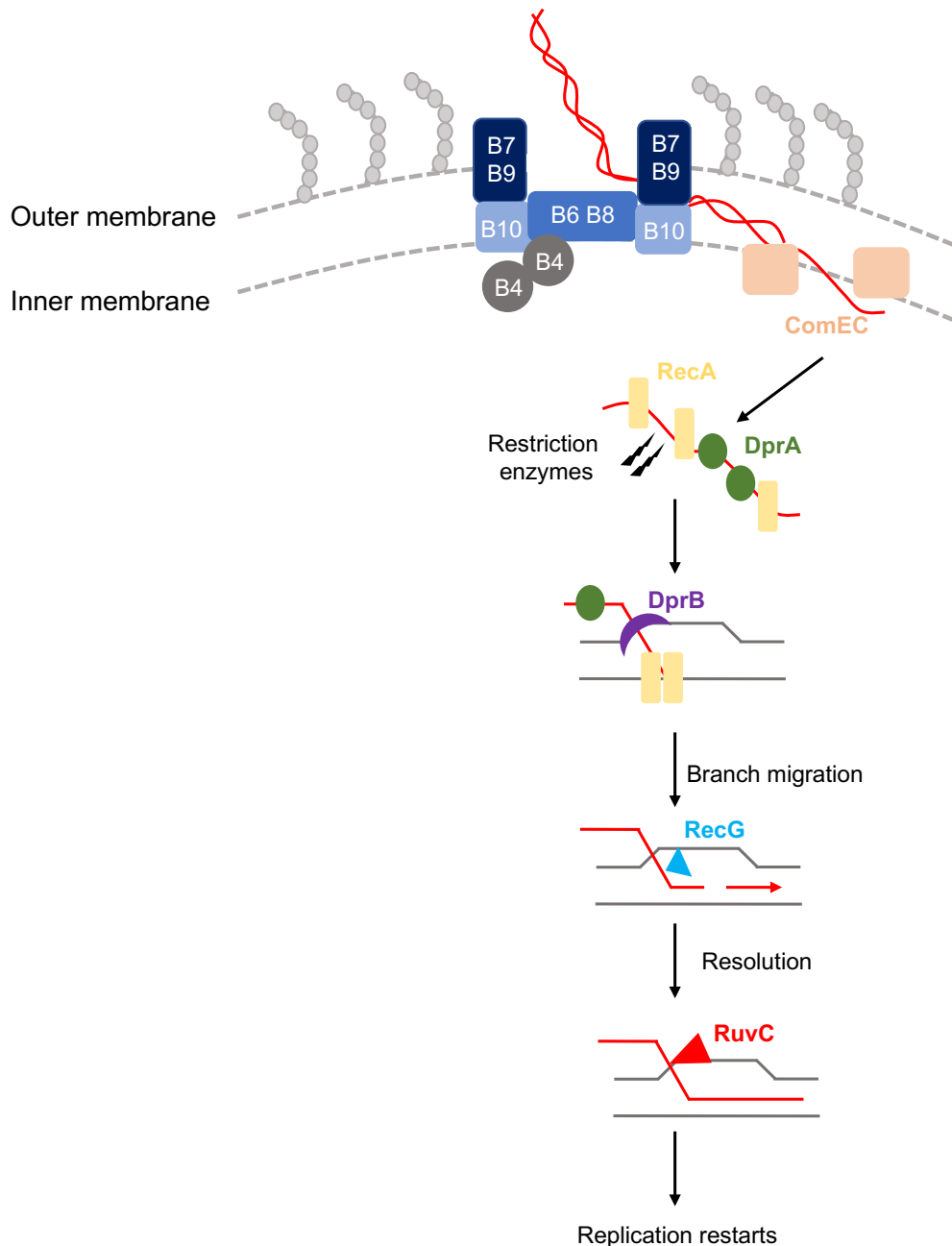
Once *H. pylori* crosses the mucus layer and reaches the epithelial surface, it can actively manipulate the host tissues to promote its own persistence by secreting (mainly) two protein effectors: VacA and CagA. The vacuolating cytotoxin A (VacA, 140 kDa) is harbored by all characterized *H. pylori* strains and interestingly presents sequence variation in several domains that are linked to expression levels and virulence (Palframan et al., 2012). VacA is able of upregulating chemokine expression in human eosinophils (causing inflammation) (Kim et al., 2007), increasing the paracellular epithelial permeability of polarized cells (Papini et al., 1998) and of using it to go through the disrupted tight junctions to inhibit T-cell proliferation (Sundrud et al., 2004). The second toxin is the cytotoxin-associated gene A (CagA, 128 kDa), originally identified as an immunodominant antigen from patients infected with highly virulent *vacA* alleles (Covacci et al., 1993; Tummuru et al., 1993). CagA is translocated into host cells by the Cag type VI secretion system (T4SS), a needle-like encoded on the

cag pathogenicity island (cag-PAI) (Chang et al., 2018; Kusters et al., 2006; Murata-Kamiya, 2011; Olbermann et al., 2010). Once inside, CagA can get phosphorylated in the EPIYA motif by the SRC and ABL family kinases (Nagase et al., 2011) and tyrosine-phosphorylated by the focal-adhesion kinase Src (Kwok et al., 2007) leading to host phenotypes that confer nutritional benefits (Tan et al., 2011; Wu et al., 2013).

### 1.1.2. Natural competence and homologous recombination

*H. pylori* is a so-called naturally competent bacterium. Natural competence is defined as the ability to take up DNA from the environment and recombine it into the genome, leading to genetic transformation. Regarding the first step of competence, *H. pylori* is the sole competent bacterium described so far that uses a type IV secretion system (T4SS, the Com apparatus, shown as B7B9, B6B8, B10 and B4 in Figure 1.2) as an alternative to the type IV pilus apparatus (which is not present on its genome) to transport DNA through the inner and outer membranes (Fischer et al., 2001; Hofreuter et al., 2001; Mell and Redfield, 2014). This strategy is surprising, as T4SS are more often used to translocate DNA (*i.e.*, conjugative plasmids) (Grohmann et al., 2003) and proteins (*e.g.*, effectors) (Alvarez-Martinez and Christie, 2009; Boschioli et al., 2002; den Hartigh et al., 2008) for virulence purposes. The *H. pylori* T4S Com apparatus has two main particularities: 1) it has an opposite polarity when compared to the other T4SS (the Com T4SS translocate DNA into the cell instead of exporting it), and 2) it works with only one ATPase motor instead of the three found in other T4SS (*e.g.*, the VirB/D4 T4SS of *Agrobacterium tumefaciens* (Atmakuri et al., 2004)) (for a review see, Christie et al., 2005). However, it uses at least one component for DNA transport that is highly conserved between all T4SS, the ComEC protein (shown in light orange in Figure 1.2).

Events occurring during natural transformation can be divided into two groups, 1) the presynaptic events, and 2) the postsynaptic events. Although the exact molecular underpinnings of competence and recombination are unclear, DNA uptake through the T4S Com apparatus likely leads to single-stranded cytosolic DNA fragments. This kind of DNA substrate is different from the one that is normally used during DNA repair, and may, therefore, be recombined through a distinct pathway (*i.e.*, independent of the RecOR and AddAB proteins, (Humbert et al., 2011; Marsin et al., 2008, 2010)). As shown in Figure 1.2., once a single-stranded DNA enters the cytoplasm, DprA (DNA processing chain A) and RecA (recombinase A) cooperatively bind to it (Mortier-Barrière et al., 2007) (shown to be essential for *H. pylori* natural transformation, Schmitt et al., 1995; Smeets et al., 2000; Thompson and Blaser, 1995). Interestingly, the DprA protein in *H. pylori* has been shown to have a double function, it can not only inhibit the activity of restriction enzymes but also stimulate methyltransferases, thus, modulating the R-M barrier during inter-strain natural transformation (Dwivedi et al., 2013). Together with DprA and RecA, a putative resolvase homolog, DprB (shown in purple in Figure 1.2), is required for high-frequency natural transformation in *H. pylori* (Humbert et al., 2011) but not for DNA repair, suggesting that all three compose the presynaptic pathway for the homologous recombination of DNA substrates taken in by the T4S Com apparatus in *H. pylori*.



**Figure 1.2. *H. pylori* natural competence.** Double-stranded DNA is taken into the periplasm by the ComB apparatus (B7B9, B6B8, B10 and B4) and transported into the cytosol through the ComEC channel. If the DNA is from an unrelated strain, restriction enzymes may cleave it limiting the size of the DNA. RecA homology searching leads to the formation of a D-loop structure that is modified and extended during replication by RecG-mediated branch migration. Resolution can occur mediated by RuvC which cleaves the Holliday junction or by RuvC-independent mechanism in which RecG pushes the Holliday junction until the end of the invasive DNA fragment. After resolution, replication can restart unless a second 3' end invasive DNA fragment is present leading to the generation of double crossover products (not shown). The Figure was adapted from (Dorer et al., 2011).

As shown in Figure 1.2., the two major actors of the postsynaptic events are the ATP-dependent DNA helicase RecG (shown in light blue in Figure 1.2), and the crossover junction endodeoxyribonuclease RuvC (shown in red in Figure 1.2), however, their role is still nowadays a controversial issue. Postsynaptic homologous recombination (HR) factors need to control the length of

the DNA that is to be integrated into the genome (Humbert et al., 2011). It is known that DNA length and sequence homology are key factors determining transformation efficiency. However, there is a certain degree of tolerance to mismatches, as a 4% difference was shown to have no effect on *H. pylori* transformation efficiency (Humbert et al., 2011). Furthermore, during transformation with heterologous DNA, longer integration lengths (~1 to 8 kb) have been measured when a whole gene was transferred (Humbert et al., 2011) than when one point mutation was integrated (~0.9-1.8 kb) (Lin et al., 2009), reflecting that longer flanking homologous DNA regions are needed for the integration of long DNA fragments. Consistently with this, RecG is able to increase the length of the integrated DNA by helping branch migration, whereas RuvC decreases the length by resolving the Holliday junction (Figure 1.2). Altogether, it seems that the HR pathway used for *H. pylori* natural transformation is different from that of DNA repair. Interestingly, RuvC defective mutants show reduced colonization (Loughlin et al., 2003), suggesting an important role of natural transformation during infection.

## 1.2. Transcriptional regulators in *H. pylori*

Bacteria use transcriptional regulatory networks (TRN) to transduce environmental signals into a coordinated expression of their genomes, resulting in an appropriate adaptive response. Usually, prokaryotic TRNs are highly organized, multilayered, hierarchical structures involving hundreds of transcription factors (TFs), which make them highly challenging to study (Babu et al., 2009). The small *H. pylori* genome (1.67 Mb) encodes for a rather narrow range of transcription factors, including only three sigma factors (i.e., RpoD, FliA and RpoN) (Tomb et al., 1997), which make of it an interesting model for the study of TRN in bacterial pathogens (Danielli et al., 2010; Scarlato et al., 2001). Indeed, to date, only 17 *bona fide* TFs have been characterized in *H. pylori*, including only three regulatory two-component systems (Scarlato et al., 2001; Tomb et al., 1997). Four main TRN origins have been identified in *H. pylori*, and are related key physiological responses during infection: 1) acid acclimation, 2) heat and stress response, 3) taxis and motility, and 4) metal ion homeostasis (Danielli et al., 2010).

In order to survive in the highly acidic conditions of the human stomach, *H. pylori* has a specialized set of genes termed acid acclimation genes (over 100 genes (Pflock et al., 2006), including the aliphatic amidase *amiE* and *amiF* (Pflock et al., 2006) and the *ure* urease operon (Pflock et al., 2004, 2005)). Transcription of this set of genes is controlled by the housekeeping  $\sigma^{80}$  factor and regulated by the essential response regulator ArsR (Acid responsive signaling). ArsR together with the ArsS sensor kinase compose one of the three two-component systems of *H. pylori* (Beier and Frank, 2000). Interestingly, the nickel response regulator NikR was reported to control the expression of a subset of acid-responsive genes (Contreras et al., 2003). Now, considering that at low pH, the solubility of metal ions such as nickel increases, it has been proposed that acidic pH induces NikR expression by an increased bioavailability of nickel ions (van Vliet et al., 2002). Further, the ArsRS system is also regulated by the iron-responsive regulator Fur (ferric uptake regulator), which enters in competition with NikR to bind to the promoter region of *asrR* (Roncarati et al., 2016), connecting the acid acclimation with the metal ion homeostasis responses. Transcriptional regulation of stress and motility responses is mediated by the two transcriptional repressor proteins HrcA and HspR, which repress the expression of the *groES-groEL* and *hrcA-grpE-dnaK* operons (Roncarati et al., 2007). Flagellar and chemotaxis responses are more challenging to characterize, as often the genes involved are unclustered and scattered within multicistronic operons (Alm et al., 1999; Tomb et al., 1997). Furthermore, they are regulated by at least 7 different TFs, including the housekeeping  $\sigma^{80}$  factor (*rpoD*),  $\sigma^{54}$  (*rpoN*),  $\sigma^{28}$  (*fliA*), a NtrC-like response regulator FlgR, FlgS and CheA/CheY/CheY2 regulating chemotactic responses (Foynes et al., 2000; Jimenez-Pearson et al., 2005).



Interestingly, the *H. pylori* genome also encodes for two, poorly characterized but essential, response regulators that lack the histidine kinase component (so-called orphan response regulators), the HP1021 and HP1043. The HP1021 orphan response regulator has been shown to regulate the transcription of a gene cluster involved in the acetone metabolism (Pflock et al., 2007). The HP1043 (as the AsrR) is a member of the OmpR/PhoB subfamily of orphan response regulators, which is uniquely found in Epsilonproteobacteria and lacks the phosphorylation domains. It was initially described as a cell cycle regulator, controlling the expression of the methyl-accepting chemotaxis *tlpB* gene in response to growth phase (Delany et al., 2002). Strikingly, in 2016 it was related to the growth-phase dependent transcription of CncR1 by binding to the PcagP promoter (Vannini et al., 2016). CncR1 is one of the only two *trans*-encoded small RNA (sRNA) so far characterized in *H. pylori*, and will be further discussed in the ‘Small regulatory RNAs in *H. pylori*’ subsection of this Chapter. Another recent study identified 37 new genes harboring a conserved HP1043 nucleotide sequence motif in their promoter or intragenic regions by using Chromatin Immunoprecipitation-sequencing (ChIP-seq) (Pellicciari et al., 2017). Intriguingly, most of them correspond to genes involved in translation, thus a potential role of HP1043 provoking growth arrest by arresting protein synthesis have been described (Pellicciari et al., 2017). But additionally, many of the newly identified targets are sRNAs, including RepG, which regulates the chemotaxis receptor TlpB (Pernitzsch et al., 2014), IsoA1, which acts as antitoxin inhibiting the translation of the AapA1 toxin (Arnion et al., 2017, see appendix 8.3 in Chapter 8), and IsoA3, antitoxin of the same family as IsoA1, impeding the expression of the AapA3 toxin, which will be the focus of this thesis. All of these sRNAs will be extensively described in the ‘Small regulatory RNAs in *H. pylori*’ subsection of this Chapter. Additionally, the potential role of HP1043 in IsoA3 regulation will be further discussed in Chapter 5 section 5.2.1.

### 1.3. Post-transcriptional regulation in *H. pylori*

Bacteria are extremely adaptable organisms, able to efficiently optimize biological processes in response to changes in the environment. Once a gene is transcribed, its expression is strongly determined by a myriad of events that occur in the time-window between transcription termination and translation initiation, the so-called post-transcriptional regulation events. Which, when and how these events act on the mRNA is fate changing. These events mainly include mRNA stability and translatability, although additional factors such as mRNA cellular location can also play a role.

#### 1.3.1. mRNA stability

In almost all domains of life, information-carrying messengers are constantly actively destroyed. This can be seen as a waste of energy and resources; however, it is in the controlled destruction of RNAs that fast and accurate responses to environmental (and internal) stimuli can occur. Additionally, by rapidly destroying not useful transcripts, the costs of translation are reduced and ribonucleotide recycling is beneficial. Thus, it is an evolutionary requirement for any organism to develop mechanisms for RNA degradation. Bacteria have developed two main types of ribonucleases: 1) endonucleases and 2) exonucleases. The number and composition of such battery can strongly vary between clades.

#### *Endoribonucleases*

Considering all bacteria species, there are three major ribonucleases with endonuclease activity, including RNase E (and its homolog RNase G), RNase Y and RNase III. Additional minor endonucleases with specific functions, such as the RNase P, which matures the tRNA 5' ends and noncoding regions of some mRNAs (Li and Altman, 2003), and the RNase Z, which eliminates aberrant 3' ends on tRNAs (Dutta and Deutscher, 2009) as well as of some mRNAs (Perwez and Kushner,

2006), can also be found. RNase E is considered one of the most important activities governing mRNA decay. It is a homotetramer carrying the catalytic domain in the highly-conserved amino-terminal domain (McDowall and Cohen, 1996), a membrane-binding helix in the poorly-conserved carboxy-terminal domain (Khemici et al., 2008), two arginine-rich RNA-binding domains, and a region that serves as a scaffold for the assembly of a ribonucleolytic complex known as the RNA degradosome (Carpousis et al., 1994; Causton et al., 1994; Ehretsmann et al., 1992; Py et al., 1994; Vanzo et al., 1998), composed by the RNase E, the polynucleotide phosphorylase (PNPase), the RNA helicase B (RhlB) and the enolase (Miczak et al., 1996; Py et al., 1996). RNase E preferentially cuts AU-rich single-stranded RNA internally (McDowall et al., 1994). Strikingly, despite being able to act far away from the 5' end, a strong preference for 5' monophosphorylated ends has been reported (Mackie, 1998). This phenomenon is explained by the presence of a 5' end binding pocket in the catalytic domain that is only able to accommodate 5' monophosphate but not di- or triphosphate, serving therefore as a 5' end state sensor (Callaghan et al., 2005). Several studies have demonstrated the crucial role of the RNase E in both, mRNA (Khemici and Carpousis, 2004; Kushner, 2002; Lopez et al., 1999) and small RNAs (in a Hfq-dependent manner) (Masse et al., 2003; Morita et al., 2005) degradation, however, *H. pylori* lacks both, RNase E and Hfq (as 50% of all bacteria species). These characteristics make of *H. pylori* an appealing model for the study of novel post-transcriptional regulation mechanisms.

Despite lacking RNase E, *H. pylori* possess RNase Y and RNase III activities. RNase Y has been described as the functional homolog of RNase E in bacteria lacking this activity. That of *Bacillus subtilis* possesses a transmembrane domain, a disordered coil-coil-domain, an RNA-binding KH domain, and a catalytic HD domain (Lehnik-Habrink et al., 2011). Several studies have pointed at RNase Y as a major regulator of mRNA decay (Laalami et al., 2013) and virulence in the pathogens *Streptococcus pyogenes* (Chen et al., 2013) and *Staphylococcus aureus* (Marincola et al., 2012). Unlike RNase E and RNase Y, RNase III is a double-strand-specific ribonuclease. It functions as a dimer with an endonucleolytic domain composed by two central catalytic sites that act independently (yielding to the characteristic 2-nt overhang at the 3' end (Gan et al., 2006; Meng and Nicholson, 2008)), and double-stranded RNA-binding domain (Blaszczyk et al., 2004). Importantly, at least two turns of an RNA helix in length (~20 bp) are required for its activity *in vivo* (Robertson, 1982), thus being most natural stem-loop structures too short to be targets. Nevertheless, RNase III plays a role in ribosomal RNA maturation as well as mRNAs (e.g., autoregulation of the *rnc-era-recO* operon (Matsunaga et al., 1996) and CRISPR RNAs (Deltcheva et al., 2011)). Interestingly for us, it has been shown that RNase III is essential for the degradation of toxin-encoding mRNAs belonging to type I toxin-antitoxin (TA) systems in *B. subtilis* (Durand et al., 2012). However, we have shown that this is not the case in *H. pylori* (Arnion et al., 2017, see appendix 8.3 in Chapter 8) (see Paper III for further discussion).

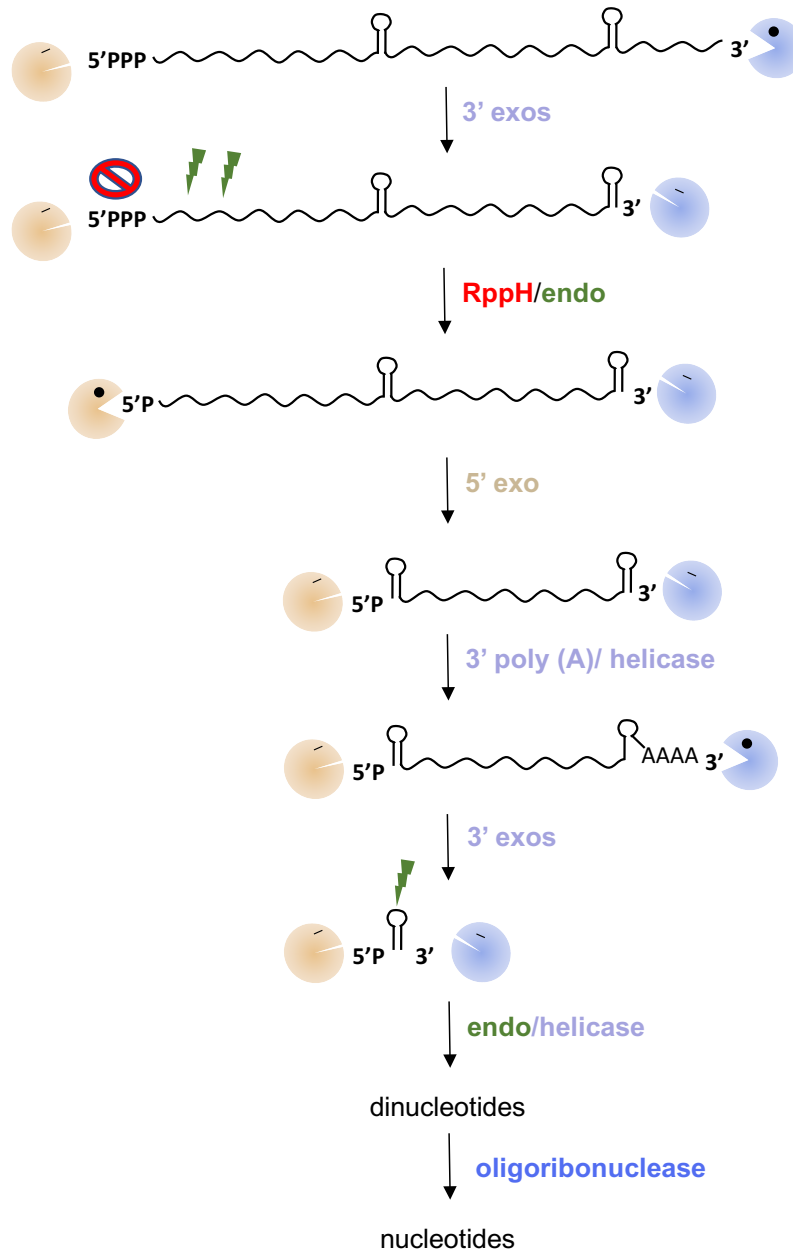
### Exoribonucleases

As degradation pathways are often initiated by endonucleolytic cleavages, exoribonucleases are required to quickly degrade decay intermediates that lack termini protection. There are three main types of exoribonuclease activity: 1) 3' phosphohydrolytic (reversible, uses orthophosphate as nucleophile to produce nucleoside diphosphates), 2) 3' hydrolytic (irreversible, yielding to monophosphate products), and 3) 5' hydrolytic (same mechanism as the 3' hydrolytic but on the 5' end of the target mRNA). The two major examples of 3' phosphohydrolytic exonucleases are the RNase PH, involved in tRNA maturation in *E. coli* (Deutscher et al., 1988), and the polynucleotide phosphorylase PNPase, involved in mRNA decay of RNAs with a single-stranded 3' end (Chen et al., 1991); both of them being members of the PDX family of enzymes (Zuo and Deutscher, 2001). *H. pylori* only possesses PNPase activity. Interestingly, its polymerase activity can add a single-stranded adenine-rich 3' end tail that facilitates its 3' hydrolytic activity (Xu and Cohen, 1995). Importantly, because the catalytically active



sites are located inside of a central channel (Narczyk et al., 2018; Štefanić et al., 2017; Symmons et al., 2000), the 3' end of the RNA must thread partway through the channel to reach them. As a result of this architecture, PNPase degrades the 3' end of RNA processively until encountering a base-paired structure with sufficient thermodynamic stability (Chen et al., 1991). This phenomenon turned out to be of extreme importance for the co- and post-transcriptional regulation of the *aapA3*/IsoA3 type I TA system here studied, and will be extensively discussed in Chapters 6 and 7. In *E. coli*, PNPase can be assisted by the unwinding activity of the RNA helicase RhlB (Py et al., 1996) and lead to the release of a 5' terminal dinucleotide (generally exonucleases cannot achieve RNA degradation to the mononucleotide level). Despite the absence of a RhlB homolog in *H. pylori*, another RNA helicase, the RhpA, has been associated to be part of a so-called minimal RNA degradosome together with the 5' exoribonuclease RNase J (El Mortaji et al., 2018; Redko et al., 2013). The role of RhpA in the *aapA3*/IsoA3 mRNA processing will also be studied on this thesis in the Chapter 4 section 4.3 (Paper IV).

The two majors 3' hydrolytic exonucleases are members of the RNR superfamily, the RNase II and RNase R. Despite sharing structural similarities (Cheng and Deutscher, 2002), RNase R has the intrinsic capacity of RNA unwinding, which allows its action with no need of external RNA helicase activity (Awano et al., 2010). *H. pylori* only encodes for the RNase R gene, which has been proposed to regulate the expression of motility and apoptosis-inducing genes (Tsao et al., 2009). Additionally, *H. pylori* have a 5' exonuclease activity (absent in *E. coli*), the RNase J, which is essential and which depletion has been shown to lead to massive changes in mRNA abundance, suggesting its involvement in the decay pathways of a large number of mRNAs (Redko et al., 2016). As above-mentioned, a minimal RNA degradosome composed of RNase J and the RNA helicase RhpA, has been proposed to exist in *H. pylori* (despite the lack of RNase E) (Redko et al., 2013). In this thesis, the possible cooperative activity of both proteins will be further studied in Chapters 6 and 7. Further, as in the case of the RNase E, RNase J has been shown to have a preference for 5' monophosphate single-stranded ends (Mathy et al., 2007; Richards et al., 2011). Indeed, the RNase J is a dimer of dimers in which each subunit contains a bipartite metallo- $\beta$ -lactamase domain, a  $\beta$ -CASP domain, and a carboxy-terminal domain. An RNA-binding channel leads deep inside the protein to a catalytic site at each dimer interface, where a monophosphorylated but not a triphosphorylated 5' end can be accommodated (Dorléans et al., 2011; Mathy et al., 2007). Because this channel continues past the catalytic center and emerges on the other side, explains the ability of RNase J to also act as an endonuclease (Even et al., 2005). This endonuclease activity turned out to be crucial for the degradation of the IsoA3 sRNA antitoxin and will be studied in the Chapter 4 section 4.3 (Paper IV). *H. pylori* has a Nudix hydrolase, the RppH, which is able to convert the 5' terminal triphosphate of bacterial transcripts to a monophosphate, triggering RNA decay by deprotection of the 5' end. In a recent study, its role in mRNA decay was examined by RNA-seq, and over 63 potential targets were identified, suggesting an important role of RppH in the post-transcriptional regulation of *H. pylori* (Bischler et al., 2017). In this thesis, the role of RppH is studied in the context of the IsoA3 sRNA decay mediated by RNase J (see Chapter 4 section 4.3 (Paper IV)).



**Figure 1.3. Possible pathways of mRNA decay in *H. pylori*.**

Figure 1.3 summarizes possible pathways of mRNA decay in *H. pylori*. In primary transcripts, the 5' exonuclease activity of RNase J is impeded by the 5' triphosphate state. The 3' exonuclease PNPase (together with other 3' exonuclease activities) can process the mRNA until reaching a stem-loop with high thermodynamic stability that serves as an energetic roadblock. At this point, the RppH can convert the 5' triphosphate to monophosphate, triggering the activity of the 5' exonuclease RNase J. Alternatively, the same results can be achieved by an endonuclease activity. Similarly to the PNPase, the RNase J gets blocked by highly stable secondary structures. At this point, PNPase can use its polymerase activity to add a poly-A tail that will facilitate its 3' exonucleolytic activity. Alternatively, a helicase could help any of the exonucleases to unwind the highly stable secondary structures and trigger their activity. When very short fragments are left, the oligoribonuclease act yielding to mononucleotides products.

### 1.3.2. Translational regulation

Translation has a great and immediate impact in the definition of the proteome (Picard et al., 2012), which is the ultimate element of adaption responses. Interestingly, several studies have highlighted a poor correlation between mRNA and protein levels. Indeed, mRNA concentration only explains about 50% of the protein concentration in *E. coli* (Corbin et al., 2003; Lu et al., 2007) and even less in other bacteria (e.g., *Lactococcus lactis* (Dressaire et al., 2010)), and protein stability playing a minor role (Maier et al., 2014). Therefore, the regulation of translation plays a major role in the generation of fast and optimized adaptation responses. Although several factors influence both, translation initiation and elongation rates, most regulatory mechanisms target the initiation step, making it the rate-limiting step of protein synthesis in bacteria (Gold, 1988; Kennell and Riezman, 1977). Canonical translation initiation involves the base-pairing interaction between the Shine-Dalgarno (SD) sequence in the mRNA with the anti-SD sequence located at the 3' end of the 16S rRNA embedded in the 30S ribosomal subunit (Shine and Dalgarno, 1974; Steitz and Jakes, 1975). The 30S ribosomal subunit, together with the Initiation Factor 2 (IF-2), that kinetically assists the specific interaction between an anticodon present in the fMet-tRNA<sup>fMet</sup> and an optimally-spaced (Chen et al., 1994) start codon (AUG or UUG/GUG) in the mRNA, form the so-called translation initiation complex (Hartz et al., 1989; Simonetti et al., 2013; La Teana et al., 1993).

#### *Consequences of transcription/translation coupling*

Due to the lack of nuclear compartmentalization, transcription and translation (and degradation) processes can be coupled in time and space in bacteria (Burmann et al., 2010; Das et al., 1967; Proshkin et al., 2010). Thus, the physical interaction between the major components of all the three processes is a key determinant of gene expression in bacteria. Interestingly, two studies used time-resolved single cell imaging in *E. coli* to reveal a possible partial compartmentalization, on which ribosomal subunits can co-transcriptionally initiate translation in the nucleoid while at the periphery and membrane proximity polysomes are formed (Bakshi et al., 2012; Sanamrad et al., 2014). The universally conserved TF NusG, as well as the ribosomal protein S10, have been proposed to be involved in the interaction between the RNA polymerase and the 30S ribosomal subunit (Burmann et al., 2010; McGary and Nudler, 2013). Indeed, co-transcriptional translation (especially the pioneer ribosome) is thought to be beneficial as it avoids spontaneous backtracking of the RNA polymerase, aiding to read-through roadblocks (Herbert et al., 2006; McGary and Nudler, 2013; Proshkin et al., 2010). One additional implication of transcription/translation coupling is that it may make possible to transcribe at different speeds throughout the mRNA, being that of the 5' UTR and start codon slower (Al-Hashimi and Walter, 2008). This is interesting as the mRNA folding dynamics have been shown to be influenced by the elongation speed of the RNA polymerase (Wickiser et al., 2005). Indeed, several studies have highlighted the role of the 5' UTR structure in translation (Simonetti et al., 2009), and more specifically, the key role of the SD sequence accessibility for translation initiation (Bhattacharyya et al., 2018; Espah Borujeni et al., 2014; Kozak, 2005), which has even been proposed to dictate N-terminal codon bias in *E. coli* (Bhattacharyya et al., 2018). In this thesis, the role of the SD sequence accessibility in the AapA3 toxin mRNA has been studied (see Chapter 4 section 4.2 (Paper III) for results, and Chapter 5 for further discussion) and was revealed to be an essential feature allowing the hostage of TA systems in bacterial genomes (for a review see (Masachis and Darfeuille, 2018) in the section 1.4.1 of this Chapter).

### **Regulatory proteins**

A large number of RNA-binding proteins (RBPs) also positively or negatively regulate translation initiation. For instance, binding of the S15 ribosomal protein (r-protein) to the *rpsO* mRNA (coding for the r-protein S15) results in SD inaccessibility in *Thermus thermophilus* (Serganov et al., 2003). Similarly, the carbon storage coordinator CsrA, inhibits translation initiation of a large set of genes (e.g., involved in carbon metabolism, quorum sensing and virulence) by binding to GGA motifs and competing with 30S ribosomal subunits to bind to the SD sequence (Duss et al., 2014; Fields et al., 2016). However, translation regulation mediated by RBPs will not be the focus of this thesis.

### **Regulation by sRNAs**

There is a large myriad of small regulatory RNAs (sRNAs) that act at the translational level to regulate the expression of their target mRNAs. In the next section, the biogenesis, mode of action and role within biological networks of sRNAs in bacteria will be discussed.

#### **1.3.3. Small regulatory RNA in bacteria**

Small regulatory RNAs are typically 50-200 nt length and do not code for proteins (non-coding RNAs), with the exception of the so-called dual-function sRNAs that also code for small proteins (i.e., the RNAIII (Balaban and Novick, 1995) and Psm-mec RNA (Kaito et al., 2011) in *S. aureus*, SRI in *B. subtilis* (Gimpel et al., 2010), SgrS in *E. coli* (Wadler and Vanderpool, 2007), and PeIRNA in *Streptococcus* (Mangold et al., 2004), for a review see (Gimpel and Brantl, 2017)). The role of RNA as regulators was discovered long ago (e.g., RNA I, inhibitor of ColE1 plasmid replication (Stougaard et al., 1981; Tomizawa et al., 1981), a small ~70 nt RNA inhibiting the translation of Tn10 transposase (Simons and Kleckner, 1983), and MicF, the first chromosomally-encoded sRNA, which inhibits the translation of the outer membrane protein OmpF (Mizuno et al., 1984)). However, their prevalence and biological significance initially received poor attention. Some years later, with the increasing availability of sequenced bacterial genomes, bioinformatic searches, together with microarray analyses, identified several new sRNAs in diverse bacterial species (Livny and Waldor, 2007; Raghavan et al., 2011; Tsui et al., 2010; Weinberg et al., 2007), including *H. pylori* (Sharma et al., 2010). Further studies, demonstrated that sRNAs not only are widely present in bacterial genomes but also play crucial roles in the physiology and adaptive responses, such as quorum-sensing (for a review see (Svenningsen, 2018)), carbon metabolism (Papenfort and Vogel, 2014) and even in bacteria-phage communication (for a review see (Altuvia et al., 2018)). However, due to space limitations this topic will not be extensively discussed on this thesis.

##### **1.3.3.1. Biogenesis**

Although most sRNAs are known to be transcribed as independent genes, there is an increasing number of sRNAs reported to be derived from the 5' and 3' UTR sequences of mRNAs (Kawano et al., 2005). A recent study, described the existence of a prominent 3' UTR-derived sRNA biogenesis pathway on which the RNA chaperone Hfq guides RNase E activity on sRNA-precursors to generate the sRNAs able to interact with their target mRNAs (Chao et al., 2017). This pathway was previously observed for the *S. enterica* CpxQ sRNA (Chao and Vogel, 2016). More recently, a study of the MicL sRNA, which despite being transcribed from an internal promoter, it needs to be cleaved to be functional (Guo et al., 2014), has reported on the existence of a similar pathway and interestingly proposed to be directed by stem-loop structures that modulate RNase E activity (Updegrove et al., 2018).

### 1.3.3.2. Modes of action

Few sRNAs act by binding to their protein target (*e.g.*, the CsrB and CsrC sRNAs have multiple GGA motifs and thereby reduce the amount of CsrA protein able to interact with its target mRNAs (Weilbacher et al., 2003), for a review see (Storz et al., 2011)), however, the vast majority of sRNAs so far characterized, regulate gene expression by base-pairing with their target RNA. Base-pairing sRNAs can be divided into two groups: 1) *cis*-encoded sRNAs, and 2) *trans*-encoded sRNAs.

#### *Cis-encoded sRNAs*

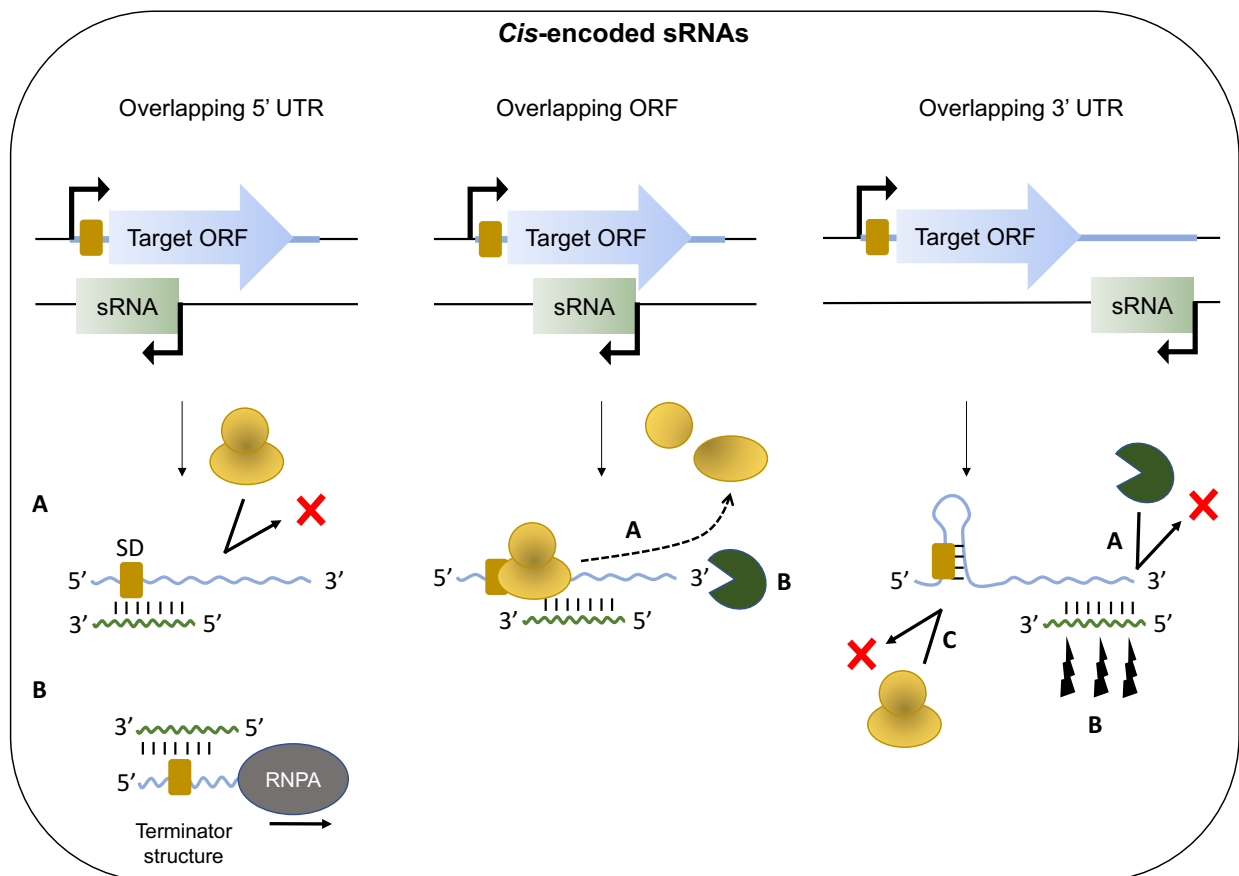
Cis-encoded sRNAs (or antisense RNAs, asRNAs) are encoded on the opposite DNA strand of their target RNA and therefore share extensive regions of complementarity (75 nucleotides or more) with their targets (Wagner et al., 2002). However, despite being encoded in the same DNA region, they are independently transcribed as discrete RNA species and function in *trans* as diffusible molecules. As previously mentioned, the first asRNA discovered was RNAI in the ColE1 plasmid acting as an inhibitor of primer formation (Stougaard et al., 1981; Tomizawa et al., 1981), followed by the identification of a sRNA inhibiting Tn10 transposase activity (Simons and Kleckner, 1983). Soon after, one more sRNA, the Sok sRNA was discovered as being essential for the stability of the R1 in *E. coli* (*i.e.*, acting as a Post-Segregational Killing (PSK) elements) (Gerdes et al., 1986; Greenfield and Weaver, 2000). Interestingly, subsequent studies revealed that the RNAI of the ColE1 plasmid is in fact the toxin component of a type I toxin-antitoxin (TA) system (the RNAI (re-named *fst*)/RNAII, (Franch et al., 1999)). Furthermore, the Sok sRNA was identified as the antitoxin component of the *hok*/Sok type I TA system in *E. coli* (Thisted et al., 1994). Type I TA systems are part of the main focus of this thesis and will, therefore, be extensively discussed throughout this work.

As shown in Figure 1.4, *cis*-encoded sRNAs can be organized in three main groups according to their operon organization: 1) overlapping the 5' UTR of the mRNA, 2) overlapping the Open Reading Frame (ORF), and 3) overlapping the 3' UTR. An example of 5' UTR-overlapping sRNA (Figure 1.4, left panel) is the antitoxin SymR, which inhibits the translation of the SOS-response toxin SymE in *E. coli* by occluding its Ribosome Binding Site (RBS) (defined as the region within the mRNA that is protected by the 30S ribosomal subunit in the initiation complex) (Figure 1.4, left panel A) (Kawano et al., 2007). 5' UTR overlapping sRNAs can also function by promoting transcription termination through the formation of stable RNA heteroduplexes that mimic terminator structures, or by promoting the formation of terminator hairpin structures (Figure 1.4, left panel B). An example of this mechanism can be found in the case of the *icsA* mRNA of *Shigella flexneri*, whose 5' UTR anti-terminator structure is abolished by the sRNA RnaG (Giangrossi et al., 2010). Interestingly, it has been recently shown the transcriptional attenuation caused by the RnaG sRNA can be alleviated by the master activator of virulence genes VirF through the inhibition of RnaG/*icsA* interaction (Giangrossi et al., 2017). Alternatively, sRNA binding to the 5' UTR can disrupt an existing anti-terminator structure, leading to a positive regulation of the mRNA expression.

Other asRNAs are encoded overlapping the coding region of mRNAs (Figure 1.4, central panel). In these cases, sRNA binding to the mRNA can stop translation elongation and lead to ribosome recycling (Figure 1.4, central panel A) or/and cause mRNA destabilization (Figure 1.4, central panel B). An example of this latter mechanism is observed in the regulation of the *IsiA* photosynthesis gene expression by the *IsrR* sRNA in the cyanobacteria *Synechocystis* (Duhring et al., 2006). Last, some sRNAs overlap the 3' UTR of mRNAs (Figure 1.4, right panel). In these cases, positive regulation can be observed. Indeed, the binding of the sRNA to the 3' UTR of the mRNA can lead to mRNA stabilization, by, for instance, avoiding exoribonuclease activity (Figure 1.4, right panel, A). An

example of this can be found in the regulation of the Fst toxin by the RNAII sRNA (Greenfield et al., 2000). Indeed, RNAII binding to the Fst mRNA leads to translation inhibition due to the sequestration of the translation start codon, but at the same time, it promotes Fst mRNA stabilization (reviewed in Brantl and Jahn, 2015). Another mRNA stabilization example is found in the regulation of the *gadXW* mRNA (operon) by the stationary phase-induced sRNA GadY. It has been shown that GadY base pairing to the *gadXW* mRNA leads to the cleavage of the duplex between the *gadX* and *gadW* genes by RNase III and GadX mRNA stabilization (Opdyke et al., 2004; Tramonti et al., 2008). Alternatively, sRNA binding, if leading to the formation of an extended RNA duplex, could target the mRNA for degradation by the double-stranded-specific ribonuclease RNase III (Figure 1.4, right panel, B).

Further, sRNA binding to the 3' UTR can also repress mRNA expression, by, for instance, inhibiting translation by promoting the formation of stable RBS sequester structures that do not exist in the original mRNA folding (Figure 1.4, right panel C). An example of this regulatory mechanism is found in the case of the Fst mRNA regulation by the RNAII sRNA (as above-mentioned), as well as in the regulation of the BsrG toxin mRNA by the SR4 antitoxin sRNA. Indeed, SR4 does not directly interact with the RBS, but rather generates an extended intramolecular SD sequestration (4 bp in the free BsrG mRNA is extended to 8 bp upon SR4 binding) (Jahn and Brantl, 2013; Jahn et al., 2012). Alternatively, sRNA binding to the 3' UTR can lead to the disruption of an SD-occluding structure, and thus, to the de-repression of the mRNA expression. Further discussion in the regulation of toxin mRNAs by antitoxin sRNAs can be found in Masachis and Darfeuille, (2018) in the section 1.4.1 of this Chapter.



**Figure 1.4. Mechanisms of action of cis-encoded sRNAs.** Cis-encoded sRNAs can overlap the 5'UTR of the mRNA (left panel), the Open Reading Frame (ORF, central panel), or the 3' UTR of the mRNA (right panel). 5'UTR overlapping sRNAs can regulate gene expression by occluding the Shine-Dalgarno (SD) or



Ribosome Binding Site (RBS) in general (**left panel, A**) or by inducing the formation of terminator structures that lead to the premature transcription termination of the mRNA and thus to transcription attenuation (**left panel, B**). Disruption of an anti-terminator structure can also be possible, leading to a positive regulation (not shown). ORF-overlapping sRNAs can impede translation elongation, leading to ribosome recycling (**central panel, A**), and/or promote mRNA degradation by cellular ribonucleases (**central panel, B**). 3'UTR overlapping sRNAs can have a positive regulatory effect by increasing mRNA stability through the protection from exonuclease degradation (impeded by stable double-stranded structures) (**right panel, A**). However, they can also repress mRNA expression by increasing mRNA decay mediated by endonucleases (**right panel, B**) and/or inhibit translation by prompting the formation of SD sequester structures that are not present in the absence of sRNA binding (**right panel, C**). Alternatively, sRNA to the 3'UTR can also disrupt an SD-occluding structure, leading to a de-repression of the mRNA translation.

---

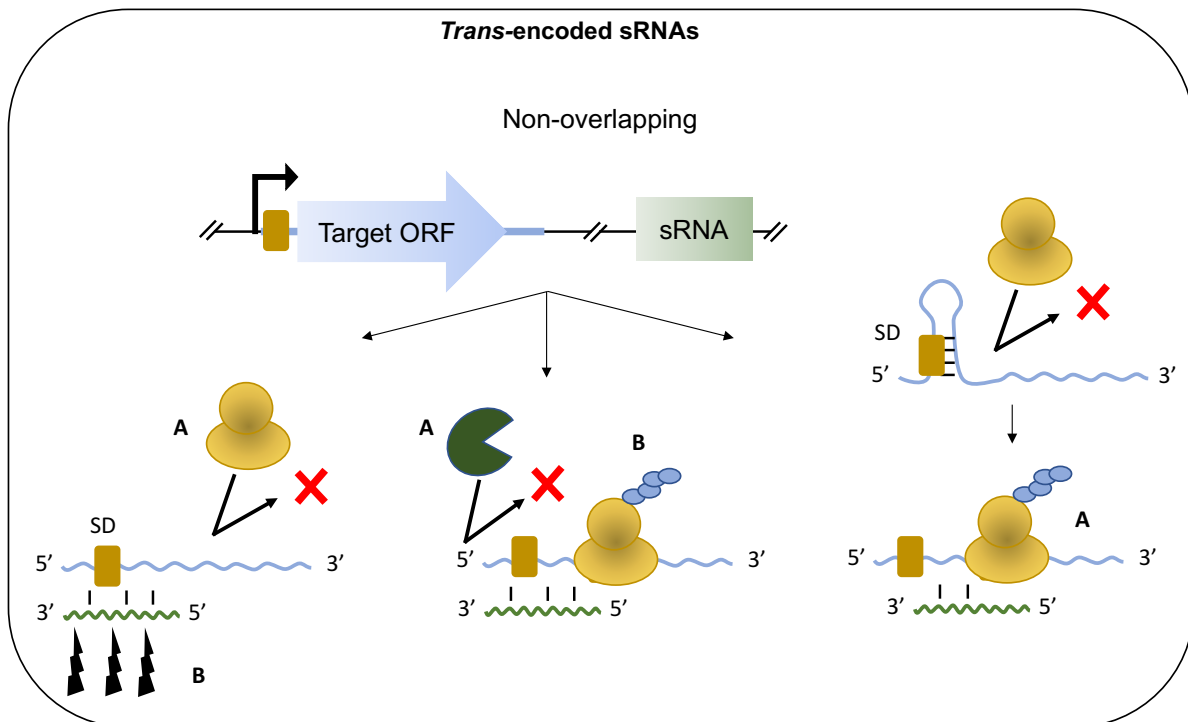
### *Trans-encoded sRNAs*

Contrary to the *cis*-encoded sRNAs, *trans*-encoded sRNAs are encoded elsewhere in the genome and share limited complementarity with their target mRNA, interacting by short and imperfect base-pairing. Nevertheless, this gives them interaction flexibility, and consequently, *trans*-encoded sRNAs can typically target multiple mRNAs and are often found as global modulators of physiological responses (review in Masse et al., 2005; Massé et al., 2007; Valentin-Hansen et al., 2007). A major consequence of the limited contact (~10-25 nucleotides, in discontinuous patches (e.g., the only 4 nt required for interaction of the SgrS sRNA with the PtsG mRNA out of the possible 23 bp, Kawamoto et al., 2006)) between *trans*-encoded sRNA and their targets is that, in most cases, the RNA chaperone Hfq is required to facilitate the interaction. This is true for all characterized *trans*-encoded sRNA in *E. coli*, however, there are some exceptions, such as the repression of OmpA expression by the VrrA sRNA in *Vibrio cholerae*, which do not require Hfq activity thanks to the presence of an exceptionally long interaction region (Song et al., 2008). Interestingly, Hfq has been proposed to act as a platform for RNA binding, increasing the local concentrations of sRNAs and mRNAs (diminishing the impact of molecule diffusion) (Aiba, 2007), and has been exploited as a tool for the identification of sRNAs in *E. coli* (Zhang et al., 2003) and *S. typhimurium* (Chao et al., 2012).

Most of the *trans*-encoded sRNA regulation characterized so far is negative, and mainly achieved by the direct occlusion of the RBS on the target mRNA (Figure 1.5, left panel A) (Gottesman, 2005). However, sRNAs can also target far up- or downstream regions from the canonical RBS and start codon of their target mRNA. For instance, the GcvB sRNA in *Salmonella* is able to inhibit translation of the *gltI* mRNA by binding to a C/A-rich translation enhancer element located at -57 to -45 upstream the translation start codon (Sharma et al., 2007). On the other side, sRNA binding to a region other than the RBS can also lead to mRNA degradation, directly (i.e., RNA duplex degradation by RNase III), or indirectly (i.e., caused by the lack of translating ribosomes (Deana and Belasco, 2005; Iost and Dreyfus, 1995)) (Figure 1.5, left panel B). This is the case for the MicC sRNA in *S. typhimurium*, which targets the coding sequence of the *ompD* mRNA (via  $\leq 12$  bp with the 23-26 codons) and triggers *ompD* mRNA degradation by RNase E (Pfeiffer et al., 2009). In this case, the mRNA destabilization effect is indirect, due to the lack of translating ribosomes. Other sRNAs, such as SgrS and RyhB, promote degradation of their target mRNAs by exposing otherwise hidden RNase E cleavage sites (Morita et al., 2005).

Interestingly, some sRNA can have a positive effect in the expression of their target genes, by, for instance, increasing the stability and/or translation of their target mRNAs (Figure 1.5 central and right panels). An example of sRNA increasing the stability of its target mRNA is the FasX sRNA of *Streptococcus*. The FasX sRNA basepairs with the 5' end of the secreted virulence factor streptokinase *ska* mRNA and thereby prevent its degradation, which ultimately results in an increased overall translation of the Ska protein (Liu et al., 2012). Similarly, in *Salmonella*, the RydC sRNA selectively

activates the longer of two isoforms of the *cfa* mRNA (coding for a cyclopropane fatty acid synthase) by binding to its 5' end and directly interfering with RNase E-mediated decay (Fröhlich et al., 2013) (Figure 1.5, central panel A and B). Additionally, and similarly to some *cis*-encoded sRNAs, a set of *trans*-encoded sRNAs activate the expression of their target mRNAs by disrupting SD-occluding structures, directly (*i.e.*, binding to the RBS region), or indirectly (*i.e.*, binding up-or downstream the RBS region but inducing a conformational change on it) (Figure 1.5, right panel A). Importantly, despite most sRNAs act through a predominant regulatory mechanism, different target mRNAs may be regulated through different pathways; *i.e.*, having positive or negative effects, and mediated by any of the above-mentioned mechanisms (*i.e.*, the Qrr 1-4 sRNAs, central regulators of quorum-sensing pathways (Feng et al., 2015; Papenfort and Bassler, 2016; Shao et al., 2013)).



**Figure 1.5. Mechanisms of action of *trans*-encoded sRNAs.** *Trans*-encoded sRNAs do not overlap their target gene. They can negatively regulate gene expression by binding to ribosome binding site region or the Shine-Dalgarno (SD) sequence, inhibiting translation (**left panel A**), or by promoting mRNA degradation, directly (*i.e.*, exposing otherwise hidden cleavage sites) or indirectly (*i.e.*, through the deprotection of the mRNA due to the lack of translating ribosomes) (**left panel B**). Alternatively, they can positively regulate their targets, by impeding mRNA degradation (*e.g.*, blocking the priming of exonucleases) (**central panel A**), which can ultimately lead to an overall increased mRNA translation (**central panel B**), or by disrupting an SD-occluding structure present in the free-state mRNA, which will also result in increased levels of mRNA translation (**right panel A**).



### 1.3.3.3. Small regulatory RNAs in *H. pylori*

During a long period of time, due to its small genome size, the apparent lack of the RNA chaperone Hfq, and the conservation of very few transcriptional regulators and housekeeping RNAs (*i.e.*, only the SRP RNA, the M1 RNA, and tmRNA are conserved (Sharma et al., 2010)), *H. pylori* was thought to lack riboregulation. Only recently, with the development of a differential RNA sequencing approach (sRNA-seq) that specifically selects for the 5' end of primary transcripts, a global map of transcriptional start sites and operons was defined and revealed a complex yet compact *H. pylori* transcriptome structure (Sharma et al., 2010). Importantly, this study also uncovered the existence of massive antisense transcription as well as an unanticipated high number of sRNAs (< 60). Such sRNAs represent potential regulators of *cis*- and *trans*-encoded mRNA targets and suggest the existence of a rich and unexplored transcriptional riboregulation in *H. pylori* (for a review see, Pernitzsch and Sharma, 2012).

To date, due to their relatively recent discovery, very few regulatory sRNAs (two *trans*- and two *cis*-encoded sRNAs) have been characterized in *H. pylori*. The first *trans*-encoded sRNA to be characterized was the RepG sRNA (HPnc5490), which can post-transcriptionally mediate both, the activation and the repression, of the TlpB chemotaxis receptor through a novel mechanism involving the interaction with a variable homopolymeric G-repeat (hypermutable simple sequence repeat, SSR) in the TlpB mRNA (Pernitzsch et al., 2014). Soon after, a second *trans*-encoded sRNA, named CncRI (HPnc2630), was characterized. The CncRI sRNA is encoded in the Cag-PAI and binds to the *fliK* mRNA (encoding for a flagellar checkpoint protein) to negatively regulate motility and adhesion to host cells. Interestingly, the observed phenotypic effects are indirect and linked to the regulation of  $\sigma(54)$ -dependent genes (Vannini et al., 2016). Among the *cis*-encoded sRNAs, the first one to be characterized was the 5' ureB sRNA, which downregulates ureAB (part of the urease gene cluster) expression by enhancing transcription termination 5' of ureB (Wen et al., 2013). The second *cis*-encoded sRNA, IsoA1, was recently characterized and identified as an RNA antitoxin in a type I TA system (Arnion et al., 2017, see appendix 8.3 in Chapter 8). Indeed, we recently reported on the existence of a family of type I TA systems, namely, aapA/IsoA (standing for antisense associated peptide family A/Inhibitor of small QRF family A) in *H. pylori*, representing the first family of type I TA systems characterized in epsilonproteobacteria. A member of this family will be the focus of this thesis, thus, in the next sections of this Chapter, type I TA systems and their regulation will be further discussed.

## **1.4. Type I Toxin-antitoxin systems**

### **1.4.1. Paper I: Type I toxin-antitoxin systems: regulating toxin expression via Shine-Dalgarno sequence sequestration and small RNA binding**



---

# **Type I Toxin-Antitoxin Systems: Regulating Toxin Expression via Shine-Dalgarno Sequence Sequestration and Small RNA Binding**

SARA MASACHIS<sup>1</sup> and FABIEN DARFEUILLE<sup>1</sup>

<sup>1</sup>ARNA Laboratory, INSERM U1212, CNRS UMR 5320,  
University of Bordeaux, F-33000 Bordeaux, France

**<< Diffusion non-autorisée / unauthorized copying or distribution >>**

### 1.4.2. Biological role of type I TA systems

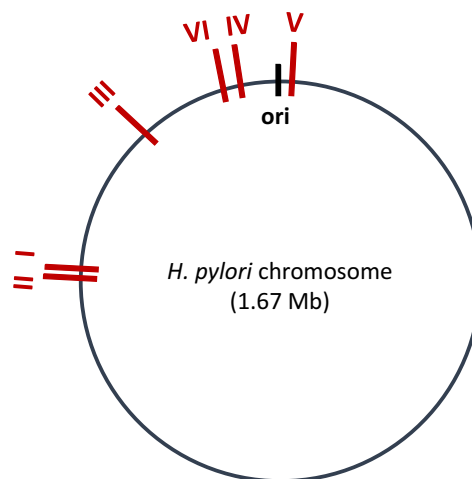
The biological role of TA systems is strongly linked to the mode of toxin expression (*e.g.*, constitutive or inducible) as well as to the molecular effects of it (*e.g.*, cellular target; reversible or irreversible effect). For instance, toxins of plasmid-encoded TAs are usually constitutively expressed as highly stable mRNAs. Plasmid loss during cell division rapidly leads to the depletion of the pool of unstable antitoxin RNA, and subsequently to cell death of the daughter cell having lost the TA-encoding plasmid (PSK mechanism). Thus, plasmid-encoded type I TAs are able to act as plasmid stabilizer elements as shown for the *hok/Sok* of the plasmid R1 (Gerdes et al., 1986), the *srnB*/*SrnC*-RNA of the plasmid F and the *pndA*/*PndB*-RNA of the R483 plasmid (Nielsen et al., 1991) in *E. coli*, and for the *fst*/RNAII of the pAD1 plasmid in *E. faecalis* (Greenfield and Weaver, 2000). Many of the TA systems found if plasmids are also found in bacterial chromosomes, however, the function of chromosomally-encoded type I TA systems is enigmatic.

The SymE antitoxin (SOS-inducible *yjiW* gene with similarities with the MazE type II toxin) from *E. coli* has been suggested to recycle damaged mRNAs upon SOS response and to act as a defense mechanism against infection by RNA phages (Kawano, 2012). Some type I TAs in *B. subtilis* (*e.g.*, BsrG, YonT and TxpA) are located on prophages and have been proposed to act as stabilizer elements in a reminiscent way to PSK (Durand et al., 2012). Other toxin-encoding mRNAs, such as BsrG, BsrE, and BsrH possess a putative ResD (ResD/ResE two-component systems involved in global aerobic and anaerobic responses in *B. subtilis*) binding site upstream their promoters, and have been suggested to be induced under oxidative stress conditions, which role could be to cause bacteriostasis to limit oxygen composition (Durand et al., 2012). In a similar way, the toxins TxpA, BsrG, BsrE, BsrH, and YonT have been reported to be induced upon glucose starvation, which could be beneficial if they cause bacteriostasis and reduce glucose consumption (Jahn et al., 2012; Nicolas et al., 2012). Other chromosomally-encoded type I TAs play a role in persister cell formation, by forming pores in the cell membrane that lead to membrane depolarization, ATP depletion, and subsequently, to a decrease of macromolecular synthesis rates and entry into a persister state, as demonstrated for the SOS-inducible *tisb*/IstR (Berghoff et al., 2017; Dörr et al., 2010; Gurnev et al., 2012; Vogel et al., 2004) and the *hok/Sok* (Pedersen and Gerdes, 1999; Verstraeten et al., 2015; Wilmaerts et al., 2018) systems in *E. coli*. However, the role of TAs in persister cell formation is currently highly controversial, as a recent study has highlighted that previous observations of the role of 10 chromosomally-encoded type II TAs in persistence and antibiotic resistance (retracted) was an artifact caused by a  $\Phi$ 80 bacteriophage infection (Goormaghtigh et al., 2018; Harms et al., 2017).

## CHAPTER 2. Aims of this thesis

### 2.1. Research background: a new family of type I TA systems in *H. pylori*

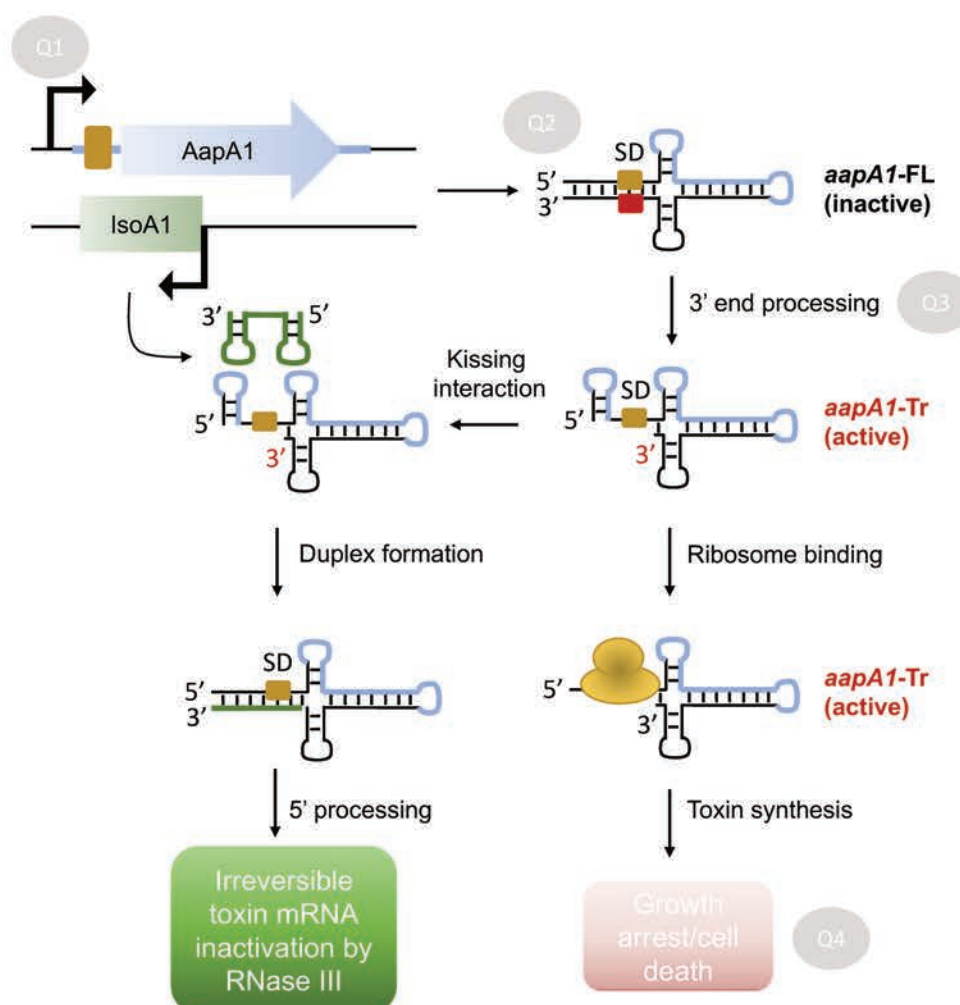
In 2010, a transcriptome-wide screen for sRNAs revealed the existence of massive antisense transcription and >60 new sRNAs in *H. pylori* (Sharma et al., 2010). To date, only four of them are characterized and are presented in the [section 1.3.3.3 of Chapter 1](#). Among the sRNAs with the highest levels of expression, there was a family of six homologous *cis*-encoded antisense RNAs (named IsoA1 to IsoA6, standing for inhibitor of small QRF family A, locus 1 to 6) expressed in the opposite DNA strand to a novel class of small mRNAs (named AapA1 to AapA6, standing for antisense associated peptide family A) (Sharma et al., 2010). Both, the AapA mRNA and the IsoA sRNA are constitutively transcribed, representing small expression modules repeated at six chromosomal loci ([Figure 2.1](#)). Additionally, *in vitro* translation assays showed that each AapA small mRNA codes for a small protein of 30 amino acids, which translation is repressed by the IsoA sRNAs in a loci-specific manner (no cross-regulation). Considering their operon organization, these loci were hypothesized to be chromosomally-encoded type I TA systems. Further characterization confirmed this hypothesis, making the *aapA*/IsoA family the first family of type I TAs identified in epsilon proteobacteria (Arnion et al., 2017; Sharma et al., 2010).



**Figure 2.1. Location of the *aapA*/IsoA modules in the *H. pylori* chromosome.** The *aapA*/IsoA modules are present at six different chromosomal loci in the *H. pylori* chromosome (loci I to VI).

In our recent work (Arnion et al., 2017, see [Appendix 8.3 in Chapter 8](#)), the module present at the locus I (*aapA1*/IsoA1) of the *H. pylori* was characterized. We showed that the *aapA1* gene codes for a small protein whose expression is toxic for *H. pylori*, and revealed several layers of post-transcriptional regulation that act in concert with the antitoxin to prevent toxin expression (as shown in [Figure 2.2](#)). The transcription of the *aapA1* gene generates a full-length mRNA of 250 nt denoted *aapA1*-FL (full-length) that is translationally inactive thanks to a long-distance interaction between the 5' and the 3' ends of the mRNA. This inactive message is highly stable and is slowly processed at its 3' end to generate a truncated transcript of 225 nt that becomes translatable. This active transcript, denoted *aapA1*-Tr (truncated), base-pairs with the IsoA1 asRNA to form an extended RNA heteroduplex that is rapidly degraded by the RNase III. In absence of IsoA1, the AapA1 toxin is

synthesized, leading to growth arrest and cell death. Overall, with this work, we characterized a new family of type I TA systems in *H. pylori* and highlighted the importance of mRNA folding in type I TAs regulation.



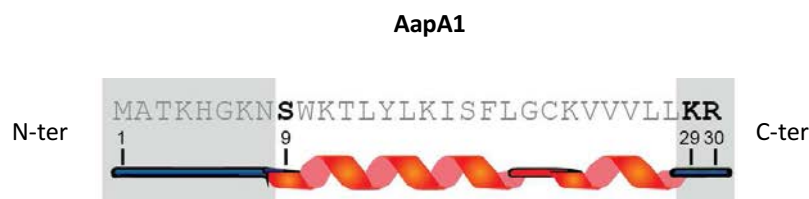
**Figure 2.2. Model of the *aapA1/IsoA1* system regulation.** This model is based on experimental data presented and described in our previous study (Arnion et al., 2017, see appendix 8.3 in Chapter 8). The *aapA1/IsoA1* operon organization is shown in the upper part of the figure. Thick arrows indicate promoters. The region encoding the AapA1 ('antisense associated peptide family A') ORF in the transcript is shown with a thick blue line. The Shine-Dalgarno (SD) sequence is represented by a yellow rectangle, while the anti-SD sequence is shown in red. The translationally active and inactive states of the different mRNA species are also indicated. The IsoA1 ('inhibitor of small ORF family A') antitoxin is shown green. Q1-4 refer to the research questions that are described in the 2.2 section of this Chapter. See text for details.

This newly identified family of type I TAs shares many similarities with other previously characterized type I TAs in enterobacteria. For instance, the *aapA1/IsoA1* system shares common features with the overlapping *cis*-encoded antisense RNA regulating TA systems characterized in *E. coli* (Kawano, 2012). Similar to the *hok/Sok* and the *ldr/Rdl* TA systems, the *aapA1/IsoA1* system encodes a small toxic protein (size < 50 aa) that is synthesized from a highly stable mRNA and which production is repressed by a small unstable asRNA. Moreover, as for *hok/Sok*, the *aapA1/IsoA1* also requires a 3' end processing to get activated (Thisted et al., 1994). However, in contrast to these two TA systems, AapA1 translation does not seem to require the translation of an overlapping small ORF



(i.e., the Mok and LdrX leader peptides required for Hok and Ldr translation, respectively) (Kawano et al., 2002; Thisted and Gerdes, 1992). Another major difference with other TA systems relies on the fact that the IsoA1 antitoxin is directly targeting the translation initiation region of the toxin gene. This specificity is due to its genomic organization, as it is fully complementary to the 5' UTR of the toxin-encoding mRNA, similarly to what was previously shown for the *symE/SymR* system in *E. coli* (Kawano et al., 2007). Finally, our bioinformatics analysis revealed that this TA locus belongs to a large family of type I TA systems (named *aapA/IsoA*) that is not only present on chromosome but also associated with mobile genetic elements (MGEs). This is an interesting observation, as the *aapA/IsoA* module may comply most of the characteristics (e.g., constitutive transcription, stable toxin-encoding mRNA and unstable antitoxin, and killing activity) present in TA systems acting as plasmid-stabilizer elements, suggesting a possible MGE-stabilizing function following a PSK-like mechanism.

Additionally, our group has recently characterized the AapA1 peptide (Korkut et al., in preparation). Similar to most type I toxins (with the exception of SymE and RalA) (Brantl and Jahn, 2015; Fozo et al., 2008; Wen and Fozo, 2014), the *aapA1* gene codes for a small hydrophobic protein that contains an alpha-helical transmembrane domain (TM). Indeed many type I toxins (as shown for the Hok (Verstraeten et al., 2015), DinQ (Weel-Sneve et al., 2013), Fst (Göbl et al., 2010), Ibs (Mok et al., 2010) and TisB (Unoson and Wagner, 2008)) act by altering membrane integrity (Brielle et al., 2016), as described in the section 1.4.2 in Chapter 1. To date, only a few of such TM domains have been experimentally validated (e.g., TisB and Ibs in *E. coli* (Grage et al., 2016; Mok et al., 2010; Steinbrecher et al., 2012), Fst in *E. faecalis* (Göbl et al., 2010), and SprA1 in *S. aureus* (Sayed et al., 2012)). We showed that the AapA1 peptide, despite having a small hydrophobicity index, has a strong affinity to the prokaryotic membrane and it is able to localize a soluble tag at the inner membrane of the bacterium. Plasmon Waveguide Resonance (PWR), solid-state Nuclear Magnetic Resonance (NMR) and molecular dynamics simulations, revealed that despite being a small protein (30 amino acids), the AapA1 is organized in three domains: 1) a disordered, positively charged N-terminal domain (8 amino acids), 2) a single alpha-helix domain spanning the membrane (~19 amino acids), and 3) a short and basic C-terminal domain (Figure 2.3). Furthermore, an alanine-scanning of AapA1 peptide in *E. coli*, revealed that the N-terminal domain is not required for toxicity, and these results are supported by phylogenetic studies showing that this region is the less conserved part of the AapA1 ORF.



**Figure 2.3. The AapA1 peptide is organized in three domains.** This Figure has been adapted from Korkut et al., in preparation. See text for details.

## 2.2. Understanding the regulation and biological role of type I TAs in *H. pylori*

The work carried out in the lab before the start of this thesis have raised several questions:

Q1) AapA1 toxicity was only tested using inducible systems, whether its endogenous expression is constitutive or not was unknown. Indeed, to date, most TA systems are studied using synthetic expression vectors and heterologous organisms (Fozo et al., 2008; Jahn and Brantl, 2013; Unoson and Wagner, 2008). However, this can lead to misinterpretations of the physiological effects of toxin expression (which would never be that high levels under endogenous conditions).

Additionally, the use of artificial expression systems does not allow the study of either transcriptional or post-transcriptional regulation. Considering that most of the regulation of type I TA systems occurs post-transcriptionally (due to the RNA nature of the antitoxin), this is a major limitation for the understanding of type I TA systems regulation.

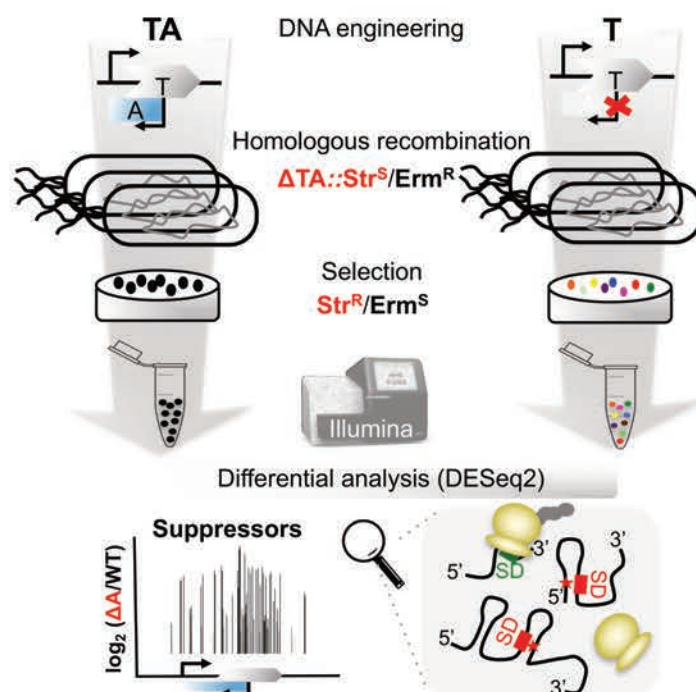
Q2) The AapA1 primary transcript was demonstrated to be inert for interaction with both, ribosomes and IsoA1 sRNA due to the presence of a long-distance interaction between both mRNA ends. However, since transcription/translation is coupled in bacteria, translation can start before the end of the message has been transcribed. How premature toxin translation (and antitoxin binding) is avoided during transcription remained unknown.

Q3) What is(are) the enzyme(s) responsible for the 3' end processing of the AapA1 mRNA?

Q4) Although the AapA1 peptide was demonstrated to target the inner membrane, whether its toxic action is a direct consequence of its interaction with the membrane (*e.g.*, by forming pores) or with other cellular components, remained enigmatic.

### 2.3. Our experimental strategy

1) To develop an approach to study chromosomally-encoded TA systems under native conditions. The approach, termed FASTBAC-Seq (Functional AnalysiS of Toxin-antitoxin systems in Bacteria by deep-Sequencing), combines a selection of genetic toxicity suppressors (*i.e.*, allowing survival in absence of antitoxin) with Next-Generation Sequencing (NGS) technique, and enables the high-throughput identification of *cis*-encoded (*i.e.*, within the TA locus) gene expression and functionality (toxicity) determinants of toxin-encoding genes with nucleotide resolution (see Figure 2.4).



**Figure 2.4. Pipeline of the FASTBAC-Seq approach.** FASTBAC-Seq is based on the lethality of toxin expression in antitoxin absence. The first step is to engineer the TA loci to inactivate the antitoxin promoter. Then, two PCR pools are generated, carrying the wild-type (TA) or the antitoxin-inactivated (T) TA locus. PCR products are introduced in the bacterial chromosome by homologous recombination. Importantly, the approach requires the prior generation of a erythromycin-resistant/streptomycin-sensitive

strain (Str<sup>S</sup>/Erm<sup>R</sup>), on which the TA locus of interest has been replaced by the counterselection marker *rpsL<sub>G</sub>-erm* in a streptomycin-resistant genetic background (*rpsL* K43R), as described by (Dailidiene et al., 2006; Pernitzsch et al., 2014). Unmarked transformants (having replaced the *rpsL<sub>G</sub>-erm* cassette by the TA-PCR products) are selected in the presence of streptomycin, pooled, and genomic DNA is extracted and used to amplify the TA locus. PCR products are used sequenced by Illumina paired-end sequencing approach and sequence datasets are analyzed using a differential analysis approach (e.g., DESeq2). Sequence analysis reveals with nucleotide resolution toxicity suppressor mutations, that represent gene expression and functionality determinants.

---

2) Apply FASTBAC-Seq to study the *aapA*/IsoA module present in the *H. pylori* chromosomal locus III (the *aapA3*/IsoA3 module). Such deep characterization aims to uncover the molecular underpinnings behind the regulation of the AapA3 toxin expression, beyond antitoxin RNA binding (e.g., transcription, mRNA decay, translation, and protein function). Thus, aiming to gain insights into the research questions presented above (Q1 to 4, [section 2.2 in Chapter 2](#)).

3) Understanding how this type I TA system is regulated under endogenous conditions may shed light into more challenging questions, such as what is its biological role (if any), or how can it be exploited with biotechnological (e.g., boost protein production) or biomedical (e.g., alternative antimicrobial strategies) purposes.



## CHAPTER 3. Thesis map

### 3.1. Organization

This thesis is organized into eight Chapters. In **Chapter 1**, the biological background of gene expression regulation in bacteria is reviewed, with the main focus on the study of post-transcriptional gene regulation and type I toxin-antitoxin (TA) systems in the RNase E and Hfq-lacking bacterium *H. pylori*. Additionally, a review document discussing the underlying mechanisms of type I TA systems regulation is presented (**Paper I**). In **Chapter 2**, the main objectives of this work are presented and justified. The organization of the thesis and the main scientific outputs as well as work contributed by others and collaborations is shown in **Chapter 3**.

Results are presented in **Chapter 4**, which is composed of three sections, corresponding to the three main publications fruit of this work. In **section 4.1**, the **Paper II** is presented. On it, we respond to the lack of tools to study chromosomally-encoded type I TA systems under endogenous conditions. A detailed protocol of the here developed approach (FASTBAC-Seq) for the identification of gene expression determinants on toxin-encoding genes belonging to type I TA systems (revealed in the form of genetic toxicity suppressors) is presented. The **sections 4.2** and **4.3**, correspond to the **Papers III** and **IV**, respectively. On them, we exploit the results obtained from the FASTBAC-Seq analysis of the *aapA3*/IsoA3 type I TA locus, hosted in the chromosome of *H. pylori*. More specifically, in **section 4.2**, we describe how the study of suppressor mutations in the *aapA3* toxin-encoding gene uncovered the existence of functional metastable RNA structures that are essential to impede the co-transcriptional translation of the toxin-encoding mRNA. In **sections 4.3** and **4.4**, we focus on the understanding of the enigmatic role of RNA processing in the regulation of the *aapA3*/IsoA3 system, and how this is determined by the RNA secondary structure.

In **Chapter 5**, the work is globally discussed and concluding remarks are presented. Additionally, remaining open questions and future perspectives are discussed. A description of the general methods used throughout the study as well as Paper-specific methods sections are described in **Chapter 6**. References are listed in **Chapter 7**. And finally, supplementary information of the **Paper III** and **IV**, as well as a copy of two studies that are closely related to this work (**related papers I and II**), is provided in **Chapter 8**.

### 3.2. Main scientific outputs

#### 1. Manuscripts published

The review document introducing to type I TA systems regulation presented in Chapter 1 section 1.4.1 was recently published elsewhere as:

**Paper I**      Masachis S. and Darfeuille F.\* (2018) *Type I toxin-antitoxin systems: regulating expression with small antisense RNAs*. Microbiol Spectr. 2018 Jul;6(4). doi: 10.1128/microbiolspec.RWR-0030-2018.

The method to characterize chromosomally-encoded type I TA systems presented in Chapter 4 section 4.1 has been recently published elsewhere as:

**Paper II**      **Masachis S.**, Tourasse N., Chabas S., Bouchez O. and Darfeuille F.\* (2018) *FASTBAC-Seq: Functional Analysis of Toxin-antitoxin systems in Bacteria by deep-Sequencing*. Methods in Enzymology: High- Density Sequencing Applications in Microbial Molecular Genetics, Vol. 612. Edited by A.J. Carpoussis. <https://doi.org/10.1016/bs.mie.2018.08.033>

## 2. Manuscripts in preparation

The work described in Chapter 4 sections 4.2 and 4.3 are in preparation for publication as:

**Paper III**      **Masachis S.**, Tourasse N., Lays C., Faucher M., Chabas S., Iost I., and Darfeuille F.\*. *Chromosomal inactivation of antitoxin RNA uncovers metastable structures embedded in a toxin-encoding mRNA*.

**Paper IV**      **Masachis S.**, Arnion H., Boissier F., Chabas S. and Darfeuille F.\*. *When mRNA folding rules decay: lessons from a type I toxin-antitoxin system*.

## 3. Additional work related to this thesis in a more indirect way can be found in Chapter 8 section 8.3 (related paper)

**Related Paper**      Arnion H., Korkut D.N., **Masachis Gelo S.**, Chabas S., Reignier J., Iost I. and Darfeuille F.\* (2017) *Mechanistic insights into type I toxin antitoxin systems in Helicobacter pylori*. **NAR**, 45(8):4782-4795

### 3.3. Contribution by others

The work presented on this thesis was performed under the supervision of the Dr. Fabien Darfeuille at the ARNA Laboratory, INSERM (Institute National de la Santé et Recherche Médicale) Unit 1212, Bordeaux, France. Parts of the work that have been contributed or performed in collaboration with others are presented below.

#### 1. Chapter 4 section 4.1 (**Paper II**)

Work contributed by others before my arrival to the group:

- The *rpsL<sub>CJ</sub>-erm* cassette was a generous gift from the Dr. Berg DE (Department of Molecular Microbiology, University Medical School, St. Louis, USA) and cloned into pSC-60 by the group of Cynthia Sharma (University of Würzburg, Germany).
- Development of the *rpsL<sub>CJ</sub>-erm* cassette-mediated deletion method in *H. pylori* was done by the engineer Sandrine Chabas (Darfeuille's Lab).

Work performed by collaborators:

- DNA Illumina sequencing was performed by Olivier Bouchez (Plateformed GeT-PlaGe- Genotoul, Centre INRA, BP52627, 31326 Castanet Tolosan).

- NGS data statistical analysis was performed by the Dr. Nicolas Jacques Tourasse (ARNA laboratory, INSERM U1212).

2. Chapter 4 section 4.2 (**Paper III**)

Work contributed by others before my arrival to the group:

- Preliminary studies on the *aapA3*/IsoA3 locus were carried out by the Postdoc Claire Lays (Darfeuille's lab).
- AapA3 RNA *in vitro* structure probing shown in Figure 4.2.2 was previously done by the Master student Marion Faucher under supervision of Fabien Darfeuille.
- RNase III deletion in the B128 *H. pylori* strain was previously done by the Dr. Isabelle Iost (Darfeuille's lab).
- Protocol for *H. pylori* polysome fractionation was developed by the Dr. Isabelle Iost.

3. Chapter 4 section 4.3 (**Paper IV**)

Work contributed by others before my arrival to the group:

- Preliminary PNPase deletion mutant and study of its role on wild type *aapA3* mRNA maturation was performed by the engineer Hélène Arnion (Darfeuille's lab).
- PNPase deletion strain in *rpsL* K43R genetic background was generated by the engineer Sandrine Chabas.

Work performed by others:

- RNase J protein purification was performed by the engineers Fanny Boissier and Snadrine Chabas (Darfeuille's lab).
- RNase J *in vitro* cleavage assay shown in Figure 4.3.6D was performed by the engineer Fanny Boissier.

## CHAPTER 4. Results and discussion

### 4.1. Paper II: FASTBAC-Seq: Functional Analysis of Toxin-antitoxin systems in Bacteria by deep-Sequencing

ARTICLE IN PRESS

## FASTBAC-Seq: Functional Analysis of Toxin–Antitoxin Systems in Bacteria by Deep Sequencing

Sara Masachis\*, Nicolas J. Tourasse\*, Sandrine Chabas\*,  
Olivier Bouchez<sup>†</sup>, Fabien Darfeuille\*,<sup>1</sup>

\*ARNA Laboratory, INSERM U1212, CNRS UMR 5320, University of Bordeaux, Bordeaux, France

<sup>†</sup>Plateforme GeT-PlaGe-Genotoul, INRA Auzeville, Castanet-Tolosan, France

<sup>1</sup>Corresponding author: e-mail address: fabien.darfeuille@inserm.fr

<< Diffusion non-autorisée / unauthorized copying or distribution >>

## 4.2. Paper III: A genetic selection reveals functional metastable structures embedded in a toxin-encoding mRNA.

In [Chapter 4 section 4.1 \(Paper II\)](#), we stated the need for tools allowing the study of chromosomally-encoded type I TA systems under endogenous conditions and described the protocol of the FASTBAC-Seq approach ([Masachis et al., 2018](#)). The first part of this Chapter will present a global view of the main results obtained from the characterization of the *aapA3/IsoA3* type I TA system using FASTBAC-Seq. Secondly, we will focus on the study of some of the identified toxicity suppressor mutations. FASTBAC-Seq revealed unanticipated regulatory elements embedded in the AapA3 toxin-encoding mRNA. Despite most suppressors located in the AapA3 ORF (and are described in [Chapter 4 section 4.1](#)), one-third mapped to non-coding regions of the toxin mRNA. Some of such mutations, targeted well-known regulatory elements (*e.g.*, SD sequence), however, others did not and their existence struck us.

Translation initiation is assumed to be the rate-limiting step of protein synthesis in bacteria ([Duval et al., 2015](#); [Kozak, 2005](#); [Milón and Rodnina, 2012](#); [Nakagawa et al., 2017](#); [Simonetti et al., 2009](#)). The canonical translation initiation mechanism involves the base-pairing interaction between the Shine-Dalgarno (SD) sequence in the mRNA and the anti-SD sequence in the 16S rRNA embedded in the 30S ribosomal subunit ([Shine and Dalgarno, 1974](#); [Steitz and Jakes, 1975](#)). Together with the Initiation Factor 2 (IF-2) that kinetically assists the specific interaction between an anticodon present in the fMet-tRNA<sup>fMet</sup> and an optimally-spaced ([Chen et al., 1994](#)) start codon (AUG or UUG/GUG), they form the so-called translation initiation complex. Several studies have previously highlighted the importance of SD accessibility for translation initiation ([Bhattacharyya et al., 2018](#); [Espah Borujeni et al., 2014](#); [Kozak, 2005](#)). In this Chapter, we will present a special set of suppressor mutations laying within the non-coding region of the AapA3 toxin-encoding mRNA. Such mutations, act as energetic traps, freezing the mRNA in translationally inert conformations. The study of these mutations revealed the existence of transient (metastable) functional RNA hairpins containing a CU-rich sequence motif. Through this motif, they co-transcriptionally mask the SD sequence and impede the toxin translation initiation during the mRNA synthesis. Our work highlights the importance of co-transcriptional folding to impede the premature toxin translation and allow the presence of these toxic genes in bacterial chromosomes.

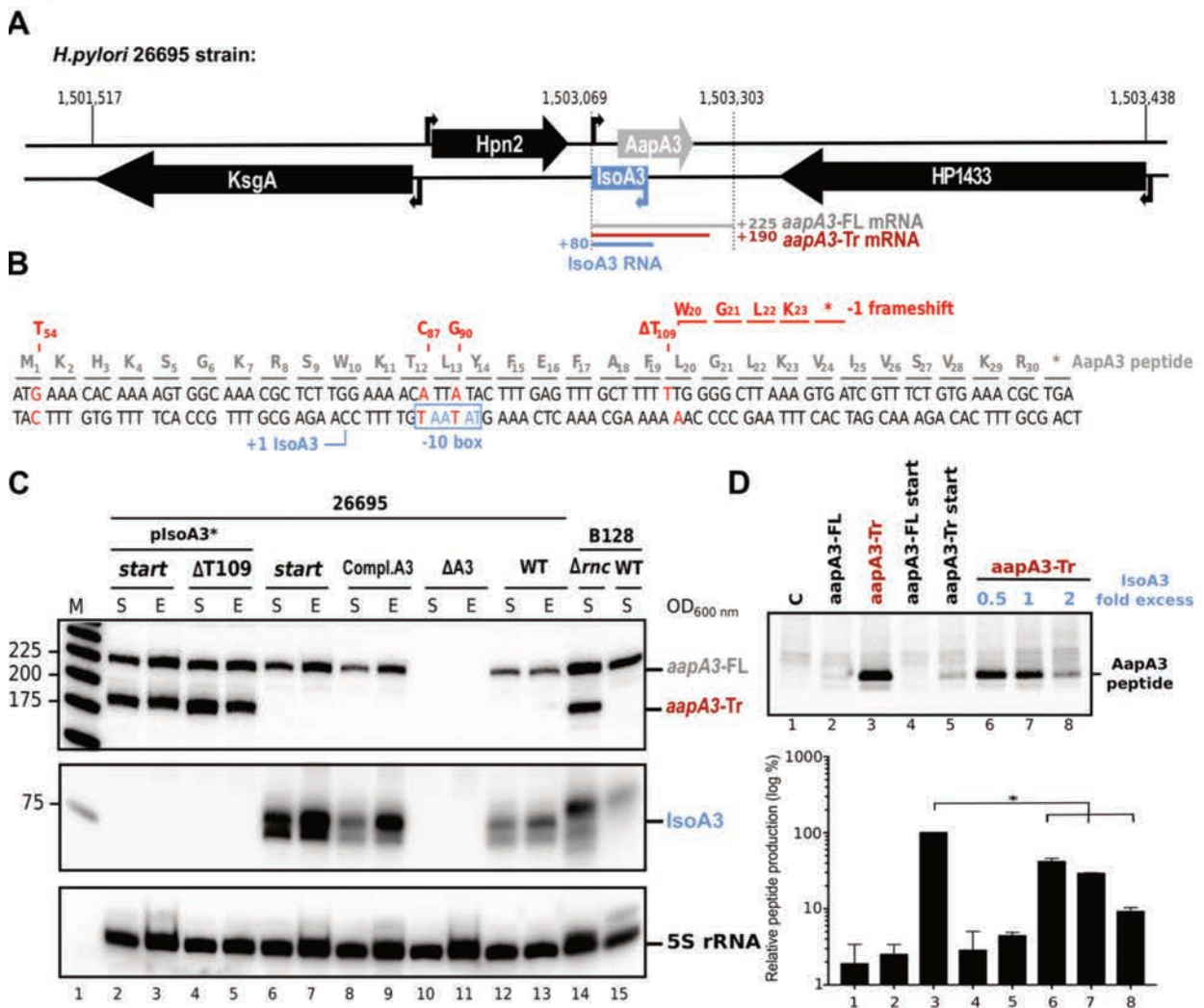
### 4.2.1. Results

#### 4.2.1.1. The small antisense RNA IsoA3 is essential to prevent AapA3 translation

We previously identified a new family of type I TA loci on the chromosome of *H. pylori*, the *aapA/IsoA* family ([Arnion et al., 2017](#), see [appendix 8.3 in Chapter 8](#)). Here, we studied the *aapA3/IsoA3* module ([Figure 4.2.1A](#); for sequence details see [Figure 8.1.1B in the Appendix](#)). As other TAs of this family, the *aapA3/IsoA3* locus codes for an antisense RNA, IsoA3 (80 nucleotides), encoded in the *cis*-overlapping orientation to a small ORF, AapA3 (30 amino acids). The AapA3 peptide shares 60% sequence homology with the AapA1 peptide, whose ectopic expression is toxic in *H. pylori* ([Arnion et al. 2017](#)). However, in contrast to AapA1, cloning of AapA3 peptide under an IPTG-inducible promoter is not viable. Therefore, we decided to ask whether AapA3 expression from the chromosome was also toxic. To do so, we inactivated the antitoxin promoter by introducing two point mutations ([Figure 4.2.1B](#), pIsoA3\* in all figures), as previously described ([Arnion et al., 2017](#)).

Chromosome manipulations were performed using the counter-selection cassette *rpsL<sub>C</sub>-erm*, which allows the generation of unmarked transformants (Dailidiene et al., 2006; Pernitzsch et al., 2014).

As expected, chromosomal inactivation of the IsoA3 promoter was not viable unless an additional mutation in the *aapA3* start codon or a -1 frameshift mutation leading to the generation of a peptide lacking the last seven amino acids was introduced (see Figure 4.2.1B for mutational strategy details). These results demonstrate that AapA3 peptide expression from its chromosomal locus is constitutive and toxic. Remarkably, none of the IsoA RNAs produced *in trans* from the five other *aapA*/IsoA chromosomal loci (I, II, IV, V and VI) counteracted the absence of IsoA3 expression demonstrating that their regulation is strictly module-specific, as previously reported (Sharma et al. 2010). The absence of IsoA3 transcript observed by Northern Blot analysis (Figure 4.2.1C, lanes 2 to 5) confirmed the successful chromosomal inactivation of the IsoA3 promoter by the introduction of only two single-point mutations on its -10 box. Locus complementation was successfully achieved with no change in the expression (Figure 4.2.1C, lanes 8 and 9).

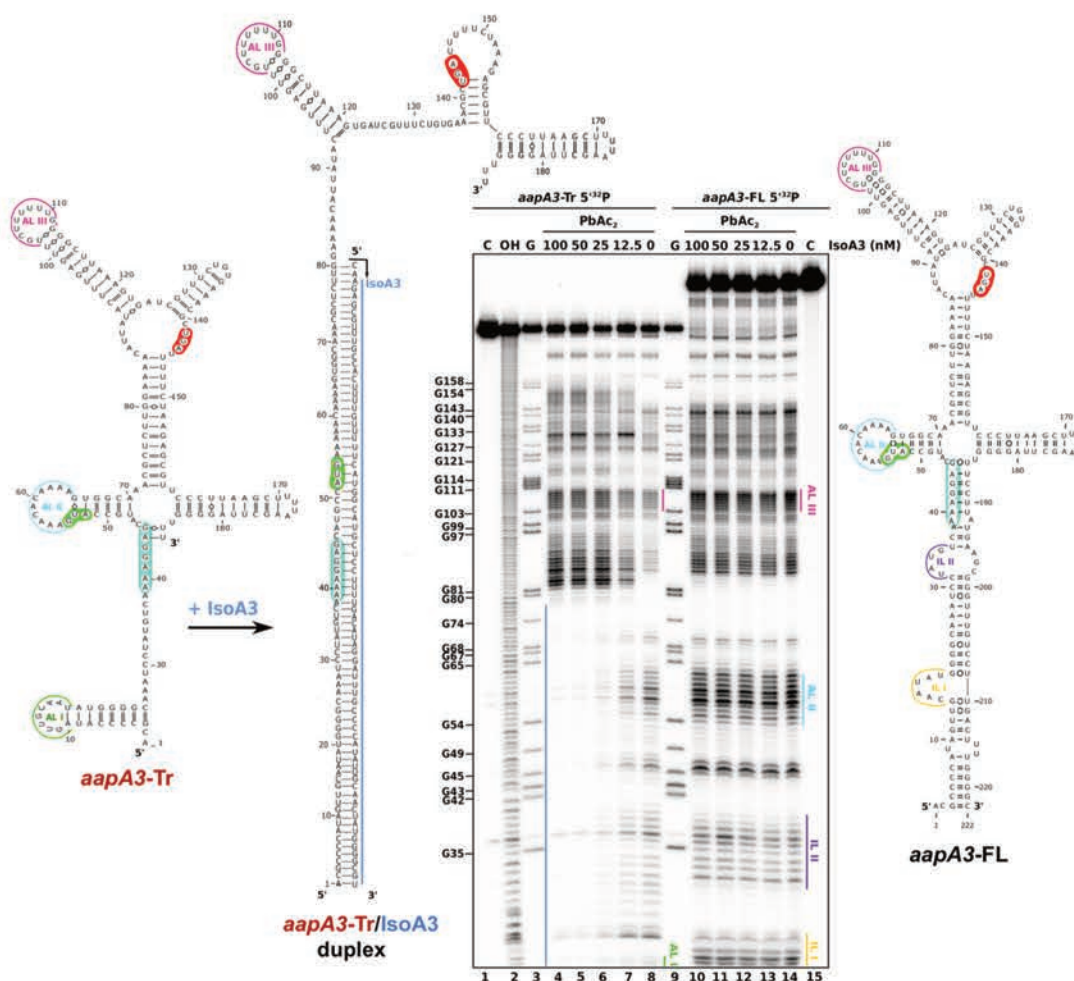


**Figure 4.2.1. IsoA3 small RNA is essential to prevent AapA3 translation.** (A) Organization of the *aapA3*/IsoA3 locus in the *H. pylori* 26695 strain. Grey arrow, AapA3 ORF; blue box, IsoA3 RNA. Small grey and blue arrows, -10 box of each transcript. Flanking ORFs Hpn2 and HP1433 are represented by black arrows. Grey, red and blue horizontal bars represent *aapA3*-FL, *aapA3*-Tr and IsoA3 RNAs, respectively. A 3' processing event that generates the *aapA3*-Tr is shown. (B) Sequence of the *aapA3* locus and the AapA3 peptide. Construction of a strain with the IsoA3 promoter mutated (A87C/A90G, pIsoA3\*) was possible in the presence of either the G54T mutation inactivating the AapA3 start codon (start) or the



$\Delta T109$  generating a -1 frameshift. **(C)** Total RNA extracted from stationary (S) or exponential (E) cultures of the indicated strains was subjected to Northern Blot analysis. The same membrane was successively probed for *aapA3*, IsoA3, and 5S rRNA transcripts. The *aapA3*-FL (full length), *aapA3*-Tr (3'-end truncated), and IsoA3 transcripts are shown. **(D)** Translation assays were performed using 0.5  $\mu$ g AapA3 RNAs. IsoA3 in 0.5, 1 or 2 molar excess was used for inhibition assays. [ $^{35}$ S]-Met was used for labeling. Control lane (C) shows the translation background obtained without exogenous mRNA. Phosphorimager-generated images were used for calculation of translation efficiencies (peak areas). Lower panel graph represents quantification of three independent experiments (\* $P < 0.0001$  values according to unpaired t-test).

In line with previous results (Arnion et al., 2017), a shorter form of the AapA3 mRNA (*aapA3*-Tr) was observed in absence of IsoA3 expression (Figure 4.2.1C, lanes 2 to 5). The exact sequence of this truncated mRNA differs from the full-length (*aapA3*-FL) by lacking approximately the last 35 nt of the 3' end, as determined by RNA-Seq (Sharma et al., 2010) and 3' RACE experiments (see Chapter 4 section 4.3). *In vitro* translation assays demonstrated that only the *aapA3*-Tr mRNA is an efficient translation substrate (Figure 4.2.1D, lane 3). Translation of *aapA3*-FL (Figure 4.2.1D, lane 2) was equivalent to that of the negative control (Figure 4.2.1D, lane 5) and close to the translational background obtained in absence of exogenous mRNA (Figure 4.2.1D, lane 1). The exclusive accumulation of this truncated and translatable toxin mRNA form in absence of IsoA3, and this absence not affecting the levels of full-length transcript demonstrates that IsoA3 RNA specifically targets *aapA3*-Tr *in vivo*.



**Figure 4.2.2.** IsoA3 inhibits *aapA3*-Tr translation by binding to its SD region. ~0.1 pmol 5' end [ $^{32}$ P]-labeled *in vitro* transcribed *aapA3*-FL and *aapA3*-Tr RNA was used. Lead probing in presence of



increasing concentrations of IsoA3 cold RNA (0-100 nM) was performed. Untreated RNA (lanes 1 and 15, denoted C) and partially alkali digested RNA (denoted OH, lane 2) served as control and ladder, respectively. Positions of all G residues revealed upon T1 digestion under denaturing conditions are indicated relative to the transcription start site of the *aapA3* gene. Cleaved fragments were analyzed on an 8% denaturing PAA gel. 2D structure predictions were generated with the RNAfold Web Server (Gruber et al., 2008) and VARNA RNA (Darty et al., 2009) was used to perform the drawing. The region involved in RNA duplex formation between IsoA3 and *aapA3*-Tr mRNA is indicated with a blue line; internal and apical loops are represented by IL and AL, respectively; AapA3 start codon is highlighted in green; AapA3 stop codon is highlighted in red; AapA3 SD sequence is highlighted in turquoise.

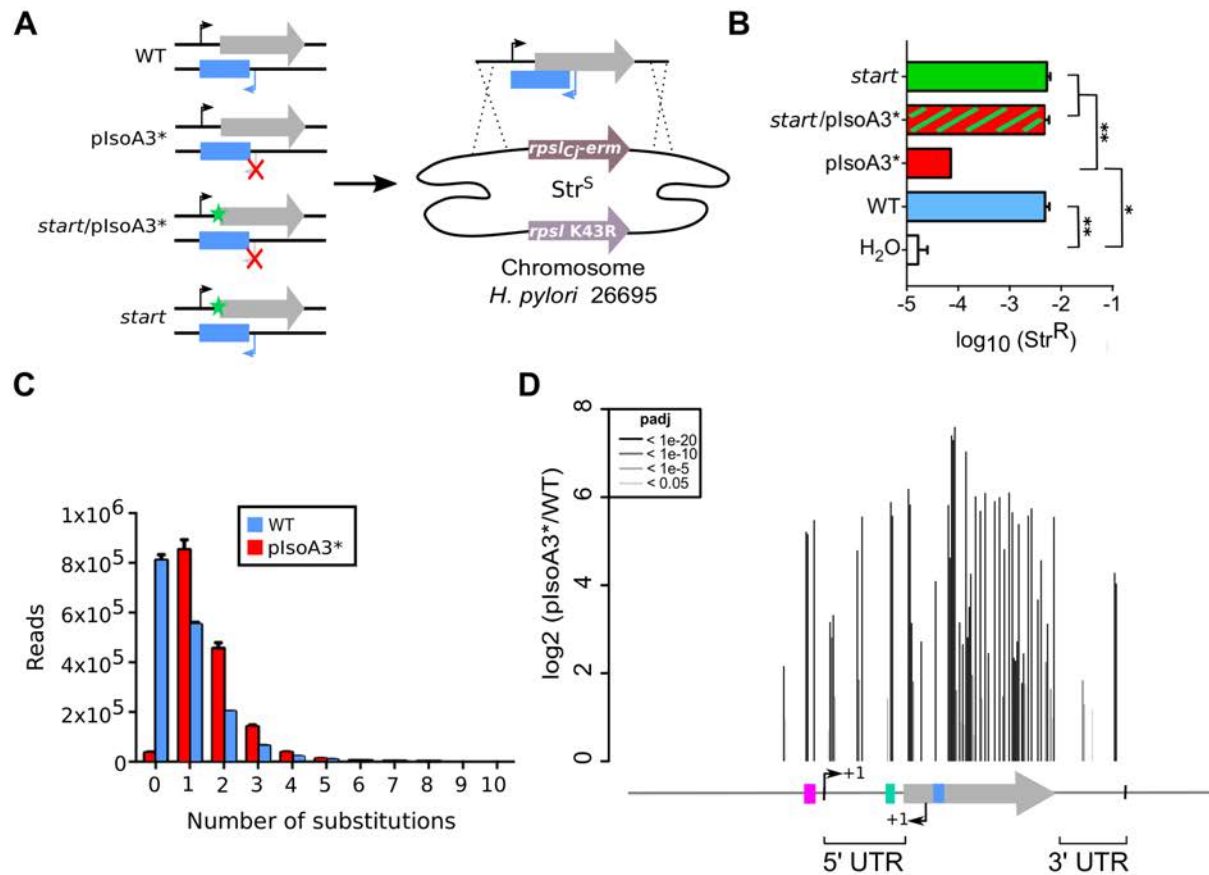
---

Furthermore, the accumulation of *aapA3*-Tr in an RNase III knockout strain pointed at a double-strand-specific mRNA decay pathway. *In vitro* structure probing of the different AapA3 mRNA forms in presence or absence of IsoA3 (Figure 4.2.2) confirmed that IsoA3 RNA only interacts with the *aapA3*-Tr mRNA (Figure 4.2.2, lanes 4 to 8). Base pairing between the *aapA3*-Tr mRNA and IsoA3 RNA creates an extended RNA heteroduplex of 80 base pairs (Figure 4.2.2, lane 4). This IsoA3/*aapA3*-Tr RNA duplex is translationally inert, as shown by *in vitro* translation assays (Figure 4.2.1D, lanes 6, 7 and 8) and by the viability of an RNase III deleted strain on which the accumulation of *aapA3*-Tr reflects the stabilization of this RNA duplex (Figure 5.1C, lanes 14 and 15), consistently with previous observations (Arnion et al. 2017).

Taken together, we demonstrate that IsoA3 small RNA continuously represses *aapA3* constitutive expression at the translational level by forming a stable RNA heteroduplex. This RNA duplex is then targeted by the double-strand-specific ribonuclease RNase III, leading to a rapid, but not essential, turnover of the translationally active toxin-encoding mRNA.

#### 4.2.1.2. Decoding AapA3 toxicity determinants with nucleotide resolution

To study AapA3 toxicity in a quantitative manner, we generated four PCR constructs carrying a wild type or mutated version of each of the TA components as described in Figure 4.2.3A. We next introduced them into a *H. pylori* strain on which the *aapA3*/IsoA3 locus had been previously replaced by the *rpsL<sub>CJ</sub>-erm* counter-selection marker ( $\Delta$ *aapA3*/IsoA3::*rpsL<sub>CJ</sub>-erm*/K43R) (Dailidiene et al., 2006; Pernitzsch et al., 2014). Markless transformants having undergone homologous recombination were then selected on streptomycin-containing plates. For each transformation, the exact number of streptomycin-resistant colonies was determined and normalized to the total number of colony forming units (Figure 4.2.3B). As expected, when no-DNA was used for transformation (H<sub>2</sub>O) only phenotypic revertants having mutated the *rpsL<sub>CJ</sub>* gene present on the *rpsL<sub>CJ</sub>-erm* cassette were selected (data not shown). In line with previous results, the lack of IsoA3 expression led to a strong reduction (1.83 log-fold) in the number of transformed strains compared to WT or the double mutant constructs (Figure 4.2.3B). These results confirmed that chromosomal expression of AapA3 is toxic and hampers the creation of recombinant strains lacking antitoxin expression. However, the number of transformants obtained in absence of antitoxin was significantly higher than the one obtained when no DNA was used for transformation (5-fold) (Figure 4.2.3B). This reflects the development of so-called suppressors strains. AapA3/IsoA3 locus sequencing of 100 of such strains revealed that all of them contained mutations in the AapA3 ORF (mainly missense and nonsense, data not shown).



**Figure 4.2.3. Unveiling local toxicity determinants with nucleotide resolution.** (A) Four different *aapA3*/*IsoA3* locus constructs were generated by PCR: wild-type, WT; *IsoA3* promoter inactivated by the point mutations A87C and A90G, pIsoA3\*; *AapA3* start codon inactivated by the point mutation G54A, *start*; and *AapA3* start codon and *IsoA3* promoter inactivated, *start*/pIsoA3\* (left panel). PCR products were used for *H. pylori* natural transformation. Homologous recombination allowed the replacement of the counter selection cassette *rpsL*<sub>CJ</sub>-*erm* (*Str<sup>S</sup>*/*Erm<sup>R</sup>*) by the PCR-generated constructs in an *H. pylori* 26695 *rpsL* mutated genetic background (*rpsL* K43R, *Str<sup>R</sup>*). Recombinant strains were selected in presence of streptomycin (right panel). (B) Number of *Str<sup>R</sup>* upon transformation was calculated. Error bars represent s.d.; n=3 biological replicates (\*\*P<0.0001; \*P<0.0005 values according to unpaired t-test). (C) Number of reads containing 0 to 10 substitutions. Error bars represent s.d.; n=3 biological replicates. (D) Positional analysis of single-nucleotide substitutions on the *aapA3*/*IsoA3* locus. Bar plot shows the  $\log_2$  of the fold change (pIsoA3\*/WT) for the 70 positions whose adjusted p-value (padj) was lower than 0.05. Bars are drawn in different shades of grey according to the p-value. 5' and 3' UTR, 5' and 3' untranslated regions; red box, *aapA3* -10 box; black arrow with "+1", *aapA3* +1 Transcription Start Site (TSS); turquoise box, *aapA3* Shine-Dalgarno sequence (SD); grey arrow, *AapA3* Open Reading Frame (ORF); black arrow with "+1" inside *aapA3* ORF, *IsoA3* TSS; blue box inside *aapA3* ORF, *IsoA3* -10 box.

To explore the complete landscape of suppressors we next scaled-up the transformation assay. The WT or inactivated antitoxin promoter (pIsoA3\*) PCR products were used as DNA substrates for transformation as described previously, and we performed the assay in three independent biological replicates for each construct. Approximately 60,000 strains selected upon transformation with each of the constructs were collected, genomic DNA was extracted, and an amplicon of 426 nt was sequenced by the Illumina paired-end approach. Consistently with preliminary results, deep-sequencing data analysis showed that 97.7% of the strains transformed with the inactivated antitoxin promoter PCR product contained mutations (pIsoA3\*, Figure 4.2.3C). A strong mutation rate was also observed when the WT PCR product was used (51.2%) (WT, Figure 4.2.3C). This demonstrates that mutations come from the Taq DNA polymerase error rate during the PCR. The strong biases can be explained by

an unanticipated technical artifact linked to the combined use of a low-fidelity Taq DNA polymerase and PCR assembly, which led to a strong mutation rate in the overlapping region used for PCR assembly (nucleotides 80, 81 and 82 from *aapA3* +1 site). This artifact strongly reduced sequencing depth and made it impossible to interpret the possible co-variation of double and triple mutations. Nevertheless, it did not alter the interpretation of single point mutations. Consequently, on this manuscript, we will focus on single nucleotide mutations having been significantly enriched (adjusted False Discovery Rate  $\text{padj} \leq 0.05$ ) in absence of antitoxin relative to WT (Figure 4.2.3D).

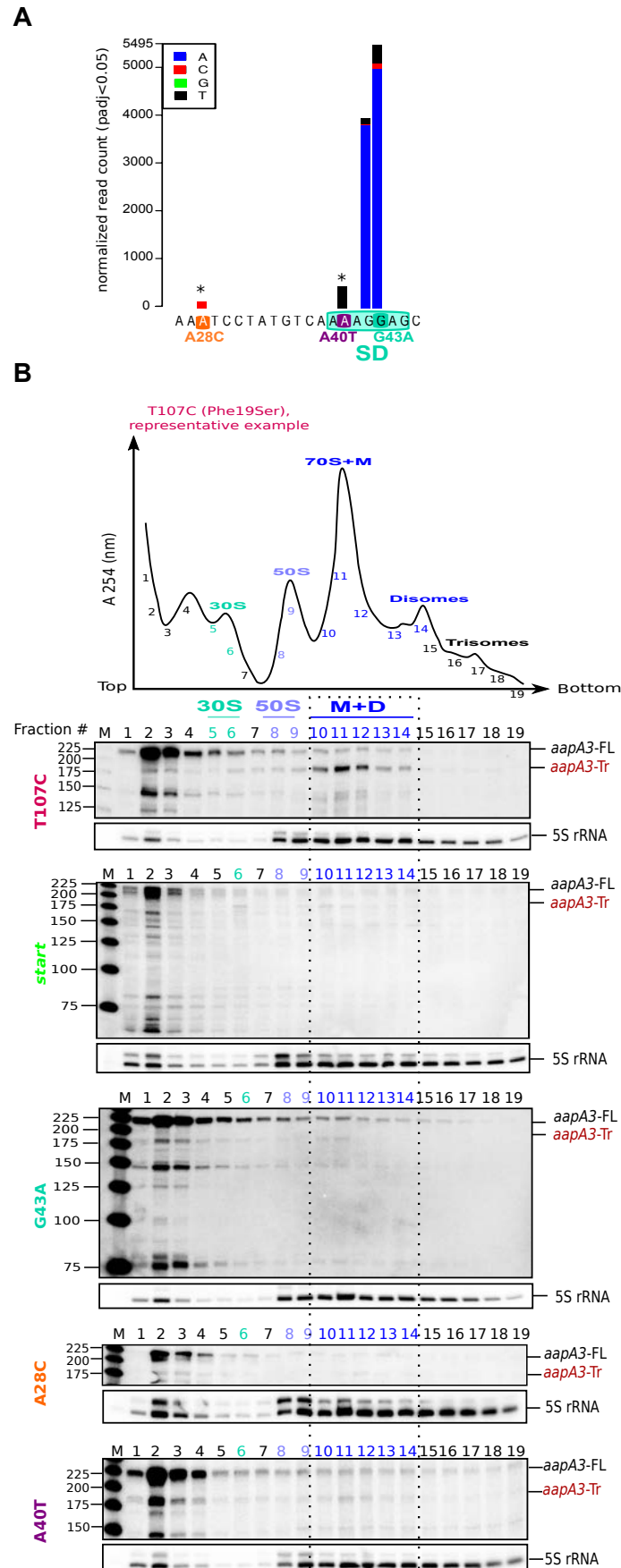
Analysis of the number of mutations per read in the complete sequencing dataset (all replicates combined) showed that more than half of the strains transformed with the pIsoA3\* PCR product (51.8% out of ~5.1 million) were mutated at a single nucleotide position (Figure 4.3.3C). This result demonstrates that single point mutations are sufficient to abolish AapA3 toxin activity and/or expression. A reduced number of pIsoA3\* strains (2.1%) had a wild-type locus sequence (Figure 4.2.3C, pIsoA3\* zero mutations), which could be explained by sequencing error rates (around 1.5%) and/or suppressor mutations lying in regions outside the loci (not sequenced). Single nucleotide mutations were overwhelmingly substitutions, favored by PCR biases. Only 4% were insertions and deletions (indels). Hierarchical clustering analysis revealed that the location and frequency of single substitutions were highly similar in the three biological replicates, indicating that locus coverage was close to optimum (Figure 8.1.2A in the Appendix). Single nucleotide indels were almost exclusively present in AapA3 ORF, where they would lead to the generation of truncated or longer forms of the peptide. Contrary to substitutions, indels were rarely found in non-coding regions, and when they were, they were not present in all three replicates. This reflects insufficient sequencing depth coming from their under-representation in the PCR products (Figure 8.1.2B and 8.1.2C in the Appendix). Importantly, depending on the type of statistical analysis, “position-specific” or “nucleotide-specific” (see Material and Methods for details), there were statistically enriched substitutions in absence of antitoxin ( $\text{padj} \leq 0.05$ ) at 70 or 72 positions within the *aapA3*/IsoA3 locus, respectively (65 positions in common between the two analyses). Positions appearing in just one of the analyses include (relative to *aapA3* +1 site): 1) in the 5' non-coding DNA, positions -26 and -7; 2) in the 5' UTR, position +28; 3) in AapA3 ORF, positions +64 and +97; and 4) in the 3'UTR, positions +146 and +177. Such positions have generally a close-to cut-off  $\text{padj}$  value, but not in all cases as for instance, position +28 (that will be extensively studied here) had a highly significant  $\text{padj}$  value of  $7.2 \times 10^{-6}$ .

Expectedly, the highest mutation densities (obtained by dividing the number of mutated nucleotides by the total number of nucleotides in the regions of interest) located in the AapA3 ORF (53%) (grey arrow, Figure 4.2.3D), as well as in well-known regulatory regions such the -10 box of the toxin mRNA (66.7%) (Figure 8.1.3 in the Appendix) and the SD sequence (42.8%) (Figure 4.2.5A). In the *aapA3* -10 box, out of the six potential targets (5'-TAGGAT-3'), suppressors mostly appeared on the two first nucleotides at the 5' end and the last nucleotide at the 3' end (Figure 8.1.3 in the Appendix). This allowed us to identify the minimal functional *aapA3* -10 box motif 5'-TANNNT-3', which is in agreement with previous studies (Sharma et al., 2010). This result also confirmed that the arbitrary chosen False Discovery Rate cut-off ( $\text{padj} \leq 0.05$ ) was stringent enough to avoid false positives but permissive enough to allow the identification of suppressors. Interestingly, a mutation upstream of the *aapA3* -10 box was significantly enriched (position -26 from the +1 site, Figure 8.1.3 in the Appendix). The existence of such suppressor is intriguing, and if confirmed, could suggest a potential transcriptional regulation of the *aapA3* gene. Remarkably, seventeen mutations were unveiled in the 5' and 3' untranslated regions lying outside the already-known regulatory elements (Figure 4.2.3D). In the present work, we will focus on the study of three of them.

#### 4.2.1.3. A single nucleotide substitution is sufficient to inhibit AapA3 toxin translation

The genetic selection of suppressors allowed us to characterize the toxin SD sequence with nucleotide resolution (Figure 4.2.4A). Among the seven mutable positions (5'-AAAGGAG-3'), the two central guanine nucleotides (at position +42 and +43 from *aapA3* +1) were the most abundantly enriched in absence of antitoxin. As expected from PCR biases, the transition G>A was preferentially enriched in both cases. The efficient SD sequence inactivation by one of these mutations (G at position +43 into A) was further validated *in vivo* (G43A, Figure 4.2.4B). Interestingly, the single-nucleotide transversion A to T at position +40 from *aapA3* +1 (A40T, Figure 4.2.4) within the SD sequence appeared as being underrepresented in the read count. The selection of a transversion despite their underrepresentation in the PCR product used for transformation (due to PCR biases), together with its 'isolated' character (none of the flanking adenine residues have been selected) was striking and drove us to study this suppressor in more depth. Moreover, an additional mutation in close proximity to the SD sequence was identified, the single-nucleotide transversion A to C at position +28 from *aapA3* +1 (A28C, Figure 4.2.4). We generated these two mutant strains for further study.

Due to their proximity to the SD sequence, we first tested if they could interfere with AapA3 translation. We assessed AapA3 translation by polysome fractionation coupled to Northern Blot analysis. A suppressor strain mutated in the AapA3 ORF at the residue 19 (Phe 19 to Ser achieved by the single nucleotide substitution T +107 to C from *aapA3* +1) was used as positive translation control (T107C, Figure 4.2.4B). Polysome fractionation of this strain provided the first *in vivo* prove that the toxin full-length mRNA (*aapA3*-FL) is translationally inert (only 3.5% located in translating fractions; T107C panel in Figure 4.2.4B, lanes 10 to 14) contrary to the 3' end-truncated isoform (*aapA3*-Tr), which appears strongly associated with monosome and disome fractions (72.9% located in fractions 10 to 14; T107C panel in Figure 4.2.4B). Its poor presence in polysomic fractions reflects on the ORF, length since a ribosome occupies a region of approximately 90 nucleotides, it seems coherent that it can accommodate a maximum of two ribosomes. Similar polysome fractionation experiments were performed on the previously generated *start*/pIsoA3\* and G43A strains, both used as negative translation controls (*start* and G43A panels, Figure 4.2.4B). Strikingly, we observed high levels of AapA3 mRNA degradation that were not observed before in standard total RNA samples. The extended time of sample collection and treatment prior to RNA extraction that exists in the polysome fractionation protocol but not in the standard total RNA extraction could explain this effect. This increased degradation was especially evident in strains producing a non-translatable AapA3 mRNA (*i.e.* *start* and G43A), as previously observed for other mRNAs (Iost and Dreyfus, 1995). Unfortunately, this effect impeded the quantification of *aapA3*-Tr band intensities on the *start* strain (shown in red as not-detectable (n.d.), *start*, Figure 4.2.4B). The fraction of *aapA3*-Tr mRNA on translating fractions on the SD suppressor strain was strongly reduced (4.2%) (G43A panel, Figure 4.2.4B, fractions 10 to 14) compared to the WT strain (9.1%) (fractions 10 to 14, Figure 8.1.4A in the Appendix) and suggests that a certain amount of AapA3 peptide may be synthesized despite the lack of a canonical AUG translational start codon. Remarkably, a very strong translation reduction was observed in the A28C and the A40T suppressors. The mutation A40T had the strongest effect with only 6.7% of *aapA3*-Tr associated with translating ribosomes (Figure 4.2.4B, A40T panel, lanes 10 to 14) compared to the 18.9% of the A28C strain (Figure 4.2.4B, A28C panel, lanes 10 to 14) (see Figure 8.1.7 quantification graphs in the Appendix). Likewise, *in vitro* translation assays confirmed the results observed *in vivo* (Figure 8.1.4C in the Appendix), demonstrating that the A28C and the A40T suppressors act by reducing AapA3 translation efficiency *in vivo*.



**Figure 4.2.4. Single-point mutations in the *aapA3* 5' UTR inhibit its translation in absence of IsoA3.**  
**(A)** Nucleotide substitutions significantly enriched (padj≤0.05, “nucleotide-specific” analysis) in pIsoA3\*



compared to WT sequences within the *aapA3* Shine-Dalgarno (SD) region. Asterisks (\*) above bars indicate transversion mutations. The A28C (adenine at position +28 from the *aapA3* TSS mutated to cytosine), A40T (adenine at position +40 from the *aapA3* TSS mutated to thymine), and G43A (guanine at position +43 from the *aapA3* TSS changed to adenine) mutations are highlighted in orange, purple and turquoise, respectively. **(B)** Cell lysate of the indicated *aapA3* mutant strains (T107C, Phe19Ser; start, AapA3 start codon mutated; G43A; A28C; and A40T) was subjected to ultracentrifugation through a sucrose gradient under polysome stabilization conditions (+ Chloramphenicol). A profile at OD254 was recorded. Peaks of the free 30S and 50S subunits, 70S ribosomes (free and translating) and polysomes are indicated. RNA was extracted from each fraction and equal volumes of each extract were subjected to Northern Blot analysis. The same membrane was successively probed for the *aapA3*, *IsoA3*, and 5S rRNA transcripts. The different transcripts are annotated as: *aapA3*-FL, *aapA3*-Tr, and 5S rRNA (loading control). M, monosomes; D, disomes.

Together, these results demonstrate that it is possible to overcome antitoxin absence by reducing, even if not to abolishment, toxin translation efficiency. Remarkably, this translation inhibition can be achieved by a single nucleotide substitution in the 5' UTR of the toxin-encoding mRNA.

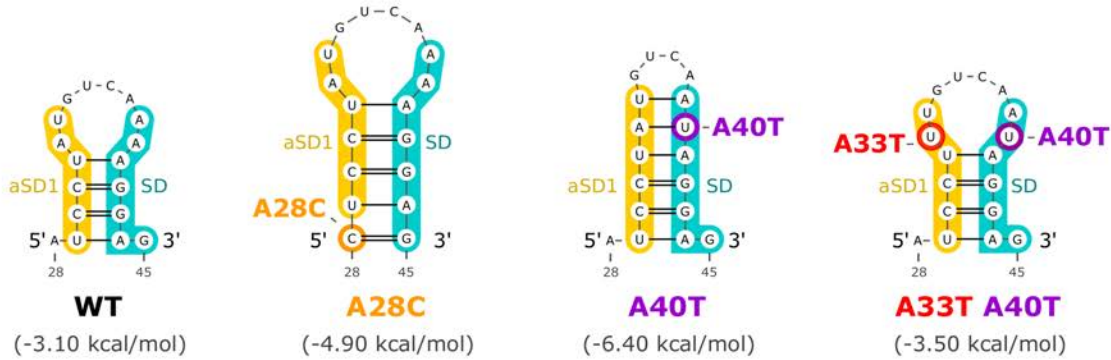
#### 4.2.1.4. The suppressors A28C and A40T inhibit toxin translation by stabilizing a metastable SD-sequester hairpin

Because the suppressor A40T locates within the AapA3 mRNA SD sequence, we next tested if it could act by SD sequence inactivation. To do so, we studied its location on the AapA3 mRNA secondary structure. In wild-type context, the SD sequence on the active toxin mRNA is accessible for interaction with both, ribosomes (Figure 4.2.1D) and *IsoA3* RNA (Figure 4.2.2, lanes 4 to 8). RNAfold (Gruber et al., 2008) secondary structure predictions suggested that A40T could stabilize the mRNA in an SD-sequestered state by extending a pre-existing anti-SD sequence embedded on a metastable hairpin by one base-pair (anti-SD sequence 1 (aSD1), Figure 4.2.5A). To investigate whether this SD-sequestering hairpin exists *in vivo*, we next performed a transformation assay. Three constructs were used: 1) the suppressor A40T; 2) a double mutant carrying a mutation in the opposite hairpin strand, expected to restore WT stem stability by disrupting the extra base-pair created by the suppressor (A33T / A40T); and 3) because the double mutation was expected to be toxic, a strain containing an additional mutation on the AapA3 start codon was generated (A33T / A40T / start, Figure 4.2.5C). Transformation assay was performed as previously described (see Material and Methods section).

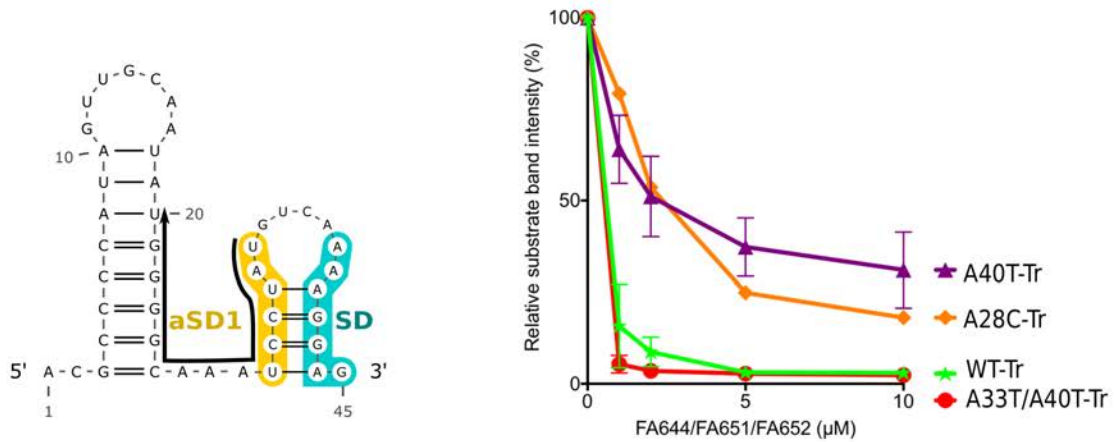
As expected, the suppressor A40T was not toxic (A40T, Figure 4.2.5D). However, a significant reduction (2 log-fold) in the number of Str<sup>R</sup>/CFU was observed with the A33T / A40T construct (A33T / A40T, Figure 4.2.5C). This effect disappeared when the toxin start codon was mutated (A33T / A40T / start, Figure 4.2.5C) demonstrating that toxicity comes from AapA3 peptide synthesis. Unfortunately, this approach could not be used for the study of the A28C suppressor since the compensatory mutation would locate within the SD sequence. Thus, we tested its SD accessibility by performing an RNase H/oligonucleotide accessibility assay (Figure 4.2.5B, see gels in Figure 8.1.5 and nucleotide list in Table 8.1.4 in the Appendix). The A40T and A40T / A33T suppressors were included in the study for comparison with a translation inhibitor mutation located outside the SD sequence. Reduced oligonucleotide accessibility was observed on A28C and A40T *aapA3*-Tr RNAs when compared to WT, demonstrating that, the A28C mutation as the *in vivo* validated A40T, inhibit toxin expression by reducing SD accessibility to ribosomes. The double mutant A33T / A40T showed WT accessibility. These results demonstrate that the suppressor A40T, despite being located on the SD sequence, acts by stabilizing SD sequestration by the addition of one base pair in a pre-existing hairpin

containing an anti-SD sequence. Thus, it acts at the mRNA structure level instead of at the sequence level.

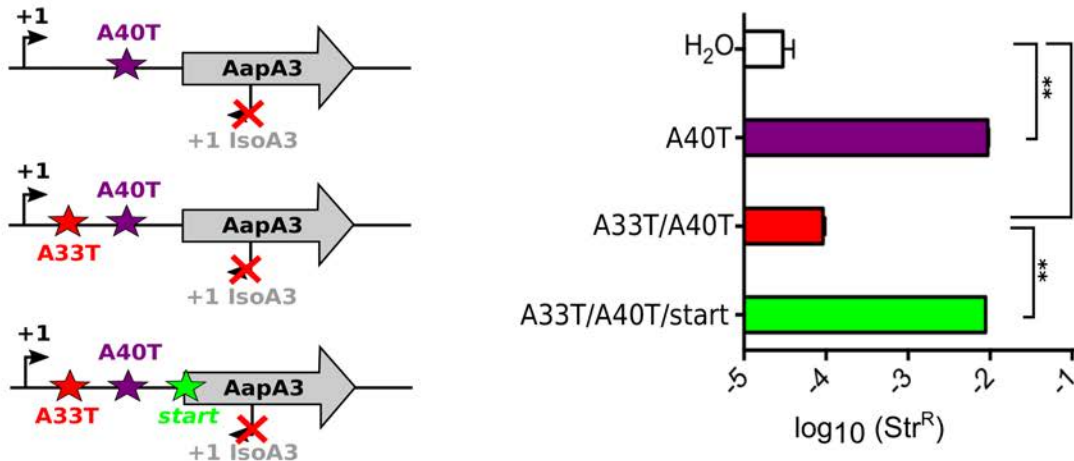
**A**



**B**



**C**



**Figure 4.2.5. Single nucleotide substitutions in the AapA3 5'UTR stabilize SD occlusion by an upstream to SD aSD sequence. (A)** Secondary structure prediction of the +28 to +45 region of the AapA3 mRNA. Shine-Dalgarno (SD) sequence is highlighted in turquoise. A CU-rich sequence element located upstream and able to mask the SD sequence creating an RNA hairpin is highlighted in yellow (anti-SD1, aSD1). The A28C and A40T suppressor mutations stabilizing this SD-occluding hairpin are shown in orange and purple, respectively. The A40T compensatory mutation A33T is shown in red. RNAfold (Gruber et al., 2008) was used for secondary structure and energy predictions, VARNA (Darty et al., 2009) was used for drawing. **(B)** Secondary structure prediction of the +1 to +45 region of the AapA3 mRNA. A black arrow represents the binding site of the DNA oligonucleotide used in the RNase H1 assays (left



panel). 30 fmol of internally labeled aapA3-Tr RNA (wild type or containing the indicated mutations) were incubated in presence of 0 to 10  $\mu$ M DNA oligonucleotide (FA644 for WT and A40T; FA651 for A33T/A40T; and FA652 for A28C, see Table S4) and subjected or not to digestion by *E. coli* RNase H1 (right panel). Digestion products were analyzed in an 8% PAA denaturing gel. Substrate consumption was quantified as relative band intensity (peak area). Error bars represent the s.d; n=2 independent experiments. (C) PCR constructs used to assess the SD-sequestering structure involving the A40T mutation by transformation assay are shown. (D) Number of StrR upon transformation with the indicated PCR constructs was calculated. Error bars represent s.d; n=3 biological replicates. (\*\*P<0.0001; \*P=0.001 according to unpaired t-test).

#### 4.2.1.5. A second metastable SD-sequester hairpin is embedded in the *aapA3* ORF

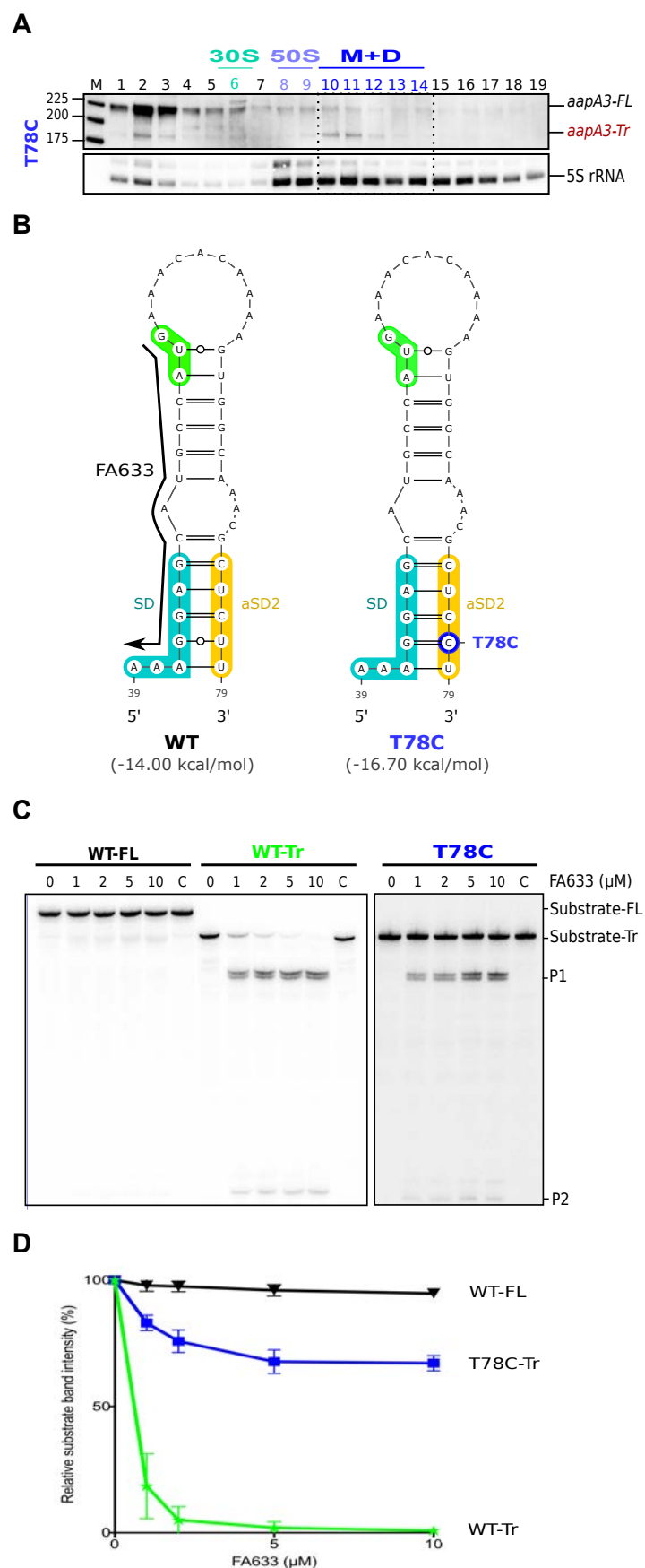
A synonymous mutation converting the AapA3 serine codon at position 9 from TCT to TCC (T to C transition at position +78 from *aapA3* +1, T78C in all figures) was selected in our suppressor screen. Located at 33 nt distance from the SD and apparently far from any other mRNA regulatory element, the existence of this mutation was intriguing. We generated the strain for further study. Northern Blot analysis showed that the T78C strain has wild-type expression and mRNA stability patterns (Figure 8.1.4B in the Appendix). We next tested its translatability *in vivo* by performing polysome fractionation coupled to Northern Blot analysis. A reduced AapA3 translatability was observed (Figure 4.2.6A, lanes 10 to 14) when compared to the positive control strain (Figure 4.2.4B, T107C panel, lanes 10 to 14). However, with 34.3% of the active toxin mRNA located in the monosome and disome fractions, the T78C mutation showed the smaller translation inhibitory effect when compared to the two previously studied suppressors (A28C and A40T, Figure 4.2.4B) (see Figure 8.1.7 in the Appendix for band intensity quantification). *In vitro* translation assays confirmed these results (Figure 8.1.4C in the Appendix) demonstrating that the T78C suppressor acts by inhibiting AapA3 translation. Interestingly, these results also indicate that a relatively high amount of peptide synthesis is tolerated by the cells, suggesting that toxicity is strongly dependent on the concentration of peptide in the cell, which could be diluted during cell division.

We next performed RNase H/ oligonucleotide protection assays to address SD accessibility on the T78C RNA. In this case, RNAfold (Gruber et al., 2008) secondary structure prediction revealed the putative stabilization of a second metastable SD-sequestering hairpin involving a second aSD (aSD2) sequence embedded on AapA3 ORF, and on which the suppressor T78C is located (Figure 4.2.6B). In order to reveal the accessibility to this region, a new oligonucleotide targeting the predicted stabilized hairpin was designed (FA633, see Table 8.1.4 in the Appendix). Remarkably, a strong reduction in SD accessibility was observed in the T78C RNA when compared to WT ( $\approx$ 58X more non-digested substrate with the maximal oligonucleotide concentration tested, Figure 8.1.5 in the Appendix). Notably, this single-nucleotide transition converts a wobble base-pair G-U into a G-C pair, allowing survival by addition of one hydrogen bond. Importantly, despite being located on AapA3 coding region, the T78C suppressor acts at the mRNA folding level.

Sequence conservation analysis of the AapA3 coding-region in 49 *H. pylori* strains (Figure 8.1.6 in the Appendix) revealed that the serine residue at position 9 (where T78C is located) is one of the most highly conserved residues in the peptide. Among all the analyzed strains, only the UM066 strain (highlighted in pink in Figure 8.1.6 in the Appendix) possesses a mutation on this codon. This mutation converts the TCT (Ser) codon into CCT (Pro) and probably abolishes peptide toxicity by disrupting the alpha-helix structure. Nevertheless, a possible conservation for an anti-SD strengthening sequence enhancing SD sequestration on an alternative secondary structure cannot be discarded. Likewise, the nucleotides corresponding to -10 box of IsoA3 small RNA (-10 IsoA3, in Figure 8.1.6 in the Appendix) are among the most highly conserved ones. This reflects the high selection pressure the

system is subjected to for avoiding premature toxin expression during transcription, and the essentiality of IsoA3 antitoxin for survival, respectively. Additionally, it is interesting to note that thirteen UC-rich motifs with potential anti-SD activity are embedded in the *aapA3* mRNA (Figure 8.1.8 in the Appendix). However, our results have clearly demonstrated that, unexpectedly, only three of them (two co- and one post-transcriptionally) are functional.

Altogether, we demonstrate that AapA3 translation can be inhibited by a single-nucleotide substitution located not only in the 5' UTR as described above but also in the coding region. Through the study of these translation inhibitor suppressors we uncovered the existence of two metastable (MeSt) SD-sequestering hairpins embedded in the AapA3 mRNA (MeSt1, involving the aSD1 sequence; and MeSt2, involving the aSD2 sequence, Figure 4.2.7) and highlighted the importance of such co-transcriptional folding for survival as it prevents premature toxin expression during transcription.



**Figure 4.2.6. A synonymous single-point mutation within a downstream to SD aSD sequence motif stabilizes SD occlusion and hampers AapA3 translation. (A)** Cell lysate of the T78C (thymine at position +78 from the *aapA3* TSS mutated to cytosine) strain was subjected to ultracentrifugation through a

sucrose gradient under polysome stabilization conditions (+ Chloramphenicol). A profile at OD<sub>254</sub> was recorded. Peaks of the free 30S and 50S subunits, 70S ribosomes (free and translating) and polysomes are indicated. RNA was extracted from each fraction and equal volumes of each extract were subjected to Northern Blot analysis. The same membrane was successively probed for the *aapA3*, *IsoA3*, and 5S rRNA transcripts. The different transcripts are annotated as: *aapA3*-FL, *aapA3*-Tr, and 5S rRNA (loading control). M, monosomes; D, disomes. **(B)** RNAfold (Gruber et al., 2008) was used for secondary structure and energy predictions and VARNA (Darty et al., 2009) for drawing. A black arrow represents the binding site of the DNA oligonucleotide used for the RNase H1 assays. The T78C mutation is shown in blue; Shine-Dalgarno sequence (SD) is shown in turquoise; anti-Shine-Dalgarno motifs (from the 5' end, aSD1 and aSD2) are shown in yellow; start codon is shown in green. **(C)** 30 fmol of internally labeled *aapA3*-FL and *aapA3*-Tr RNA (WT or T78C) were incubated in presence of 0 to 10  $\mu$ M DNA oligonucleotide (FA633) and subjected or not to digestion by *E. coli* RNase H1. **(D)** Substrate consumption was quantified as the relative band intensity (peak area). Error bars represent the s.d; n=2 independent experiments.

---

## 4.2.2. Discussion

How bacteria modulate gene expression via RNA structure has been a fascinating topic for the last 30 years. This post-transcriptional regulation is often achieved at the translation initiation step through the sequestration of the SD sequence in stable RNA hairpins that prevent ribosome binding to the translation initiation region (TIR) of the mRNA (Duval et al., 2013, 2015; Meyer, 2017). Although translation efficiency strongly correlates with the mRNA structure around the TIR (Mustoe et al., 2018), little is known about the influence of co-transcriptional folding. Bacteria could, in principle, reduce or delay the translation of a specific mRNA by playing with its secondary structure while the mRNA is being made (Lai et al., 2013; Zhu and Meyer, 2015). In this article, we identified two functional RNA hairpins within a type I toxin-encoding mRNA for which a tight control of translation is essential. The study of such RNA structures reflected the importance of co-transcriptional folding and the existence metastable RNA hairpins able to sequentially occlude SD accessibility during mRNA synthesis.

### 4.2.2.1. FASTBAC-Seq uses lethality to identify suppressor mutations

The initial aim of this work was to study the chromosomal expression of the AapA3 small toxic protein, which belongs to a new family of type I TA system recently discovered on the chromosome of the major human gastric pathogen *H. pylori* (Arnion et al., 2017). To date, most of TA systems studies, including our previous work (Arnion et al., 2017), made use of artificial expression systems to characterize the effects of toxin expression. The use of such overexpression vectors is often a source of misinterpretations, as toxin proteins would rarely be found at such high concentrations under endogenous conditions. To detect the AapA3 toxin expression at the chromosomal level, we inactivated the antitoxin promoter by introducing of two synonymous mutations, as previously described (Arnion et al., 2017). However, despite several attempts, we were unable to obtain a viable strain without having suppressor mutations in the toxin-encoding gene. Suppressor mutations have been previously reported in *B. subtilis* when two chromosomally-encoded type I toxins with killer-activity were expressed (the *txpA*/RatA (Silvaggi et al., 2005) and the *bsrG*/SR4 (Jahn et al., 2012)). But strikingly, such lethality has never been observed before in chromosomally-encoded TAs in Gram-negative bacteria. Indeed, the killer activity observed in the plasmid-encoded *hok*/Sok TA system was even believed not to be conserved in the chromosomally-encoded homologs (Pedersen and Gerdes, 1999). In the K12 reference strain, most of these homologs (*hokA*, *C* and *E*) have been inactivated by the insertion elements IS150 or IS186, located close to the toxin-encoding reading frames (Pedersen and Gerdes, 1999). These observations, together with studies reporting on the

differential expression of several TA systems in response to various stresses (*e.g.*, temperature shift, oxidative stress, starvation), suggested that chromosomally-encoded TAs may not be involved in a killer activity, but rather, in a reversible growth arrest in response to a specific stress. On the contrary, our results clearly demonstrated that the chromosomal expression of the AapA3 toxin is constitutive and lethal in the absence of the IsoA3 antitoxin. Consequently, we took advantage of this lethality, and coupled the genetic selection of suppressors to Next-Generation sequencing (FASTBAC-Seq, see [section 4.1 in Chapter 4](#)), allowing the fast identification of hundreds of intragenic suppressors mutations with nucleotide resolution ([Masachis et al., 2018](#)).

#### 4.2.2.2. A single nucleotide substitution is sufficient to abolish toxin translation

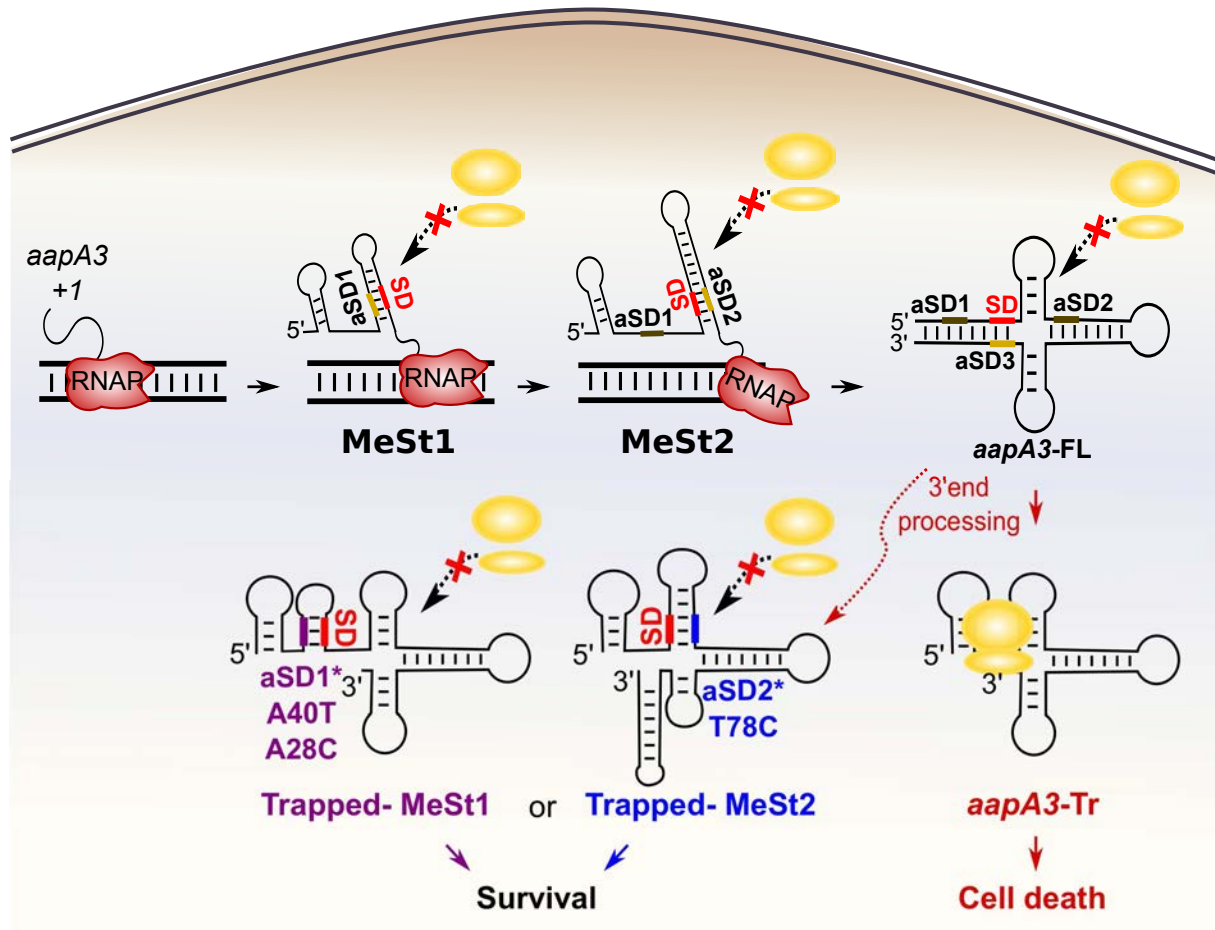
FASTBAC-Seq revealed a wide range of unanticipated *cis*-encoded toxicity determinants, affecting the toxin protein, or its expression. Among the mutations affecting the toxin mRNA expression, we identified five single nucleotide substitutions able to inhibit the AapA3 translation without affecting its stability. Three of such mutations were located in the SD sequence. Two of them, were highly enriched, substituting the guanines at positions 43 and 44 from the AapA3 TSS (G43 and G44) by either an adenine, a cytosine, or a uridine. This revealed the 5' AGG 3' and the 5' GGA 3' as the minimal functional SD sequences allowing AapA3 translation, in agreement with the previously identified *H. pylori* SD consensus sequence (5' AAGGA 3') ([Sharma et al., 2010](#)). The third mutation (A40T), was much less enriched, and remarkably, at that position, only the transversion mutation adenine to thymine was selected. Another transversion mutation was selected 14 nucleotides upstream the SD sequence. The fact that only transversion mutations were selected at these two positions indicated that, in this case, the nature of the substituted nucleotide was important, and the effect of these mutations may not directly play at the sequence level, but rather at the structure level. Our results demonstrated that these mutations create, one (for A28C) or two (for A40T), additional base pair(s) within a stem-loop structure formed between the SD sequence and an upstream complementary anti-SD (aSD) sequence (aSD1, 5' UCCU 3'). Thus, the A40T mutation, despite being located within the SD sequence, remarkably acts at the mRNA structural level and not at the sequence level.

Interestingly, the T78C mutation revealed the existence of a second aSD sequence (aSD2) located downstream the SD sequence, within the toxin coding region. This synonymous substitution (UCU→UCC, Ser codon at position 9, Ser9), creates a perfect aSD sequence (5' CUCCU 3'). Despite having no impact in the protein sequence, synonymous mutations can influence gene expression ([Kudla et al., 2009](#)). The stability of the mRNA folding close to the TIR can explain more than half of the variation in protein levels. Indeed, a strong codon bias to avoid mRNA structure close to the TIR region is often observed within the first 15 codons ([Bentele et al., 2014](#); [Bhattacharyya et al., 2018](#)). A similar aSD sequence (5' UCCU 3') has been observed in the coding sequence of the *gnd* gene in *E. coli* ([Carter-Muenchau and Wolf, 1989](#)). However, it is encoded far downstream (codon 66-71), having only a moderate effect on translation that is highly dependent on ribosome concentration. Interestingly, when this aSD sequence is placed at the codon 13, it is able to fully inhibit translation. In line with this work, we show that despite the presence of up to thirteen CU-rich sequences in the AapA3 mRNA, only three of them are able to act as aSD sequences, reflecting a strong context-dependent functionality of these translation regulatory elements.

#### 4.2.2.4. Suppressor mutations reveal functional metastable structures acting co-transcriptionally to impede premature toxin translation

The three here-studied mutations (A78C, A40T, and T78C) act within a pre-existing sequence context. Indeed, they locate within and stabilize, SD-sequester hairpin structures that were previously predicted to be formed in the co-transcriptional folding pathway of several AapA mRNAs ([Arnion et al., 2017](#)).

Importantly, despite stabilizing such structures, they do not interfere with the folding pathway of the full-length mRNA, neither affecting its transcription, stability, or 3' end maturation, and indicating that they exclusively act on the active AapA3 mRNA form (Figure 4.2.7). The selection of such stabilizing mutations, suggests that the thermodynamic stability of such SD-sequestering stem-loops in the wild-type context is not sufficient as to avoid the translation of the active AapA3 mRNA form (Figure 4.2.7).

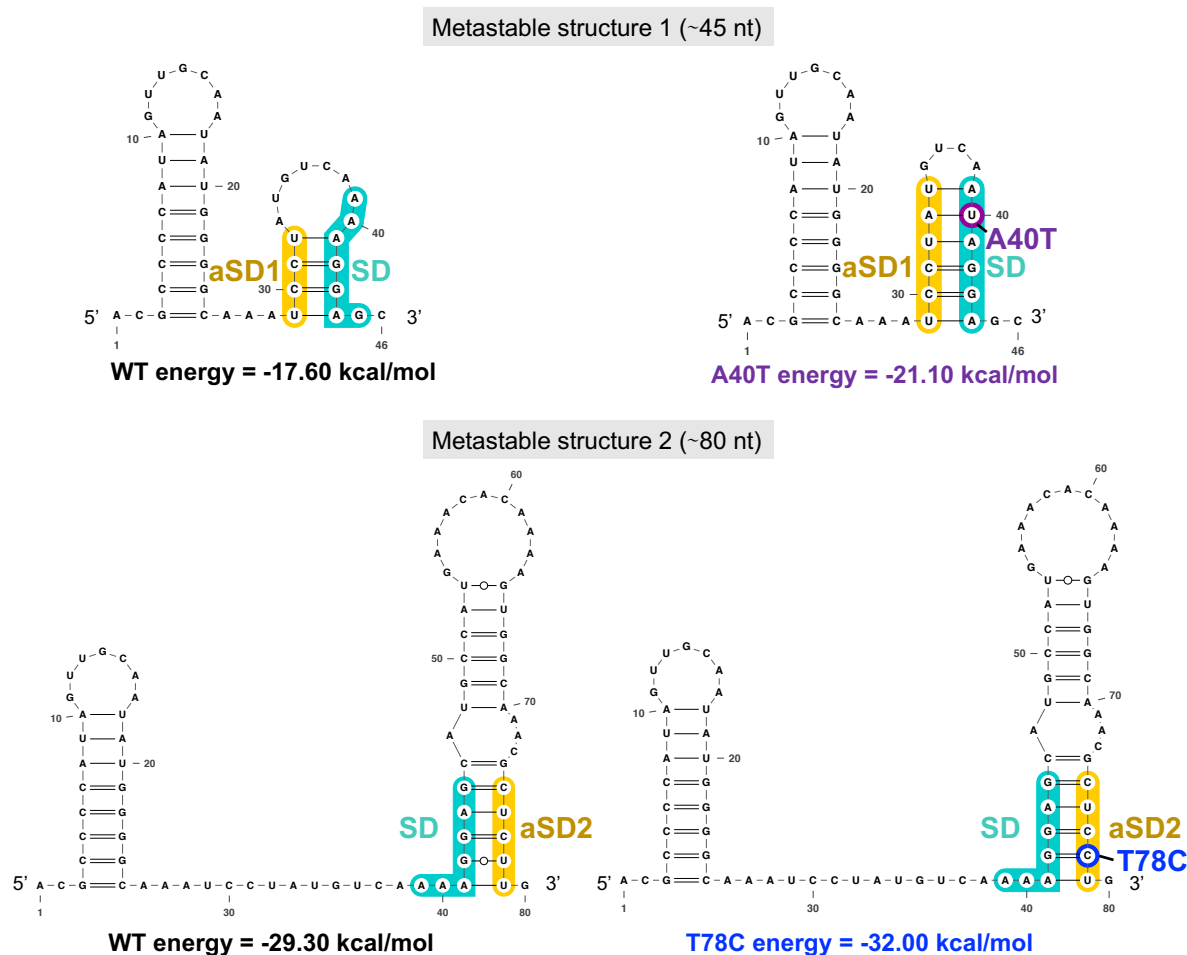


**Figure 4.2.7. Working model of the *aapA3* co- and post-transcriptional translational SD-accessibility regulation.** Co-transcriptional folding of the AapA3 mRNA leads to the sequential generation of two transient RNA hairpins (metastable structures 1 and 2, MeSt1 and MeSt2) that mask the Shine-Dalgarno (SD) sequence using CU-rich sequence elements called anti-SD sequences (aSD, aSD1 and aSD2, successively), temporally impeding ribosome accessibility. Upon transcription, the *aapA3* full-length transcript (*aapA3*-FL) is kept translationally inert through the SD occlusion by a third aSD sequence (aSD3) mediated by a 5'-3' long-distance interaction between both mRNA ends. Only upon a 3'-end nucleolytic truncation that leads to the loss of the aSD3, the toxin mRNA becomes translationally competent (*aapA3*-Tr) and can potentially lead to cell death if translated. Our results show that, in the absence of antitoxin, a single nucleotide substitution can hamper the translation of this active toxin mRNA form (*aapA3*-Tr) by stabilizing one of the two SD-occluding RNA hairpins predicted to be co-transcriptionally formed (MeSt1 using the aSD1 in the case of the A28C and the A40T mutants; MeSt2 using the aSD2 in the case of the T78C mutant). The selection of these hairpin-stabilizing suppressor mutations reflects on the existence of the two-above-mentioned functional metastable structures.

Instead, our results suggest that in the wild-type situation, such SD-sequestering structures are transiently formed to co-transcriptionally impede the premature toxin translation (metastable structures 1 and 2, MeSt1 and 2, Figure 4.2.7). This transient character is essential to ensure the proper



transcription termination and folding of the full-length mRNA, and it is achieved by hierarchically-increasing thermodynamic stabilities (see energy of wild-type metastable structures 1 and 2, [Figure 4.2.8](#)). Importantly, the suppressor mutations do not provide enough stabilization as to impede the formation of the next most stable structure (*i.e.*, the A40T mutated metastable structure 1 has an energy of -20.90 kcal/mol, while the metastable structure 2 has an energy of -29.30 kcal/mol, [Figure 4.2.8](#)). This may explain why the SD-sequester mutations are able to impede the translation of the active mRNA form without interfering with its co-transcriptional folding pathway.

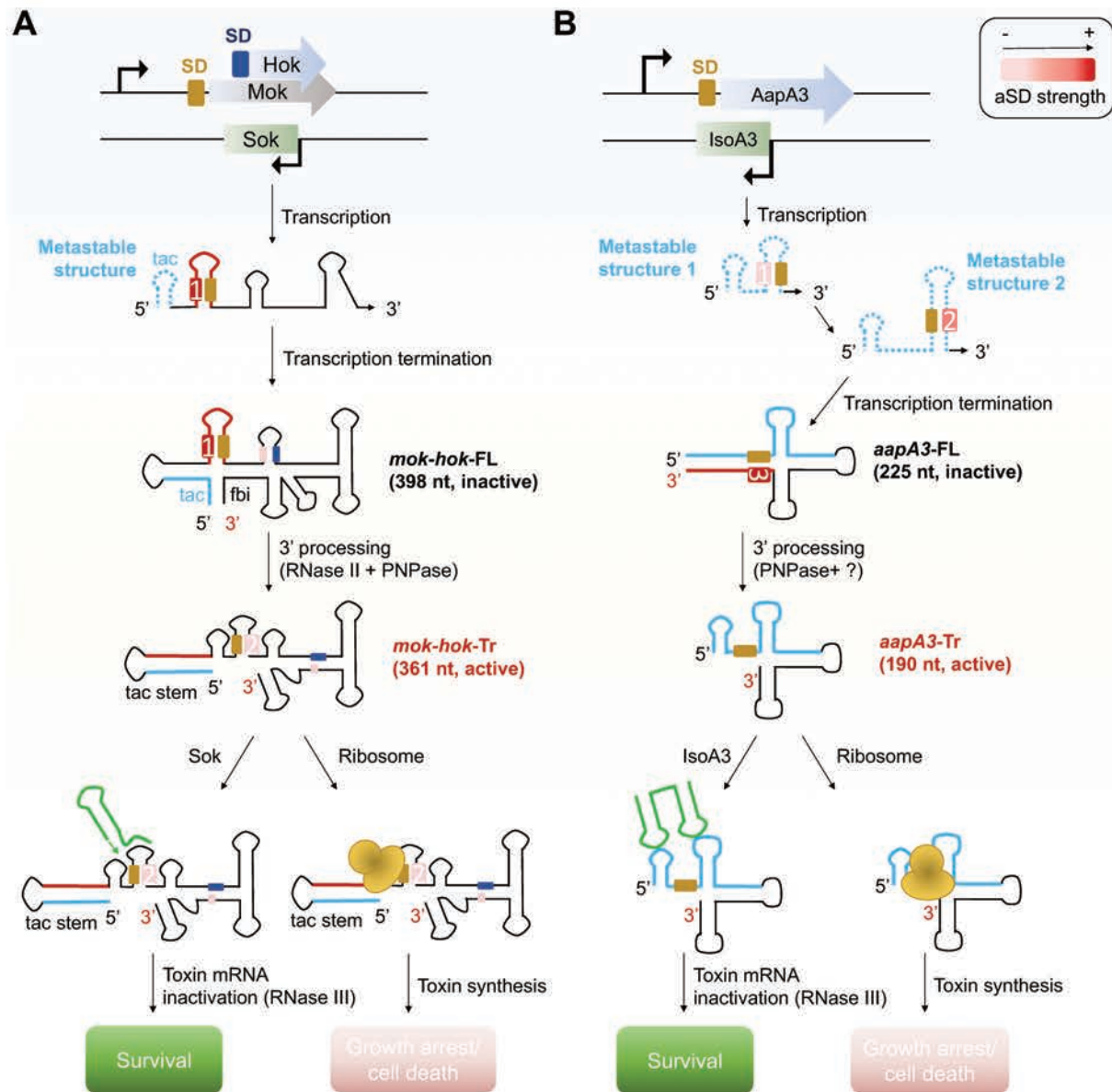


**Figure 4.2.8. The two successive AapA3 mRNA metastable structures have increasing stability and are stabilized by the A40T and the T78C suppressors.** The two metastable structures (1, ~45 nt long, upper panel; and 2, ~80 nt long, lower panel) successively formed during the AapA3 mRNA transcription are shown. Shine-Dalgarno (SD) sequence is highlighted in turquoise, anti-SD (aSD1 and aSD2) are highlighted in yellow, suppressor mutation stabilizing SD sequestration by the aSD1 in the metastable structure 1 is highlighted in purple (A40T), the suppressor stabilizing the SD sequestration by the aSD2 in the metastable structure 2 is highlighted in blue (T78C). RNAfold ([Gruber et al., 2008](#)) was used for secondary structure and minimum free energy predictions, and VARNA RNA ([Darty et al., 2009](#)) for drawing.

Another proof of the existence of such metastable RNA structures in the AapA3 mRNA is the strict conservation of the UCU Serine codon at position 9. As our results have shown, a UCC codon at that position (T78C mutation) would have rendered the AapA3 toxin expression inefficient, what would have rapidly led to the TA locus sequence degeneration and loss. The formation of metastable RNA structures has been reported in several RNA-mediated regulatory pathways, including the viral RNA



replication (Repsilber et al., 1999), RNA catalysis (Pan and Woodson, 1998), RNA editing (Linnstaedt et al., 2006), and ribosome biogenesis (Sharma et al., 2018). They are usually described as folding intermediates that work in a hierarchal manner to help an RNA molecule reach its functional conformation (*i.e.*, most thermodynamically stable conformation). A pioneering work on the *hok/Sok* type I TA system in *E. coli* described the existence of a metastable hairpin in the folding pathway of the Hok mRNA (Møller-Jensen et al., 2001; Nagel et al., 1999). Importantly, although the metastable hairpin described for the Hok mRNA was made to lock the growing mRNA into an inactive conformation, this structure was not directly involved in the sequestration of the SD sequence (see Figure 4.2.9A).



**Figure 4.2.9. Comparison of the co- and post-transcriptional translational regulation of the *hok/Sok* and *aapA3/IsoA3* type I TA systems.** The co- and post-transcriptional events controlling the Hok (A) and the AapA3 (B) toxin expression are illustrated. The genetic organization of the systems is shown at the top. Metastable structures formed during transcription are represented by dotted-lines. Upon transcription, the sequences that formed part of metastable structures are shown as blue lines. Shine-Dalgarno (SD) sequences are represented by yellow boxes. In the case of *hok/Sok*, the only the SD of the leader peptide (Mok) is shown in yellow. The SD of Hok is shown in dark blue. Anti-SD (aSD) sequences are represented

by boxes with different red intensities depending on the strength of the sequence to mask the SD, as indicated at the top right. aSD sequences sequentially masking the SD at the different stages of the regulatory pathway are represented by the numbers 1, 2 or 3, written within the aSD boxes. The translationally inert (full-length, -FL) and active (truncated, -Tr) toxin mRNA forms are indicated. The sequence context involved in the SD sequestration by the most stable aSD interaction is represented by a red line. Sok and IsoA3 sRNA antitoxins are shown in green. Ribosomes are represented in yellow. In the *hok*/Sok regulation, *tac* (translational activator), *fbi* ('fold-back-inhibition' element).

---

In our work, we provide *in vivo* evidences for the existence of functional metastable hairpins that impede the co-transcriptional toxin translation. This temporal control of gene expression is achieved through the sequential formation of two functional transient RNA structures that mask the SD sequence via CU-rich elements and that are replaced by a more stable one involving a long-distance-interaction between both mRNA ends in the full-length mRNA (Figure 4.2.9B). Similar to the *hok* mRNA, this final mRNA structure is so stable that its translational activation requires the removal of a third aSD sequence (aSD3, located within the 3' end of the mRNA), by an exonucleolytic activity (see section 4.3 of Chapter 4). The highly stable structure of the *aapA3*-FL mRNA is also similar to the cloverleaf-like structure found in the 5' UTR of the maturation (A) gene of the MS2 RNA coliphage (Groeneveld et al., 1995). Interestingly, it has been shown that in the case of the maturation (A) gene, it may take up to several minutes for the mRNA to be synthesized and properly folded (van Meerten et al., 2001), explaining the existence of functional transient structural intermediates preventing premature gene expression.

Overall, in the case of the Hok mRNA, the first aSD sequence masking the SD is the most stable one (represented by an intense-red box named 1, Figure 4.2.9A). On the contrary, in the case of the Aapa3 toxin mRNA, three increasingly-stable aSD sequences (boxes termed 1, 2 and 3, in increasingly-intense red color, Figure 4.2.9B) transiently occlude ribosome binding until reaching the most stable conformation (full-length) on which the SD is masked by the aSD3 (represented by an intense-red box named 3, Figure 4.2.9B). This is a crucial difference in the co- and post-transcriptional regulation of these systems, as mechanistically, the tightness of the co-transcriptional SD sequestration, as well as the dynamic time-window of gene expression, are drastically different (which may have functional consequences, e.g., ability or no of ribosome concentration sensing and regulation by sRNA binding). Our study represents the first *in vivo* evidence of the existence of sequential metastable structures, directly but transiently, avoiding the co-transcriptional translation of a toxin-encoding mRNA.

#### 4.2.2.5. Conclusion

To conclude, our work revealed the existence of CU-rich sequence elements that act co-transcriptionally to either transiently or stably occlude ribosome binding to the toxin-encoding mRNA. Due to the co-transcriptional translation that occurs in bacteria, these RNA hairpin structures are essential to uncouple the transcription and translation processes, and allow the presence of type I toxin antitoxin systems on bacterial chromosomes. To date, only four families of functional transient RNA structures have been described, including the Trp operon leader, the 5' UTR of the Levivirus (MS2 phage), the HDV ribozyme, and the SAM riboswitch (Zhu and Meyer, 2015). Our work is closer to the previous studies on the 5' UTR of the MS2 phage (van Meerten et al., 2001), except that in that case, the transient structure is formed to allow translation initiation to occur before the cloverleaf-like structure is formed. In both cases, a transient RNA structure is formed to exert a temporal control of translation, either negatively or positively. Although transient RNA structures can be predicted *in silico* (Meyer, 2017), their *in vivo* visualization remains challenging. Several methods have been recently reported to analyze the co-transcriptional folding of regulatory RNA, both *in vitro* (Uhm et

al., 2018; Watters et al., 2016) and *in vivo* (Incarnato et al., 2017). It would be really interesting to use these complementary techniques to analyze the formation of these metastable hairpins in real-time.

### 4.3. Paper IV: When mRNA folding rules decay: lessons from a type I toxin-antitoxin system

In the previous section (Paper III), we described how the AapA3 toxin-encoding mRNA is silenced via RNA-based regulatory mechanisms throughout its synthesis. How untranslated RNAs avoid degradation (and Rho-dependent termination) is a major question in the prokaryotic field that arises from the observation that many mRNAs, including type I toxin-encoding mRNAs (Masachis and Darfeuille, 2018), escape to transcription/translation coupling. In this paper, we aimed at exploring the mRNA stability determinants of the AapA3 toxin-encoding mRNA, and at understanding how they influence its activation pathway and the regulation of the system.

The stability of a specific RNA is strongly dependent on its secondary structure, as most exonucleases (and some endonucleases) are, to a lesser or greater extent, inhibited by them. For instance, it has been shown that PNPase, the major 3'-5'-exonuclease present in *H. pylori*, degrades the 3' end of RNA until encountering an energetically stable enough stem-loop structure (shown in *E. coli*, Spickler and Mackie, 2000). This property has even been exploited in some studies for the *in vitro* secondary structure analysis of RNAs (Chen et al., 1991). Similarly, a strong sensitivity to secondary structure has been shown for the only so far identified *H. pylori* 5'-3'-exonuclease, the RNase J (in *H. pylori* (Redko et al., 2013) and in *E. coli*, Even et al., 2005)), whose activity has also been used for RNA structure *in vitro* studies (Daou-Chabo and Condon, 2009). The key role of 5' and 3' end stem-loop structures in mRNA decay was discovered long ago (Chen et al., 1991; Emory et al., 1992; Mott et al., 1985). With the recent development of the Term-Seq approach (Dar et al., 2016), a massive existence of such protective hairpins in the *E. coli* transcriptome has been reported (Dar and Sorek, 2018).

FASTBAC-Seq revealed several suppressor mutations able to trigger the degradation of the AapA3 toxin mRNA. Such mutations locate within highly-stable hairpins that are present at both ends of the 3' end processed species. Remarkably, these mutations lead to a strong destabilization of the hairpins by disrupting only one base pair in their stem. Our work demonstrated that the stem-loop present at the 3' end functions as a thermodynamic roadblock, impeding the 3'-exonuclease activity of the PNPase. Furthermore, this protective hairpin is essential for the proper maturation of the AapA3 mRNA, as its translational activation is achieved through a 3'-exonucleolytic processing pathway led by the PNPase. On the other hand, the hairpin present at the 5' end blocks 5'-processing activities, such as RNase J. Importantly, both termini protective hairpins act in coordination, and the strict presence of both of them is essential to ensure that the AapA3 mRNA is stable enough as to be efficiently translated (*i.e.*, functional stability). Overall, this work highlights the key role of secondary mRNA structure for the stability and nucleolytic activation of the AapA3 toxin-encoding mRNA.

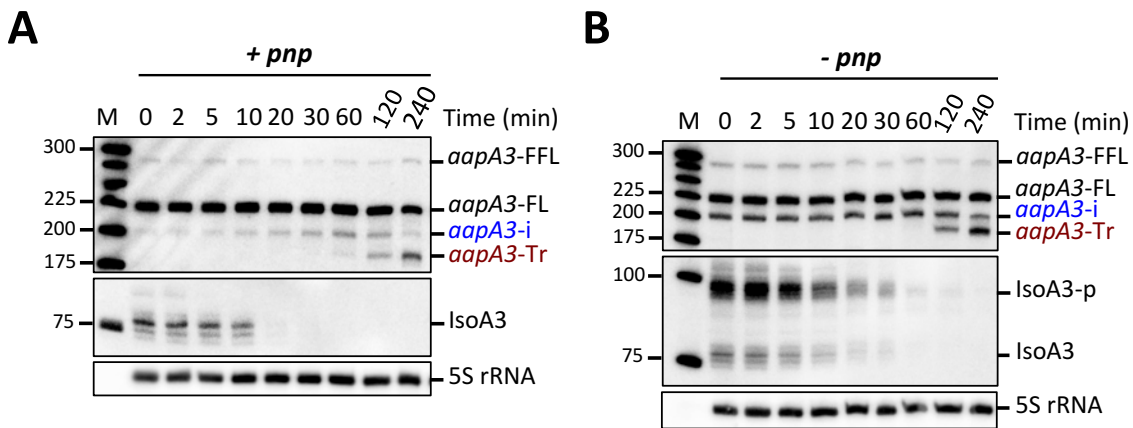


### 4.3.1. Results

#### 4.3.1.1. Ribonucleolytic processing activates the *aapA3* mRNA

We recently reported on the existence of a truncated AapA3 mRNA species that accumulates in the absence of antitoxin as well as in absence of RNase III and is the only translationally active AapA3 mRNA form (see Chapter 4 section 4.2). Based on RNA-Seq data (Sharma et al., 2010), by reminiscence to the AapA1 mRNA (Arnion et al., 2017), and considering that the RNA antitoxin targets the 5' end of the toxin mRNA, we hypothesized that this active mRNA species is the product of a 3' end ribonucleolytic activation process. However, the molecular underpinnings of this activation were unknown.

In the present study, we first aimed to confirm that the observed truncated AapA3 mRNA was the product of a nucleolytic activity and not that of transcription termination. To that aim, we performed RNA decay studies using rifampicin. Northern Blot analysis of the wild-type 26695 *H. pylori* strain revealed the presence of a transcript with longer size (*aapA3* full-full length, *aapA3*-FFL,  $\approx 280$  nt) than the so far considered to be the full-length (*aapA3*-FL,  $\approx 225$  nt) (Figure 4.3.1A). This *aapA3*-FFL species appears to be generated by transcription termination and be highly stable. The constant levels of *aapA3*-FFL mRNA indicate that the so far considered to be the AapA3 primary transcript and the most abundant AapA3 mRNA form found *in vivo*, the *aapA3* full-length (*aapA3*-FL,  $\approx 225$  nt), is generated by an independent transcription termination mechanism rather than by nucleolytic cleavage of the *aapA3*-FFL species. Furthermore, the lack of correlation between the levels of *aapA3*-FFL and *aapA3*-FL forms with those of the IsoA3 RNA (IsoA3,  $\approx 80$  nt), is in line with previous results and demonstrates that IsoA3 RNA targets none of them.



**Figure 4.3.1. Rifampicin treatment reveals an exonucleolytic activation pathway of the *aapA3* mRNA led by PNPase.** After rifampicin addition (at OD<sub>600</sub> = 1.7), aliquots of *H. pylori* 26695 WT (+ *pnp*) (A), or  $\Delta$ PNPase (- *pnp*) (B), cultures were collected at several time points (as indicated on top of the gel). RNA was extracted and subjected to Northern blot analysis. The membrane was successively probed with <sup>32</sup>P-FD38 labeled oligonucleotide and IsoA3 riboprobe to detect *aapA3* and IsoA3 transcripts, respectively. The different transcripts are annotated as: *aapA3*-FFL (full-full-length,  $\approx 280$  nt), *aapA3*-FL (full-length,  $\approx 225$  nt), *aapA3*-i (intermediate,  $\approx 200$  nt), *aapA3*-Tr (truncated,  $\approx 190$  nt), IsoA3-p ( $\approx 95$  nt) and IsoA3 ( $\approx 80$  nt). A labeled DNA marker (25 bp intervals, lane M) was used for size estimation. Proper loading was assessed by the level of 5S rRNA using the labeled oligo probe FD35.

Surprisingly, a band of an intermediate size between the *aapA3*-FL and the *aapA3*-Tr species was revealed (*aapA3*-i,  $\approx 200$  nt, Figure 4.3.1A). This intermediate form appears before the *aapA3*-Tr can be observed, accumulates in a time- and IsoA3-independent manner, and has a relatively long

half-life. This reflects a slow and IsoA3-independent decay, pointing out that *aapA3-i*, as the two longer species (*aapA3-FFL* and *aapA3-FL*), is inert for interaction with the IsoA3 RNA. Remarkably, the *aapA3-i* levels decrease when those of the *aapA3-Tr* increase, demonstrating that it is a nucleolytic intermediate on the *aapA3* mRNA activation pathway. Rifampicin assays demonstrated that the AapA3 truncated form (*aapA3-Tr*,  $\approx 190$  nt) is a ribonucleolytic product that accumulates only once the IsoA3 RNA pool has been fully degraded (Figure 4.3.1A). This result is in line with previous observations and demonstrates that *aapA3-Tr* is rapidly targeted by IsoA3 RNA and that the duplex is degraded by the double-strand-specific RNase III, as previously shown *in vivo* (see Chapter 4 section 4.2). Taken together, our results demonstrate that *aapA3* mRNA translational activation occurs by ribonucleolytic degradation of a long (and inert) primary transcript in an IsoA3-independent manner.

#### 4.3.1.2. AapA3 mRNA translational activation is led by the 3'-exonucleolytic activity of PNPase

To uncover the molecular underpinnings of the *aapA3* mRNA nucleolytic activation process, and given that *H. pylori* lacks RNase II, we tested AapA3 mRNA expression and decay in different genetic backgrounds deleted for one/or both of the two major *H. pylori* 3'-5' exonucleases, the ribonuclease R (RNase R, HP1248) (- *rnr*) and the polynucleotide phosphorylase (PNPase, HP1213) (- *pnp*). No changes in the relative abundance or length of the different AapA3 mRNA species were observed in the absence of RNase R (data not shown). However, an abundant steady-state presence of the *aapA3-i* species was observed in the *pnp* deleted background (Figure 4.3.1B). The absence of *aapA3-i*/IsoA3 linked-decay despite the high steady-state levels of *aapA3-i* in the *pnp* deleted background confirmed that the IsoA3 RNA antitoxin does not target the *aapA3-i* species. As previously observed (Figure 4.3.1A), the levels of *aapA3-i* decreased when those of *aapA3-Tr* increased (Figure 4.3.1B), demonstrating that PNPase is involved in the 3' end processing of the *aapA3-i* species to generate the active *aapA3-Tr* mRNA. However, the observation of the *aapA3-Tr* species despite PNPase absence reflects the existence of a so far unidentified redundant activity able to counteract the lack of PNPase.

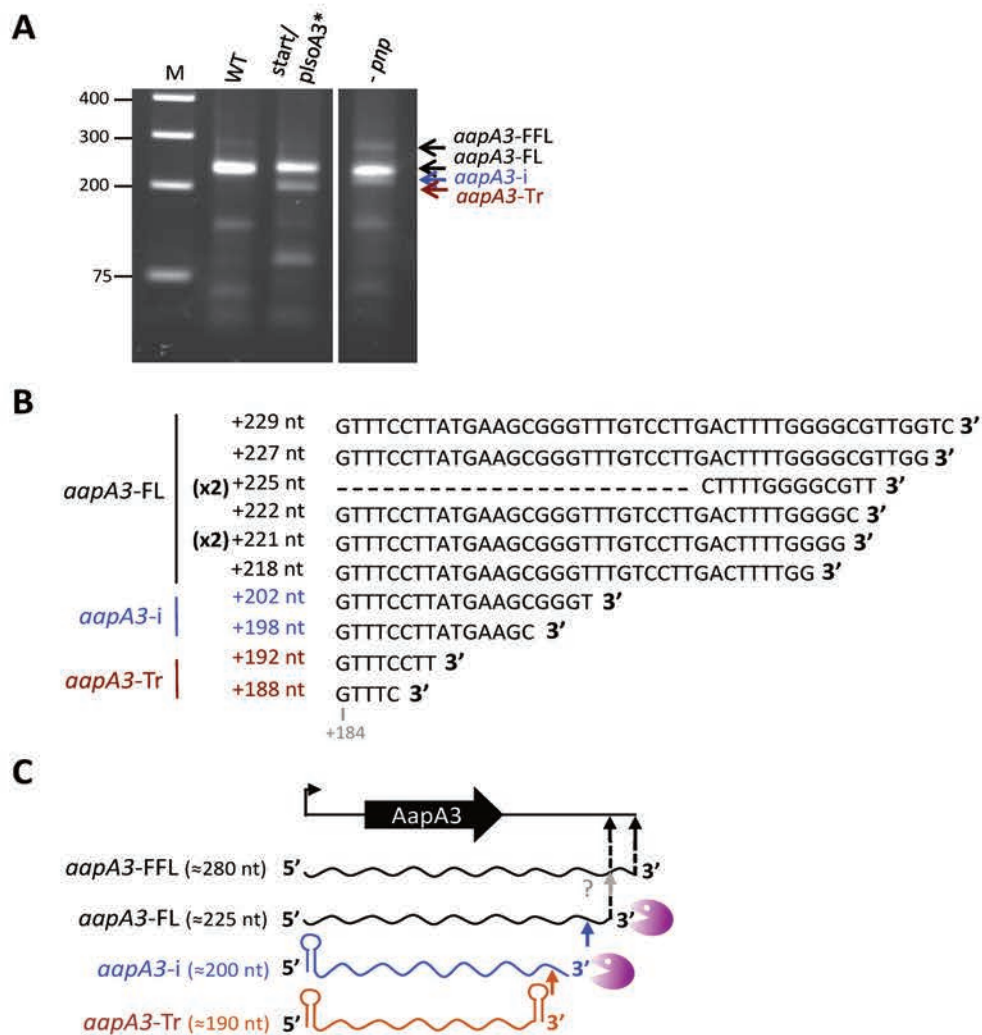
In order to identify the PNPase redundant activity responsible of the generation of the *aapA3-Tr* species in *H. pylori*, we generated a deletion strain of the endoribonuclease RNase Y (HP0760) ( $\Delta rny::aphA3$ , Table 8.2.3 in the Appendix) as well as a double deletion strain lacking both, PNPase and RNase Y activities ( $\Delta rny::aphA3/\Delta pnp::catCG$ , Table 8.2.3 in the Appendix). Rifampicin assays on the - *rny* strain revealed decreased levels of *aapA3-Tr*, and close to wild type *aapA3-i* levels (Figure 8.2.1A in the Appendix). When comparing the *aapA3* mRNA decay profile in the double - *rny*/- *pnp* mutant with that of the wild type strain (Figure 8.2.1B in the Appendix), a slight decreased on *aapA3-i* levels can still be observed, indicating that RNase Y is involved in the generation of *aapA3-i* from the *aapA3-FL* species. However, in the double - *rny*/- *pnp* mutant, the levels of *aapA3-Tr* were strongly reduced when compared to those of the simple - *pnp* mutant (Figure 8.2.1C in the Appendix). This strong reduction cannot be fully explained by the decreased *aapA3-i* levels, and indicates that RNase Y is also involved in the generation of the active *aapA3* mRNA species (*aapA3-Tr*) from the *aapA3-i* form. Nevertheless, despite the lack of RNase Y and PNPase activities, the *aapA3-Tr* species is still produced, therefore, the identity of the redundant activity remains enigmatic.

#### 4.3.1.3. 3' RACE reveals a heterogeneous population of AapA3 mRNA species

So far, we had estimated the size of the different AapA3 mRNAs based on RNA-Seq data (Sharma et al., 2010) and size markers in Northern Blot analyses. Protocols for the identification of 5' ends have been well established for several years (Argaman et al., 2001; Thomason et al., 2015; Urban and Vogel, 2007), however, the identification of 3' ends in bacterial mRNAs has proved to be more challenging due to the lack of poly-U tails in most mRNAs. Nevertheless, recently, Dar and colleagues



(2016) performed a genome-wide mapping of RNA 3' termini via term-seq (Dar and Sorek, 2018) and described an efficient Rapid Amplification of cDNA 3' Ends (3' RACE) protocol. To get a better understanding of the Aapa3 mRNA activation pathway, we used the 3' RACE protocol described by Dar et al., (2016) with slight modifications (see Methods section). Total RNA isolated from the *H. pylori* 26695 wild-type and - *pnp* strains as well as strain carrying a mutated *aapA3*/IsoA3 locus version on which the start codon of the toxin has been inactivated as well as the antitoxin promoter (*start*/pIsoA3\*) was used as template to perform cDNA amplification and PCR. As shown in Figure 4.3.2A, we were able to identify the four most abundant Aapa3 mRNA species that were previously observed in Northern Blot analysis of the specific strains, being the -FL form present in all strains, the -Tr form visible only in absence of antitoxin and the -i form accumulated only in the absence of PNPase. Importantly, this result brought the definitive proof that the Aapa3 mRNA undergoes a 3' end ribonucleolytic activation process and fully discards any kind of 5' end processing event on its activation pathway.



**Figure 4.3.2. 3' RACE reveals heterogeneous populations of the three main 3' end-processed *aapA3* mRNA species.** (A) Total RNA from the *H. pylori* 26695 WT, *start*/pIsoA3\* or  $\Delta$ PNPase (-*pnp*) strains was isolated and used to perform 3' RACE experiments (see Methods for details). 3' RACE PCR products were visualized in a 2% agarose gel stained with ethidium bromide. DNA bands corresponding to the previously observed by Northern Blot *aapA3* species are indicated. (B) Sequence alignment of the 3' RACE results on *aapA3* mRNA. 3' RACE PCR product from the *H. pylori* 26695 -*pnp* strain was cloned into pGEM-T *E. coli* vector and sequenced. mRNA lengths are shown relative to the already known 5' end

(*aapA3* +1). Heterogeneous sequence populations of the *aapA3* full-length (*aapA3*-FL), intermediate (*aapA3*-i) and truncated (*aapA3*-Tr) species are shown. (C) Schematic representation of the different *aapA3* mRNA species. The two longer forms (*aapA3*-FFL,  $\approx 280$  nt and *aapA3*-FL,  $\approx 225$  nt) generate from transcription termination. The generation of *aapA3*-FL from *aapA3*-FFL is unlikely but not discarded (grey arrow). 3' end processing led by the 3'-5'-exonuclease activity of the PNPase (represented by a purple Pac-man) successively generates the *aapA3*-i (intermediate,  $\approx 200$  nt) species with a 5' stem-loop, and the *aapA3*-Tr ( $\approx 190$  nt) with both a 5' and a 3' stem-loops.

Subsequently, PCR products were cloned and sequenced. As shown in Figure 4.3.2B, 3' RACE sequencing revealed highly heterogeneous populations of the three main AapA3 mRNA species (unfortunately, due to its low abundance in the PCR, the *aapA3*-FFL species got lost in the cloning process). The observation of such landscape of 3' ends is highly significant and reflects the mechanism of action of the actors in the AapA3 mRNA activation pathway (3'-exonucleases). If we now make RNA secondary structure predictions using the newly identified *aapA3*-FL ( $\approx 229$ -218 nt), *aapA3*-i ( $\approx 202$ -198 nt) and *aapA3*-Tr ( $\approx 192$ -188 nt) 3' ends, we can observe that the 3' ends are not arbitrary, but rather defined by energetic roadblocks for 3'-exonucleases (Figure 8.2.2 in the Appendix). Indeed, in the secondary structure of the last AapA3 mRNA activation product (*aapA3*-Tr), two highly stable stem-loop structures can be observed at both, the 5' (with five consecutive GC pairs) and the 3' (with three consecutive CG pairs) ends. The 5' stem-loop is expected to be co-transcriptionally formed to avoid the premature mRNA degradation by cellular 5'-3'-exonucleases and to re-appear after the cleavage of the first 3' end nucleotides in the full-length transcript (weakening the 5'-3' long-distance interaction). On the contrary, the 3' stem-loop is unlikely formed during transcription, as the 5'-3' long-distance interaction needs to be efficiently formed to avoid both, ribosomes and IsoA3 binding to the nascent mRNA. Thus, the 3' stem-loop is a unique feature of the final AapA3 mRNA activation state, the *aapA3*-Tr species (Figure 4.3.2C and Figure 8.2.2 in the Appendix).

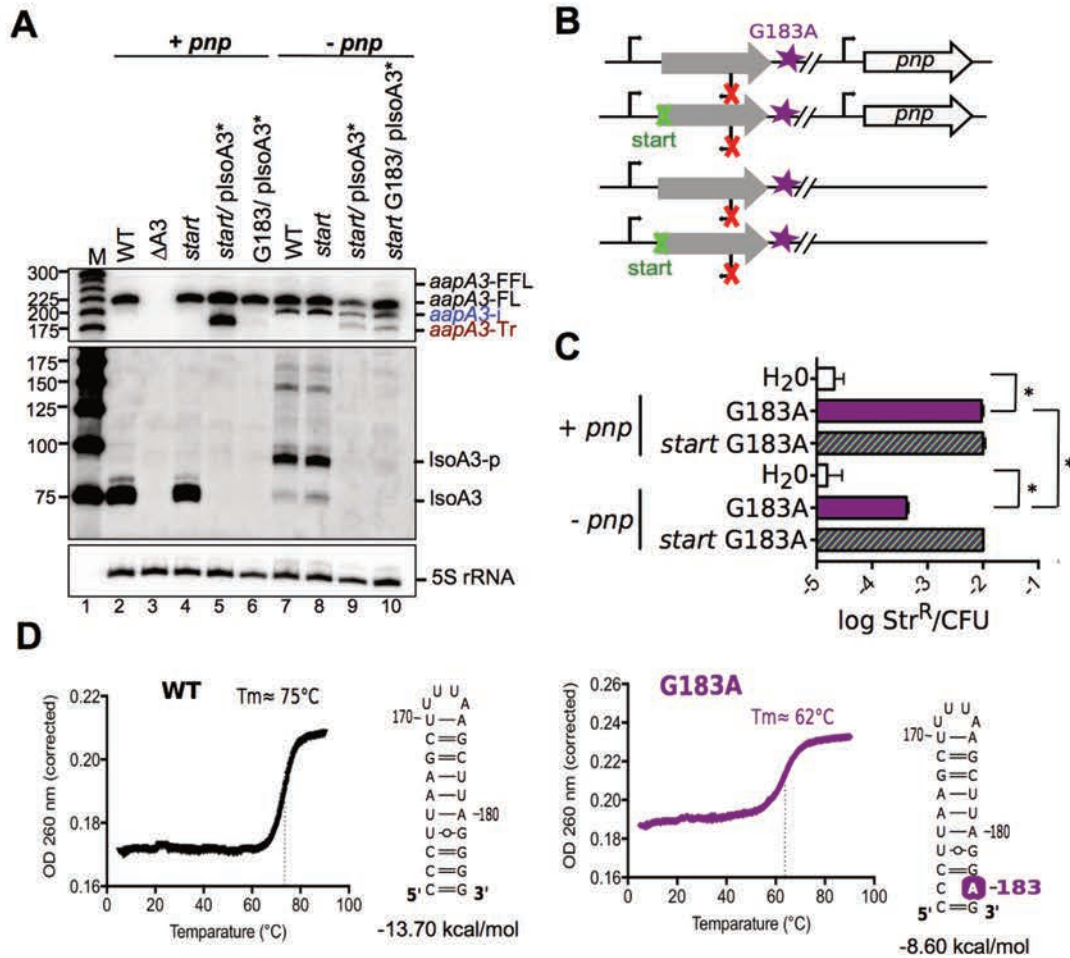
#### 4.3.1.4. The active AapA3 mRNA harbors a 3' end protective stem-loop structure

During the analysis of the toxicity suppressor mutations previously identified by FASTBAC-Seq (see Chapter 4 section 4.1), two groups of mutations struck us. Such mutations were located in the 5' and 3' non-coding regions (untranslated regions, UTRs) of the AapA3 mRNA laying outside any known regulatory element (*e.g.* Shine-Dalgarno). Interestingly, mapping the mutations to the secondary structure of the active AapA3 mRNA species (*aapA3*-Tr), we observed that they symmetrically cluster at both sides of the stems, in both, the 5' and 3' termini stem-loops which now we know, are characteristic of this AapA3 mRNA species (mutations shown in orange and purple, respectively, Figure 8.2.3 in the Appendix). To understand how these single point mutations were able to counteract the absence of the IsoA3 sRNA antitoxin, we selected one mutation on each stem-loop for deeper mechanistic characterization. The single point substitution G183A (guanine at position 183 from AapA3 TSS mutated to adenine) located in the 3' stem-loop was surgically introduced into the *H. pylori* chromosome together with two point mutations inactivating the IsoA3 antitoxin promoter, as described in Chapter 4 sections 4.1 and 4.2. Total RNA was extracted and subjected to Northern Blot analysis. Strikingly, the active AapA3 mRNA species (*aapA3*-Tr) was missing in the G183A strain despite IsoA3 absence (compare lanes 5 and 6, Figure 4.3.3A). This result was unexpected and could be explained by two events: 1) the mutation was blocking the 3' end maturation process, or 2) the mutation was prompting *aapA3*-Tr mRNA degradation by 3'-exonucleases.

Because we already identified the 3'-5'-exonuclease PNPase as one of the actors involved in the AapA3 mRNA maturation process and the G183A mutation was disrupting two hydrogen bonds in the 3' stem-loop (predicted to reduce the stem-loop stability by 5.1 kcal/mol), we decided to test the second hypothesis first and asked whether the G183A suppressor was still able to avoid toxicity in absence of PNPase activity. To that aim, the G183A/pIsoA3\* PCR construct was used to perform *H. pylori* transformation assay in wild-type or *-pnp* strains (Figure 4.3.3B). Because the G183A mutation was expected to be toxic in a *-pnp* background, a second mutated PCR construct carrying an additional mutation inactivating the AapA3 translation start codon (*start* G183A/pIsoA3\*) was included in the study (Figure 4.3.3B). As shown in Figure 4.3.3C, *H. pylori* transformation with the G183A/pIsoA3\* PCR construct was lethal in the PNPase deleted background (*-pnp*), obtaining  $\approx 1.4$  log-fold fewer transformants than in wild type background (*+pnp*). Additionally, this number was significantly higher ( $\approx 1.4$  log fold) than that obtained in the no-DNA transformation control ( $H_2O$ , Figure 4.3.3C), reflecting the (expected) emergence of suppressor strains. In line with these results, northern blot analysis revealed the recovery of the *aapA3*-Tr mRNA species in the *start* G183A/pIsoA3\* mutant in *-pnp* background (lane 10, Figure 4.3.3A), explaining its toxicity in the transformation assay (Figure 4.3.3C). These results clearly pointed at a destabilization effect of the G183A mutation in the 3' stem-loop of the *aapA3*-Tr mRNA species causing the full mRNA degradation mediated by 3'-exonucleases. Interestingly, the observation of the *aapA3*-Tr mRNA species in the *start* G183A/pIsoA3\*/*-pnp* strain was possible thanks to the existence of the, previously noticed and unidentified, PNPase redundant activity. Thus, thanks to the study of the G183A suppressor, we were able to differentiate between both activities, as the PNPase redundant activity was proven to be more sensitive to mRNA secondary structure (not being able to degrade the *aapA3*-Tr mRNA despite the presence of the G183A destabilizer mutation, lane 10, Figure 4.3.3A).

To get a more precise idea of the amplitude of the destabilizing effect that the G183A mutation has in the 3' stem-loop, we performed UV-melting experiments using chemically synthesized wild-type or G183A 3' RNA stem-loops as template. Interestingly, during RNA purification, we observed a strong migration delay on the wild-type RNA when compared to the G24A RNA, indicative of a lack of denaturation despite migrating on a denaturing PAA gel (Figure 8.2.4 in the Appendix). As shown in Figure 4.3.3D, a dramatic melting temperature ( $T_m$ ) reduction ( $\sim 10^\circ C$ ) was observed in the G183A stem-loop when compared to that of the wild-type, in line with the 5.1 kcal/mol destabilization effect predicted using RNAfold Web Server (Gruber et al., 2008). Interestingly, a decrease in the G183A curve slope when compared to wild-type can be observed (Figure 4.3.3D), reflecting a loss of cooperativity in the helix, and explaining how a single point mutation can have such a strong impact in the stability of a 24 nt-long stem-loop structure. Together, these results demonstrate that the AapA3 mRNA possess a 3' end terminal protective structure that impede the degradation by 3'-5'-exonucleases.





**Figure 4.3.3. The active *aapA3* mRNA harbors a 3' end protective stem-loop structure.** (A) Total RNA extraction of the indicated strains was performed and 10  $\mu\text{g}$  of RNA was subjected to Northern Blot. The membrane was successively probed with FD38 labeled oligonucleotide and *IsoA3* riboprobe to detect *aapA3* and *IsoA3* transcripts, respectively. The different transcripts are annotated as: *aapA3*-FFL (full-full-length,  $\approx 280$  nt), *aapA3*-FL (full-length,  $\approx 225$  nt), *aapA3*-Tr (truncated,  $\approx 190$  nt); and *IsoA3*-p (precursor,  $\approx 95$  nt) and *IsoA3* (mature,  $\approx 80$  nt). A labeled DNA marker (lane M) was used for size estimation. Proper loading was assessed by 5S rRNA using the labeled oligo probe FD35. (B) Constructs used to study the role of PNPase in *aapA3-Tr* mRNA degradation in the context of the G183A suppressor strain are shown. (C) The number of  $\text{Str}^R$  CFU/ total CFU upon transformation was calculated. Error bars represent s.d.;  $n=3$  biological replicates. (\*  $p<0.076$ , according to unpaired t-test). (D) The suppressor G183A strongly reduces the thermal stability of the 3' stem-loop of the *aapA3-Tr* mRNA. UV-melting experiments were performed on chemically synthesized RNA molecules exclusively consisting on the 3' stem-loop (24 nucleotides). RNAfold (Gruber et al., 2008) was used for secondary structure prediction and VARNA RNA (Darty et al., 2009) for drawing. Suppressor mutation G183A is highlighted in purple.

#### 4.3.1.5. A 5' stem-loop in the active AapA3 mRNA impedes 5'-exonucleolytic degradation

A second group of suppressor mutations was strikingly identified laying in the AapA3 5' UTR region. Like the above-studied mutations in the 3' UTR, those in the 5'UTR clustered symmetrically, this time, in the 5' stem-loop of the active AapA3 mRNA species (mutations shown in orange in Figure 8.2.3 in the Appendix). We selected the single point mutation G24A (guanine at position 24 from AapA3 TSS changed to adenine) for further studies. The G24A/pIsoA3\* strain was surgically generated as described in the 'Chromosomal manipulation techniques' in the Methods section. Northern Blot analysis of the G24A/pIsoA3\* revealed that, similarly to what was observed in the G183A/pIsoA3\* strain, the *aapA3-Tr* mRNA species was absent (compare lanes 5 and 6, Figure

4.3.4B). Secondary structure and stability predictions suggested a 5.5 kcal/mol destabilization effect caused by the G24A mutation (shown in orange in Figure 4.3.4A). UV-melting experiments confirmed this prediction and demonstrated that similarly to the G183A mutation, the G24A mutation leads to a  $\sim 10^{\circ}\text{C}$  decreased in the melting temperature of the stem-loop (Figure 4.3.4E). By similarity to the previously studied G183A mutation, we hypothesized that the G24A mutation was able to counteract IsoA3 absence by promoting the degradation of the active *aapA3*-Tr species by, in this case, 5'-3'-exonucleases.

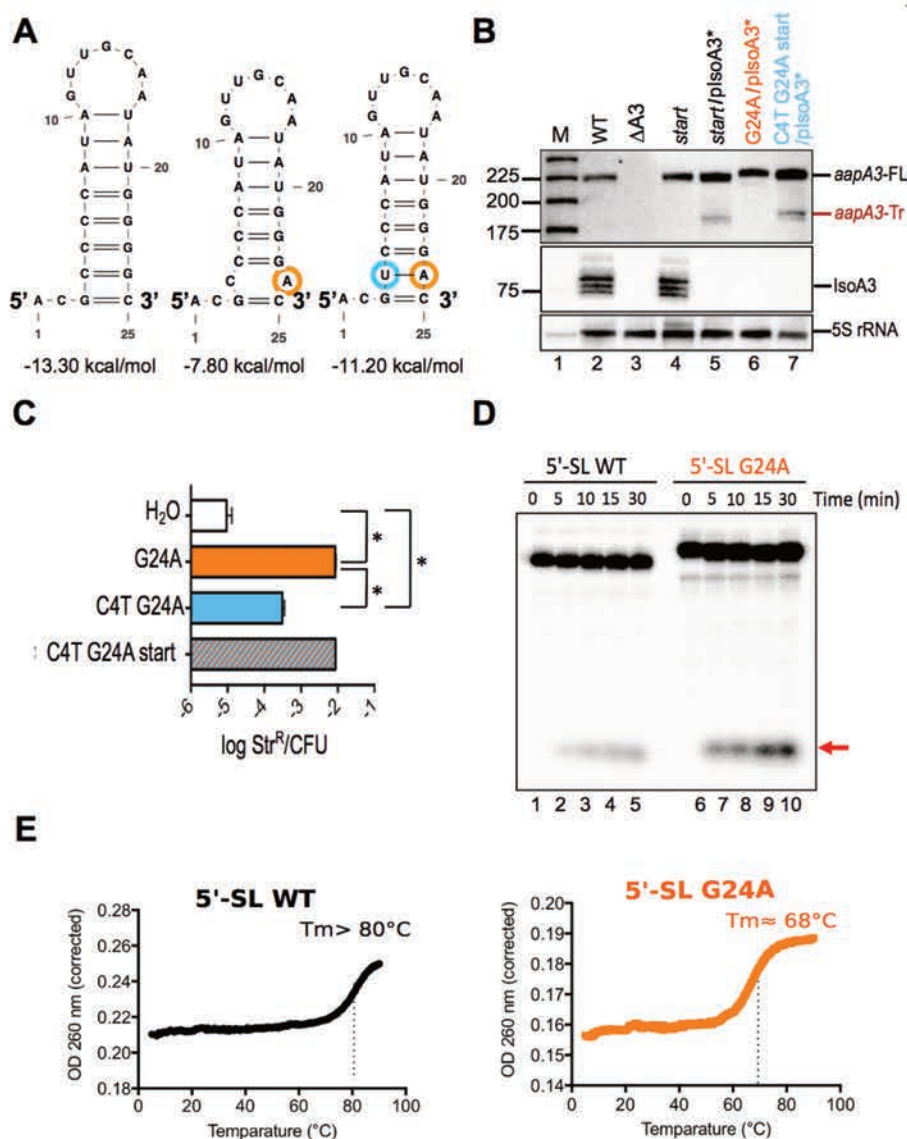
The major *H. pylori* 5'-3'-exonuclease is RNase J, however, unlike the PNPase, the *rnj* gene is essential for *H. pylori*. Thus, we decided to first test the role of the 5' stem-loop stability *in vivo* by performing *H. pylori* transformation assays. To this aim, we generated a PCR construct carrying the G24A mutation together with the C4T mutation (cytosine at position 4 from the *Aapa3* TSS mutated to thymine) and the antitoxin promoter inactivated (allowing toxicity assessment). The C4T mutation locates in the stem in front of the G24A mutation and is predicted to partially restore the stem-loop stability (see predicted kcal/mol values under the structures in Figure 4.3.4A). This restoration could only be partial as we are converting a GC pair (two hydrogen bonds) into an AT pair (one hydrogen bond). Because the C4T mutation was expected to be toxic if the G24A mutation was acting as a 5' stem-loop destabilizer, a third PCR construct carrying an additional point mutation in the *Aapa3* translation start codon (*start*) was introduced to the study. As shown in Figure 4.3.4C, when the C4T G24A PCR construct was used for transformation, a strong reduction ( $\sim 1.5$  log-fold) in the number of transformants obtained was observed when compared to when the PCR carrying the G24A suppressor mutation was used. Additionally, and as expected, this number was significantly higher ( $\sim 1.5$  log-fold) than the one obtained when no DNA was used for transformation ( $\text{H}_2\text{O}$ , Figure 4.3.4C), reflecting the apparition of suppressor strains. Further, Northern Blot analysis confirmed the recovery of the active *Aapa3* mRNA species in the C4T G24A *start*/pIsoA3\* strain (Figure 4.3.4B), explaining the toxicity observed in the transformation assays when the PCR carrying the C4T G24A mutations was used. Together, these results demonstrate that the G24A suppressor destabilizes the 5' stem-loop and promotes the degradation of the *aapA3*-Tr mRNA by 5'-exonucleases (*i.e.* RNase J).

To investigate the role of the RNase J, *H. pylori* RNase J was expressed, purified in *E. coli* cells, and used to perform *in vitro* cleavage assays. Uniformly  $^{32}\text{-P}$  labeled 5' monophosphate *in vitro* synthesized RNA corresponding to the *aapA3*-Tr 5' stem-loop with wild-type or carrying the G24A mutation was used as substrate. As shown in Figure 4.3.4D, a strong increase in the accumulation of labeled mononucleotides (UMP, exonuclease activity product) over time was observed when the G24A mutated RNA was used as substrate. This reflects a facilitated RNase J activity due to the destabilization of the stem-loop and strongly points at RNase J as the responsible of the *aapA3*-Tr disappearance in the G24A suppressor strain.

The here identified suppressor mutants in the 5' stem-loop offered us a unique opportunity to study 5' end-mediated mRNA *decay in vivo*. Although in bacteria lacking RNase E, as is the case of *H. pylori*, the existence of an RNA degradosome remains enigmatic, the presence of a 'minimal RNA degradosome' composed by RNase J and the helicase RhpA (HP0249) has been proposed in *H. pylori* (Redko et al., 2013). To test whether the observed 5'-exonucleolytic degradation of the G24A *aapA3*-Tr mRNA was dependent on the helicase activity, we generated a *rhpA* deletion strain and used it as genetic background to introduce the G24A mutation. If RhpA was required for RNase J activity on the G24A *aapA3*-Tr mRNA, such suppressor is expected to be lethal in absence of RhpA activity. As shown in Figure 8.2.5A in the Appendix, the G24A suppressor was still viable in absence of RhpA and no changes in *aapA3* expression or processing pattern were observed. A significant growth defect was observed in absence of RhpA, but this was independent on the presence of the G24A mutation



(Figure 8.2.5B in the Appendix). These results show that the G24A *aapA3*-Tr mRNA 5'-exonucleolytic degradation is independent of RhpA activity.



**Figure 4.3.4. The 5' stem-loop of the active *aapA3* mRNA avoids 5'-exonucleolytic degradation.** (A) Secondary structure prediction of the 5' end stem-loop of the *aapA3*-Tr mRNA is shown. Suppressor mutation G24A is highlighted in orange; C4T mutation restoring the base pair disrupted by G24A is highlighted in blue. RNAfold (Gruber et al., 2008) was used for secondary structure prediction and VARNA RNA (Darty et al., 2009) for drawing. (B) Total RNA extraction was performed subjected to Northern Blot analysis. The membrane was successively probed with FD38 labeled oligonucleotide and IsoA3 riboprobe to detect *aapA3* and IsoA3 transcripts, respectively. The different transcripts are annotated as: *aapA3*-FFL (full-full-length, ~280 nt), *aapA3*-FL (full-length, ~225 nt), *aapA3*-Tr (truncated, ~190 nt) and IsoA3 (~80 nt). A labeled DNA marker (lane M) was used for size estimation. Proper loading was assessed by the level of 5S rRNA. (C) The number of  $\text{Str}^R$  CFU/ total CFU upon transformation was calculated. Error bars represent s.d.; n=3 biological replicates. (\* p<0.0001, according to unpaired t-test). (D) 5'-3'-exonucleolytic RNase J activity is prompt by the 5' stem-loop suppressor G24A. Internally  $^{32}\text{P}$ -labelled 5'-P *in vitro* synthesized RNA containing the 47 first nucleotides of the *aapA3* mRNA (comprising the 5' stem-loop) with wild-type or G24A mutated sequence were subjected to *in vitro* cleavage assay with *H. pylori* RNase J. A red arrow next to the gel indicates the product of exonucleolytic activity (E) The G24A suppressor strongly reduces the thermal stability on the 5' stem-loop on the *aapA3*-Tr RNA. UV-



melting experiments were performed on chemically synthesized RNA consisting on the 5' stem loop (5'-SL, 25 nucleotides).

---

Another key factor of 5' end-mediated mRNA decay is the 5' phosphorylation state as observed for the IsoA3 sRNA here studied as well as in several previous studies (Mackie, 1998, 2000; Xu and Cohen, 1995). To determine whether the 5'-exonucleolytic degradation observed in the G24A *aapA3*-Tr mRNA was dependent in the 5' phosphorylation state, we used the RppH deletion strain as genetic background to introduce the G24A suppressor as well as an independent suppressor also located in the 5' stem-loop, the C4G mutation (cytosine at position 4 for the AapA3 TSS mutated to guanine). As shown in Figure 8.2.6A in the Appendix, the suppressors were not lethal in absence of RppH activity and no changes in the expression or processing profile were observed. These results demonstrate that, as shown in Figure 8.2.6B, the destabilizer effect of the 5' stem-loop suppressors is so dramatic that 5'-exonucleases (*i.e.*, RNase J) are able to cleave independently of the 5' phosphorylation state. At the same time, and interestingly, the results highlight that in wild-type conditions, the stability of the 5' stem-loop is so high that 5'-exonucleases (*i.e.*, RNase J) are not able to act despite the presence of a 5' monophosphate. Indeed, *in vitro* cleavage assays using 5'<sup>32</sup>-P *aapA3*-Tr RNA as substrate showed an almost complete absence of RNase J exonuclease activity, being only a few endonucleolytic cleavages in single-stranded regions visible (Figure 8.2.7 in the Appendix).

## 4.3.2. Discussion

Studying how toxin-encoding mRNAs are co- and post-transcriptionally regulated is key for the comprehensive understanding of type I toxin-antitoxin (TA) regulation as genetic systems. Here, we present the first insights into the characterization of the activation and decay pathways of the AapA3 toxin-encoding mRNA, belonging to the *aapA3*/IsoA3 type I TA locus present in the *H. pylori* chromosome. To explore the molecular underpinnings of such pathways, we made use of traditional rifampicin treatment assays, chromosome engineering, 3' RACE (Rapid Amplification of cDNA Ends), biophysics and computational predictions, and the newly developed high-throughput selection of suppressors, FASTBAC-Seq (Masachis et al., 2018).

### 4.3.2.1. Behind the scenes of a 2D-structure-based 3' end maturation of a toxin-encoding mRNA

In previous work, we identified the *aapA3*/IsoA3 type I TA module (Sharma et al., 2010) and characterized its regulation at the chromosomal level (see Chapter 4 sections 4.1 and 4.2). Considering that the AapA3 mRNA encodes a constitutively expressed toxin with self-targeting activity, it is intuitive to think that it must be optimally regulated co- and post-transcriptionally by means that complement the blind-spots of the IsoA3 sRNA antitoxin activity (for a review see (Masachis and Darfeuille, 2018) in Chapter 1, section 1.4.1). We recently reported on the existence of a truncated AapA3 mRNA species that accumulates in the absence of IsoA3 or RNase III. Indeed, this species is the only translationally active one and the IsoA3 sRNA competes with the ribosome to bind to it. Once the IsoA3 RNA is bound to its target, the RNA heteroduplex is rapidly degraded by RNase III, explaining its accumulation in the absence of this two elements (see Chapter 4 section 4.2). At that point, we hypothesized, based on RNA-Seq data (Sharma et al., 2010) and by the fact that the IsoA3 RNA targets the 5' end of the toxin mRNA, that this truncated species is the product of a 3' end processing. However, the molecular underpinnings of this process were unknown.

Here, we demonstrate that the translationally inert AapA3 primary transcript undergoes a 3' end exonucleolytic activation process (as previously observed for the AapA1 mRNA (Arnion et al.,

2017, see appendix 8.3 in Chapter 8)) led by 3'-5'-exonuclease PNPase. The endonuclease RNase Y, as well as, one or more, unidentified redundant activity(ies) are also involved. An interaction between PNPase and RNase Y has been previously described in *B. subtilis* (Salvo et al., 2016), but further investigation will be required to decipher its role in *H. pylori*. A similar 3' end processing has been described for the Hok mRNA, belonging to the *hok/Sok* type I TA system in *E. coli* (Thisted et al., 1994). Remarkably, in the case of the Hok mRNA, the actors of its activation could be fully unmasked, being the PNPase and the RNase II. The fact that the AapA3 and the Hok mRNAs share such similar activation pathways despite having no sequence homology is striking. Indeed, as discussed in Masachis and Darfeuille (2018) (Chapter 1 section 1.4.1), a major hallmark of the regulation of type I TA systems is that toxin-encoding mRNAs possess their SD sequence occluded and must undergo a nucleolytic activation step to be translated. How this activation occurs is not random and strongly depends on the operon organization of the TA system. In systems where the antitoxin is encoded overlapping the 5' end of the toxin mRNA (e.g., AapA1 (Arnion et al., 2017) and Hok (Thisted et al., 1994) mRNAs), SD occlusion is achieved by a long-distance interaction between the 5' and 3' mRNA ends, and activation occurs by 3'-trimming. On the opposite, in systems whose antitoxin is encoded in no-overlapping fashion, SD occlusion occurs within stem-loops and activation by 5' processing. However, as in the case of 3' processed mRNAs, despite several 5' ends have been mapped for some mRNAs such as the TisB (Darfeuille et al., 2007), DinQ (Weel-Sneve et al., 2013), ShoB (Kawano et al., 2005) and ZorO (Wen et al., 2014), the activities responsible for this activation remain enigmatic.

Defining mRNA termini is an essential step for the mechanistic understanding of any given mRNA. However, despite protocols for the identification of 5' ends have been well set up for years (Argaman et al., 2001; Thomason et al., 2015; Urban and Vogel, 2007), defining 3' mRNA ends in bacteria is challenging due to the lack of poly-U tails in most transcripts. Nevertheless, a recent study by Dar and colleagues., (2016) described a new genome-wide 3' end mapping strategy based on an improved 3'RACE protocol. We used this protocol with slight modifications (see Methods section) to study in more detail the AapA3 mRNA 3' end activation pathway. Remarkably, 3'RACE revealed that the three main AapA3 mRNA species, previously observed by Northern Blot analysis, are in fact heterogeneous populations of sequentially 3' processed mRNA species, in line with the here proposed 3'-exonucleolytic activation pathway. Furthermore, the different 3' ends are not defined arbitrarily but by energetic roadblocks for 3'-exonuclease activity. This result, was of high informative value and revealed a major hallmark of the active AapA3 mRNA species that was later on exploited as a tool to study mRNA decay: it possesses highly stable 5'- and a 3'- termini stem-loop structures.

Indeed, it can be striking to see the PNPase enzyme involved in an activation pathway instead of leading to mRNA decay (as originally described for the RNA I, antisense inhibitor of ColEI-type plasmids (Xu and Cohen, 1995)). However, by studying the AapA3 mRNA, we understood that this is possible thanks to the formation of a highly stable 3' stem-loop that acts as an energetic shield impeding the PNPase 3'-exonucleolytic activity at the right place, by limiting its degradation to the only  $\approx 35$  last nucleotides to get rid of the aSD sequence that occludes ribosome accessibility to the SD sequence. Thus, the *in vivo* 3' end of the active AapA3 mRNA is derived from the 3'-5'-exonucleolytic processing of the primary transcript back to the 3' terminal protective hairpin. Importantly, these results demonstrated that although some highly stable 3' end stem-loop structures are often automatically thought to act as Rho-independent terminators, in the case of the AapA3 mRNA, this structure has both, protective and 'activating' functions instead.

A similar case has been previously reported in *E. coli*. The tryptophan operon mRNA undergoes a 3'-exonucleolytic maturation mediated by PNPase and RNase II. This processing stops at a 3' stabilizer hairpin structure that was previously assumed to act as an intrinsic terminator (Mott et

al., 1985). The observation of such similar maturation patterns in these two distant mRNAs is interesting, and invites to the careful study of 3' end stem-loop structures thought to act as Rho-dependent terminators to search for hidden mRNA stability-related functions.

#### 4.3.2.2. Termini protective stem-loops ensure functional stability of the AapA3 mRNA

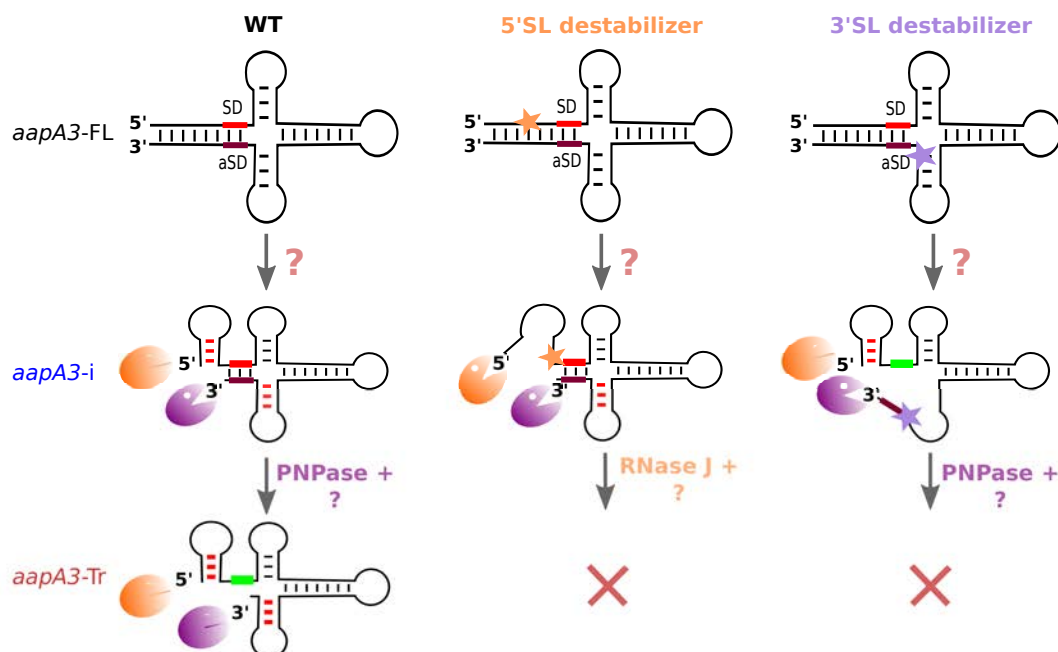
3'RACE highlighted that the active AapA3 mRNA possesses two highly stable termini stem-loop structures. FASTBAC-Seq identified two sets of mutations symmetrically clustering at both termini structures. As shown in Table 8.2.1 in the Appendix, the 5' stem-loop is highly stable, having a predicted Minimum Free Energy of -13.30 kcal/mol. From the four selected mutations in the 5' stem-loop, two of them, C4G (cytosine at position 4 from the AapA3 TSS mutated to guanine) and G24A (guanine at position 24 from the AapA3 TSS mutated to adenine), were predicted to have a strong destabilization effect of 5.6 kcal/mol and 5.5 kcal/mol, respectively. The other two, G3C (guanine at position 3 from the AapA3 TSS mutated to cytosine) and C4T (cytosine at position 4 from the AapA3 TSS mutated to thymine), were predicted to have a lower destabilization effect of 3 kcal/mol and 2.7 kcal/mol, respectively. Interestingly, the amplitude of the predicted destabilization effect correlated with the padj value that each mutation had in the statistical analysis of the deep-sequencing datasets, suggesting that low destabilization effects may lead to more dynamic or less stable toxicity avoidance phenotypes (Table 8.2.1 in the Appendix). However, this correlation was not so clear with the selected mutations in the 3' stem-loop, as the G183A here studied suppressor has the lowest padj value ( $2.24 \times 10^{-40}$ ) but is not predicted to have the strongest stem-loop destabilization effect (Table 8.2.2 in the Appendix).

We choose to study one mutation on each stem-loop, having a padj value above our cut-off ( $\text{padj} < 0.05$ ), to test whether this arbitrarily-selected statistical threshold was able to discriminate false-positives. As shown in Table 8.2.1 and Table 8.2.2 in the Appendix, a good correlation between the padj values and viability (or not) of the strains was observed. Thus, our statistical analysis efficiently identified as false-positives the C4T mutation in the 5' stem-loop ( $\text{padj} = 0.64$  and lethal, Table 8.2.1 in the Appendix) and the G176 mutation in the 3' stem-loop ( $\text{padj} = 0.2083$  and lethal, Table 8.2.2 in the Appendix). However, the discriminative power of our statistical analysis is not infallible. Some mutations may display intermediate toxicity phenotypes, as is the case of the G3C mutation, which we were able to obtain but not in a reproducible nor stable manner. Thus, the G3C mutation appears in our statistical analysis as a 'real' suppressor mutation ( $\text{padj} = 0.0039$ ) despite not having a stable suppressor phenotype (Table 8.2.1 in the Appendix). This limitation may be linked to the way the suppressors are selected, as upon 3 only days growth, the genotypic stability of a suppressor strain linked to the different toxicity levels cannot be fully assessed. Nevertheless, we estimate the number of cases following the G3C situation to be strongly limited and therefore having a minimum deleterious impact in the FASTBAC-Seq analysis.

Northern Blot analysis of the G24A in the 5' stem-loop, and the G183A in the 3' stem-loop revealed the absence of the *aapA3*-Tr mRNA species despite the lack of IsoA3. We further demonstrated that the G183A mutation is a PNPase-dependent suppressor, as in a deleted background for *pnp* the G183A mutation is lethal due to the presence of the active AapA3 mRNA species. Importantly, this observation could only be made thanks to the presence of the, so far unidentified, PNPase redundant activity, as otherwise the *aapA3*-Tr mRNA would not be generated. Moreover, the G183A mutation allowed us to discriminate between the PNPase and the redundant activity, as the latter one has proven to be more sensitive to RNA secondary structure. It is not surprising to observe this strong sensitivity to secondary RNA structure in 3'-exonucleases. Previous studies in *E. coli* have reported a strong inhibition of the 3'-5'-exonucleases RNase II and PNPase by secondary structure

(Spickler and Mackie, 2000). Nevertheless, some 3'-5'-exonucleases such as RNase R, possess an intrinsic helicase activity and can easily degrade highly structure RNA molecules (Hossain et al., 2016).

In a similar manner, we were able to demonstrate that the G24A mutation located in the 5' stem-loop is a 5'-exonuclease-dependent suppressor. Many examples have previously reported mRNA stabilization mediated by the presence of a 5' stem-loop structure (Bouvet and Belasco, 1992; Emory et al., 1992; Matsunaga et al., 1996). However, the viability of the G24A mutation in a *rppH* deleted background, demonstrate that the destabilization effect of the G24A mutation in the 5' stem-loop is independent of the 5' phosphate state. Indeed, 5'-exonucleases (*i.e.*, RNase J) are able to degrade the mutated mRNA albeit the presence of a 5' triphosphate. On the other hand, these results reflect that the stability of the wild-type 5' stem-loop is so high that 5'-exonucleases cannot act despite the presence of a 5' monophosphate. These results are surprising as RNase J it is known to be sensitive to the 5' phosphate state, as shown in the case of the IsoA3 (see Chapter 4 section 4.4), as well as in several previous studies (Daou-Chabo and Condon, 2009; Even et al., 2005; Mackie, 1998, 2000; Xu and Cohen, 1995). Going further, these results reflect to which extend the toxin mRNA is has evolved to be strictly regulated (having to undergo a 3' nucleolytic activation processing) but at the same time, remarkably highly stable, having an estimated half-life similar to the *H. pylori* doubling time. The biological relevance of this phenomenon is mysterious, however, we could hypothesize that synchronizing the half-life of a toxin-encoding mRNA with the doubling time opens the possibility to the vertical transmission of the mRNA, following an innovative variation of the previously described mechanism of Post-Segregational Killing (PSK) (Franch et al., 1997; Gerdes and Wagner, 2007; Gerdes et al., 1986; Greenfield et al., 2000, 2001; Nielsen et al., 1991; Weaver, 2012). Considering the short half-life of the antitoxin sRNA, it could represent a time-window on which the toxin mRNA can be expressed, which may be favorable in some specific situations (*e.g.*, to increase phenotypic heterogeneity under stress conditions).



**Figure 4.3.5. Functional AapA3 mRNA stability is ensured by two termini protective stem-loop structures.** In wild-type conditions, the *aapA3*-FL species is protected from degradation (and from IsoA3 and ribosomes binding) via a 5'-3' long-distance interaction. 3' end processing by a, so far unidentified enzyme (?), generates an intermediate species (*aapA3*-i) that is further processed by PNPase together with a redundant activity to generate the active *aapA3*-Tr mRNA. This species has highly stable 5'- and 3' stem-

loops that act as structural shields against exonucleolytic degradation. A single point mutation on the 5'- or 3'- stem-loop can lead to the full degradation of the active *aapA3* mRNA, mainly mediated by the 5'-3'-exonuclease RNase J or the 3'-5'-exonuclease PNPase, respectively.

---

Figure 4.3.5 summarizes the knowledge on the role secondary RNA structure for the stability and activation of the AapA3 toxin-encoded mRNA that was revealed by the study of the 5' and 3' stem-loop suppressors. In the wild-type situation (WT), the translationally inert AapA3 primary transcript (*aapA3*-FL) undergoes a 3' exonucleolytic activation process mediated by PNPase and one or more, so far, unidentified redundant activity(ies). The active AapA3 mRNA (*aapA3*-Tr) possess two highly stable termini step-loops that act as energetic shields against exonuclease activity (despite 5' monophosphate state) and ensure its functional stability. In the context of a 5' stem-loop (5' SL) destabilizer suppressor, 5'-exonucleases degrade the *aapA3*-i and/or -Tr mRNA species, independently of the 5' phosphate state. Importantly, such mutations do not lead to *aapA3*-FL degradation, indicating that the co-transcriptional mRNA folding is fast enough as to reach the stable 5'-3' long-distance interaction conformation before 5'-exonucleases can act. In a similar manner, 3' stem-loop (3'SL) destabilizer mutations, trigger the degradation of the *aapA3*-i and/or -Tr mRNA species by 3'-exonucleases, including the PNPase.

#### 4.3.3. Conclusion

The existence of 5' and 3' protective RNA structures is increasingly recognized as a major determinant of mRNA decay in bacteria (e.g., being involved in the differential operon decay, Dar and Sorek, 2018) (Dar et al., 2016; Emory et al., 1992; (Mott et al., 1985)). As described in the section 4.2 of this Chapter (Paper III), the AapA3 mRNA must be kept untranslated, co- and post-transcriptionally, to avoid cell death. Untranslated RNAs can rapidly become unstable. The results of this work showed that termini protective stem-loop structures need to be present at both ends of the AapA3 toxin-encoding mRNA to ensure its functional stability. In the absence of one of such protective hairpins, the translation efficiency at the population level is strongly reduced, leading to toxin production levels that lie below the lethality threshold. Hence, even though translation initiation is assumed to be the rate-limiting step of protein production in bacteria, our work shows that, in the case of the AapA3 mRNA, mRNA stability is the limiting factor. Furthermore, this work highlighted the key role of such protective structures, not only in the AapA3 mRNA stability, but also for its proper 3'-exonucleolytic activation. Indeed, this work revealed that the 3' stem-loop of the AapA3 mRNA does not act as an intrinsic terminator, but instead, it determines the position until which the maturation process occurs and generate an active mRNA.

As discussed in our recent review (see section 1.4.1 in Chapter 1, Masachis and Darfeuille, 2018), understanding how the activation of type I toxin-encoding mRNAs occurs is key for comprehension of the regulation of these systems. However, even though the existence of such activation mechanism has been reported in a few cases, including the 3' end activation of the Hok toxin mRNA (Franch et al., 1997), and the 5' end activation of the TisB (Darfeuille et al., 2007), DinQ (Weel-Sneve et al., 2013), ShoB (Kawano et al., 2005) and ZorO (Wen et al., 2017) toxin mRNAs, in most cases the identity of the activities involved remain enigmatic. The future development of tools such as FASTBAC-Seq (Masachis et al., 2018), may help revisit the post-transcriptional regulatory mechanisms of previously-identified type I TA systems, and shed light into the molecular underpinnings behind the strict regulation of type I toxin-encoding mRNAs expression.



## 4.4. On the decay of a sRNA antitoxin

In the previous section (Paper IV), the role of mRNA structure in the high stability of an untranslated toxin-encoding mRNA was studied. However, the expression output of a toxin belonging to a type I TA system, do not only depend on its self, but in the molecular ratio between the toxin mRNA and the RNA antitoxin counterpart. The levels of RNA antitoxins can be post-transcriptionally adjusted by RNA degradation. In our recent studies, we have shown that IsoA1 sRNA, as well as IsoA3 sRNA, are degraded by the RNase III when forming an RNA heteroduplex with their toxin-encoding mRNA counterparts (see section 4.3 Chapter 4, and Arnion et al., (2017), section 8.3 in Chapter 8). However, the functional pools of sRNA antitoxins are, firstly and strongly, dependent on the stability of the sRNAs themselves (*i.e.*, in ‘free-state’). The activities involved in this process vary depending on the bacterial species. For instance, RNase E is involved in the functional inactivation of the Sok RNA in *E. coli* (Dam Mikkelsen and Gerdes, 1997), meanwhile RNase Y and PNPase control several sRNAs in *B. subtilis* (*e.g.*, SR4 and RatA RNA antitoxins, Durand et al., 2012; Jahn and Brantl, 2013; Silvaggi et al., 2005).

Interestingly, an RNA-sequencing study of a *H. pylori* RNase J depleted strain, recently suggested that between the several targets of RNase J, there was the IsoA3 sRNA antitoxin (Redko et al., 2016). Additionally, another recent global RNA-seq study, demonstrated that the IsoA3 sRNA is stabilized in a *H. pylori* strain deleted for the RppH (Bischler et al., 2017). The *H. pylori* RppH (HP1228, HpRppH) is an RNA pyrophosphohydrolase from the Nudix family of enzymes, that converts 5’ triphosphates and diphosphates to monophosphates, triggering 5’ end exonucleolytic RNA degradation. Because RNase J activity has been previously shown to be sensitive to the RNA 5’ end phosphorylation state (Mathy et al., 2007; Richards et al., 2011), the observations shown in the two-above-mentioned studies were consistent, and motivated a deeper characterization of the molecular activities underlying the IsoA3 sRNA antitoxin decay mechanism.

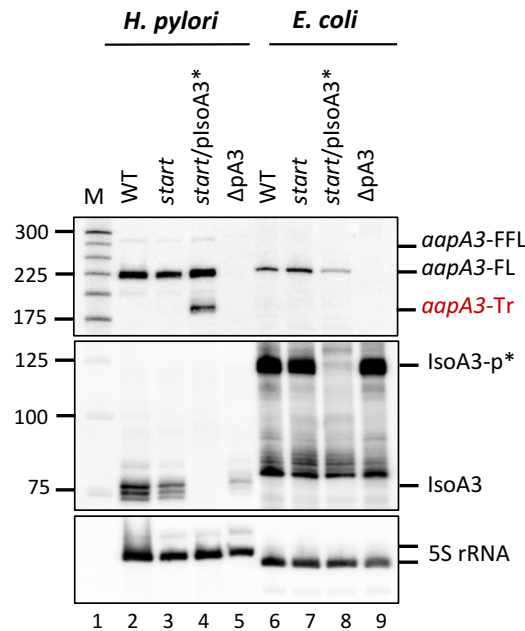
### 4.4.1. Results

#### 4.4.1.1. PNPase is involved in the maturation of the IsoA3 sRNA antitoxin

During the study of the AapA3 mRNA activation pathway, Northern Blot analysis of rifampicin assays in a *H. pylori* PNPase deleted strain revealed a strong accumulation of a long IsoA3 transcript (IsoA3 precursor, IsoA3-p,  $\approx$  95 nt) (see Figure 4.3.1B in the section 4.3.1 of this Chapter). This result indicated that, as shown for other type I antitoxin sRNAs (Durand et al., 2012), the 3’-exonuclease PNPase is involved in the processing of IsoA3 *in vivo*. Additionally, relatively high levels of IsoA3-p can be observed even upon the apparition of the *aapA3*-Tr mRNA (higher levels than the mature IsoA3 RNA ( $\approx$  80 nt)), indicating that IsoA3-p is totally or partially impaired to interact with the *aapA3*-Tr mRNA, and suggesting that PNPase processes 3’end of IsoA3-p to generate the mature IsoA3 sRNA form. Importantly, the observation of the mature IsoA3 RNA species despite the lack of PNPase activity reflects the existence of a redundant activity able to partially compensate its absence, and generate enough (although in lower levels than in wild-type conditions) mature IsoA3 RNA as to avoid AapA3 toxicity (as shown in Figure 4.3.1B in the section 4.3.1 of this Chapter). Whether this redundant activity is the same as the one acting on *aapA3*-i mRNA remains unknown, nevertheless, some clues came from the study of IsoA3 expression in *E. coli* cells.



When the *H. pylori* *aapA3*/IsoA3 module is expressed in *E. coli* cells (from its endogenous promoter), a strong accumulation of a long IsoA3 species is observed (IsoA3-p\*,  $\approx 125$  nt, lanes 6, 7 and 9, [Figure 4.4.1](#)). Strikingly, although a full-length AapA3 mRNA of a similar size to the one observed in *H. pylori* can be observed in *E. coli* cells, the mature IsoA3 sRNA ( $\approx 80$  nt) is absent. This result strongly suggests that the *E. coli* PNPase activity is not able to mature the IsoA3 sRNA. Furthermore, the full lack of mature IsoA3 sRNA indicates that *E. coli* may lack the PNPase redundant activity previously observed in *H. pylori* (or it may as well be inactive on this sRNA). In line with these results, the active *aapA3*-Tr mRNA is absent in *E. coli*, despite the inactivation of IsoA3 expression (compare the lanes 2-4 and 6- 8, [Figure 4.4.1](#)). Together, these results strongly point at the lack of a 3'-exonuclease activity able to process the AapA3 or IsoA3 RNAs in *E. coli*, highlighting the importance of the use of endogenous conditions when studying post-transcriptional gene expression regulation. Nevertheless, the non-viability of an *E. coli* strain harboring a wild type AapA3 peptide and lacking IsoA3 expression (data not-shown), demonstrated that despite not processing the AapA3 mRNA, toxin expression could be achieved from the *aapA3*-FL species. This phenomenon is probably due to the coupling of transcription and translation processes that may take place in the fast-growing *E. coli* but not in the slow grower *H. pylori*.



**Figure 4.4.1. *E. coli* PNPase activity on the AapA3 and IsoA3 RNAs is strongly impaired.** 10  $\mu$ g total RNA from the wild-type or the indicated *aapA3*/IsoA3 mutated versions *H. pylori* 26695 strains, and *E. coli* TOP10 strain harboring a pGEM-T vector carrying the different indicated *aapA3*/IsoA3 locus versions was subjected to Northern Blot. Membrane was successively probed for IsoA3 and 5S rRNA transcripts as described in. Transcripts are annotated as: *aapA3*-FFL (full-full-length,  $\approx 280$  nt), *aapA3*-FL (full-length,  $\approx 225$  nt), *aapA3*-Tr (truncated,  $\approx 190$  nt), IsoA3-p\* (*E. coli*,  $\approx 125$  nt) and IsoA3 (*H. pylori*,  $\approx 80$  nt). *Start*, AapA3 translation start codon mutated (ATG $\rightarrow$ ATT by the single point mutation G54T, from AapA3 TSS). pIsoA3\*, IsoA3 -10 box inactivated (5' TATAAT 3' $\rightarrow$ 5' TACAAG 3', by the point mutations A87C A90G, from AapA3 TSS).  $\Delta$ pA3, AapA3 -10 box inactivated (5' TAGGAT 3' $\rightarrow$ 5' TAGGAC 3', by the single point mutation T-7C, from AapA3 TSS).

Although these results demonstrate that PNPase plays a major role in the maturation of the IsoA3 sRNA, rifampicin assays in a *H. pylori* 26695 RNase Y and RNase Y/PNPase deletion strains revealed decreased levels of both, the IsoA3-p and the IsoA3 RNAs ([Figure 8.2.1 in the Appendix](#)). This result is sticking and suggest that RNase Y is involved in the generation not only of IsoA3 ( $\approx 80$

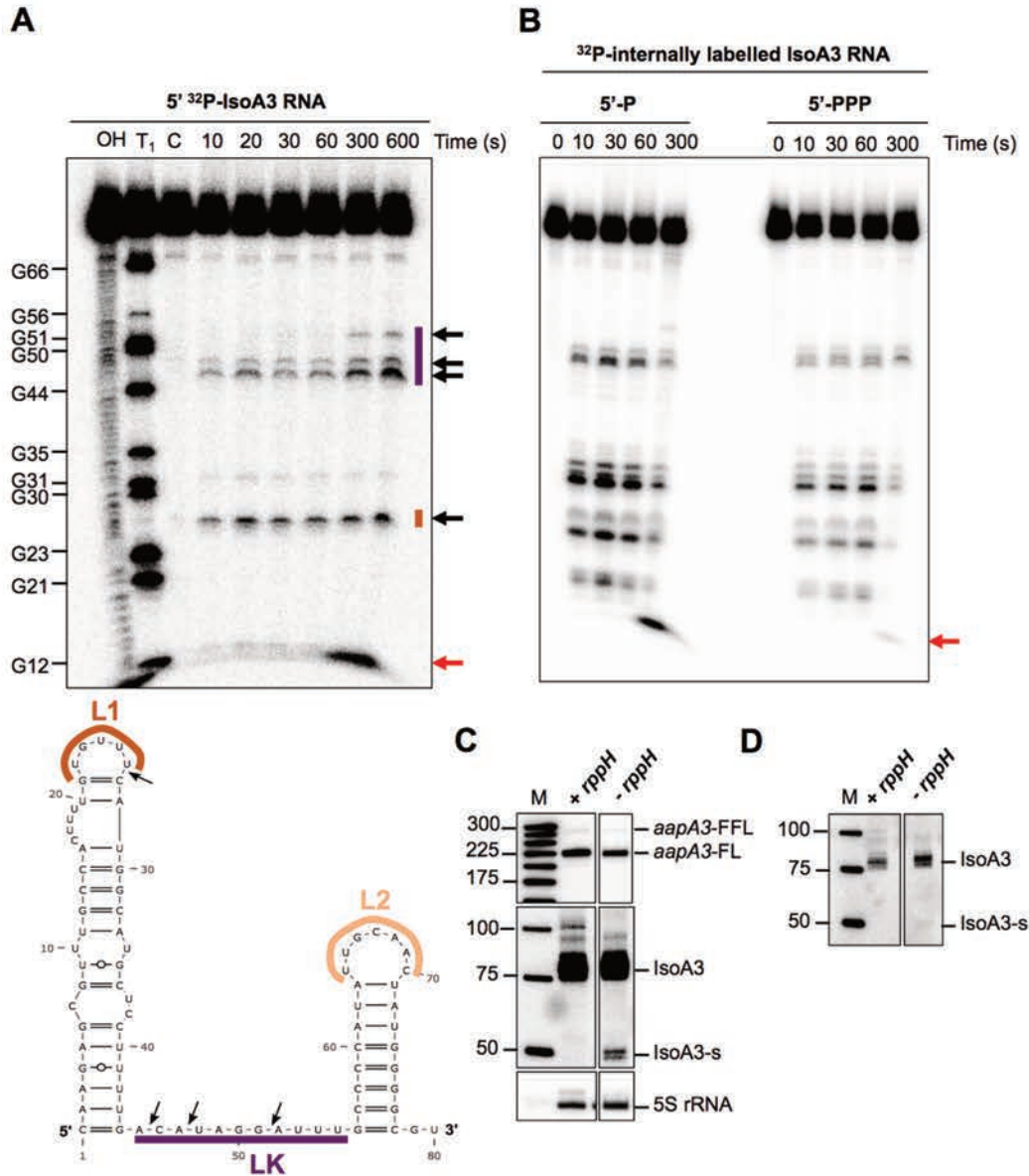
nt), but also of the IsoA3-p ( $\approx 95$  nt). Nevertheless, and similarly to what was observed for the AapA3 mRNA maturation, the presence of the mature IsoA3 species despite the lack of RNase Y and PNPase activity reflects the existence of, so far enigmatic, additional actors on its maturation pathway.

#### 4.4.1.2. RNase J is involved in the IsoA3 sRNA decay and senses its 5' phosphate *in vivo*

To investigate the actors of IsoA3 decay, we studied the role of the major *H. pylori* 5'-3'-exonuclease and endonuclease RNase J, recently suggested to be involved in IsoA3 degradation by RNA-seq studies (Redko et al., 2016). Because the *rnj* (HP1430) is an essential gene for *H. pylori*, the *H. pylori* RNase J protein was expressed and purified in Rossetta2 *E. coli* cells. On a first test, 5' end  $^{32}$ -P labeled IsoA3 RNA was used as substrate to differentiate and map the potential exo- and endonuclease activities of RNase J. As shown in Figure 4.4.2A, three major endonucleolytic cleavages appear within the first 10 seconds of incubation and their intensity increase over time. After 300 seconds (s), two more cleavage products become visible. Interestingly, if we map the cleavages in the IsoA3 RNA secondary structure prediction, we observe that they are in single-stranded regions, and mainly, in the so-called IsoA3 linker (LK, Figure 4.4.2A). The LK region, stands between the two IsoA3 stem-loop structures (annotated as L1 and L2 in Figure 4.4.2A), and may be involved in the locus-specific AapA3 mRNA recognition. From the cleavages that lay outside the LK region, one is located in the L1, and the second one, which appears only upon 300 s of incubation, correspond to released labeled mononucleotides product of exonuclease activity (indicated by a red arrow next to the gel; from this gel, it could have also been an endonucleolytic cleavage close to the nucleotide at position 10, but the experiment in Figure 4.4.2B further evidence that they are labeled mononucleotides product of exonucleolytic activity).

Numerous studies point to the importance of the 5' end in the regulation of mRNA decay in bacteria (Mackie, 1998, 2000; Xu and Cohen, 1995). We next investigated whether the observed IsoA3 processing events were dependent on the 5' phosphorylation state (as recently proposed by RNA-seq studies (Bischler et al., 2017)). To that aim, uniformly  $^{32}$ -P-labeled triphosphorylated and monophosphorylated IsoA3 RNAs were used as substrates. As shown in Figure 4.4.2B, RNase J activity was strongly reduced in the presence of a 5' triphosphate end (red arrow next to the gel indicates labeled mononucleotide product of 5' exonucleolytic activity). This result demonstrates that similarly to other mRNA targets (e.g., mRNAs from the T-box family, Even et al., (2005)), RNase J activity on IsoA3 sRNA is sensitive to the 5' phosphorylation state.

We next sought to investigate the effect of the IsoA3 sRNA 5' end state *in vivo*. To that aim, we generated a *H. pylori* 26695 deletion strain for the pyrophosphohydrolase gene *rppH* (HP1228) and studied IsoA3 expression by Northern Blot. As shown in Figure 4.4.2D, there is a mild stabilization effect of IsoA3 in the absence of RppH (- *rppH*), as previously reported by Bischler et al., (2017). However, a shorter IsoA3 RNA species accumulates in - *rppH* background (IsoA3-s,  $\approx 50$  nt, Figure 4.4.2C). Strikingly, the size of this species corresponds to the most abundantly observed endonucleolytic cleavage products (cleavage in the LK region). This result allowed us to identify *in vivo* an endonucleolytic cleavage in the IsoA3 LK sequence, and demonstrate that RNase J senses the 5' phosphorylation state as part of the decay pathway of the IsoA3 sRNA.



**Figure 4.4.2. IsoA3 sRNA processing by RNase J is sensitive to the 5' phosphorylation state.** (A upper panel) RNase J shows endonucleolytic activity on IsoA3 substrate *in vitro*. *H. pylori* RNase J was purified from *E. coli* Rossetta2 cells. ~0.1 pmol 5' end [ $^{32}\text{P}$ ]-labelled *in vitro* transcribed IsoA3 RNA was incubated with 28  $\mu\text{M}$  of purified RNase J at 37°C and reaction were stopped at the indicated time points by addition of one volume of 2X Loading Buffer II. Untreated RNA (C) and partially digested RNA (OH) served as control and ladder, respectively. Positions of all G residues revealed by T1 treatment under denaturing conditions are indicated relative to the transcription start site of the IsoA3 gene. Reaction products were analyzed on an 10% denaturing PAA gel. Main endonucleolytic products are indicated by black arrows and colored lines indicate positions in the secondary RNA structure. Labeled released mononucleotides product of exonucleolytic activity is indicated by a red arrow next to the gel. (A lower panel) RNAfold Web Server (Gruber et al., 2008) was used to make secondary structure prediction of IsoA3 RNA and VARNA RNA (Darty et al., 2009) was for drawing. Loops are indicated on the structure as L1 (loop 1) and L2 (loop 2). The linker region (LK) is indicated by a purple line. (B) Uniformly  $^{32}\text{P}$ -labelled 5'-P or 5'-PPP *in vitro* synthesized IsoA3 RNA was subjected to *in vitro* cleavage assay with *H. pylori* RNase J. Reactions were stopped by the addition of 10  $\mu\text{l}$  2X Loading Buffer II at the different time points indicated on top of the gel. Cleavage products were run on a 6% denaturing PAA gel and analyzed using a Pharos FX phosphorimager. Labeled mononucleotide product of exonucleolytic activity is indicated by a red arrow next to the gel. (C) Total RNA was extracted subjected to Northern Blot analysis. The membrane was successively probed with FD38 labeled oligonucleotide and IsoA3 riboprobe to detect *aapA3* and IsoA3 transcripts, respectively. A labeled DNA marker (lane M) was used for size estimation. Proper loading was



assessed by the level of 5S rRNA using the labeled oligo probe FD35. Transcripts are annotated as: *aapA3*-FFL (full-full-length,  $\approx$  280 nt), *aapA3*-FL (full-length,  $\approx$  225 nt), IsoA3 ( $\approx$  80 nt) and IsoA3-p ( $\approx$  50 nt). **(D)** A mild IsoA3 stabilization observed in absence of RppH activity reveals the presence of a short species (IsoA3-s,  $\approx$  50 nt). Lower exposition image of the membrane shown in (C).

---

#### 4.4.2. Discussion

We have previously demonstrated that IsoA3 is an essential sRNA for *H. pylori* (see Chapter 4 section 4.2). Considering this, it is expected that the post-transcriptional regulation of this sRNA is tightly controlled. Indeed, the levels of sRNA antitoxins are often post-transcriptionally controlled by RNA decay, as shown for the Sok sRNA, controlled by the 5' end state sensitive RNase E in *E. coli* (Gerdes, 2016), as well as for the RatA sRNA controlled by RNase Y and RNase J1 in *B. subtilis* (Durand et al., 2012). Interestingly, a stabilization effect of the IsoA3 sRNA in a *H. pylori* deletion strain for the phyrophosphohydrolase *rpph* gene was previously observed (Bischler et al., 2017), indicating the existence of a 5'-exonuclease-mediated RNA decay pathway. Here, we were able to show that the IsoA3 sRNA decay is mediated by the 5'-3' exonuclease and endonuclease activity of RNase J in a 5' phosphate state-sensitive manner. These results highlight the importance of the 5' end state in mRNA decay, as previously shown, for instance, in *E. coli* for RNase E activity (Mackie, 1998) and in *B. subtilis* for RNase J1/J2 activity on mRNAs from the T-box family (Even et al., 2005).

Interestingly, we observed that RNase J endonucleolytic activity on IsoA3 RNA was stronger than the 5'-3'-exonucleolytic activity, even when the 5' end was monophosphate. More specifically, it was mainly directed to a single-stranded region named IsoA3 linker region (LK) (as previously described for the IsoA1 sRNA, Arnion et al., (2017)). These results reflect a strong sensitivity of the RNase J to secondary RNA structure, line with previous studies (Even et al., 2005; Mackie, 1998), and support the possible exploitation of RNase J for RNA structure *in vitro* studies, as the one performed by Daou-Chabo and Condon (2009).

Strikingly, the IsoA3 LK region is involved in the binding to the AapA3 mRNA and may also be important for the locus-specificity interaction between the different IsoA sRNAs and their toxin mRNA counterparts in the *aapA*/IsoA family of type I TA systems in *H. pylori*. Thus, we propose that RNase J endonucleolytic cleavage of IsoA3 leads to its direct functional inactivation. This kind of antitoxin inactivation mechanism has previously been reported for the Sok sRNA, mediated in this case, by the RNase E in *E. coli* (Dam Mikkelsen and Gerdes, 1997). The similarities of these mechanisms are intriguing, and, considering the pattern building tendency of nature, it may reflect the existence of such regulatory events in many other type I TA systems.

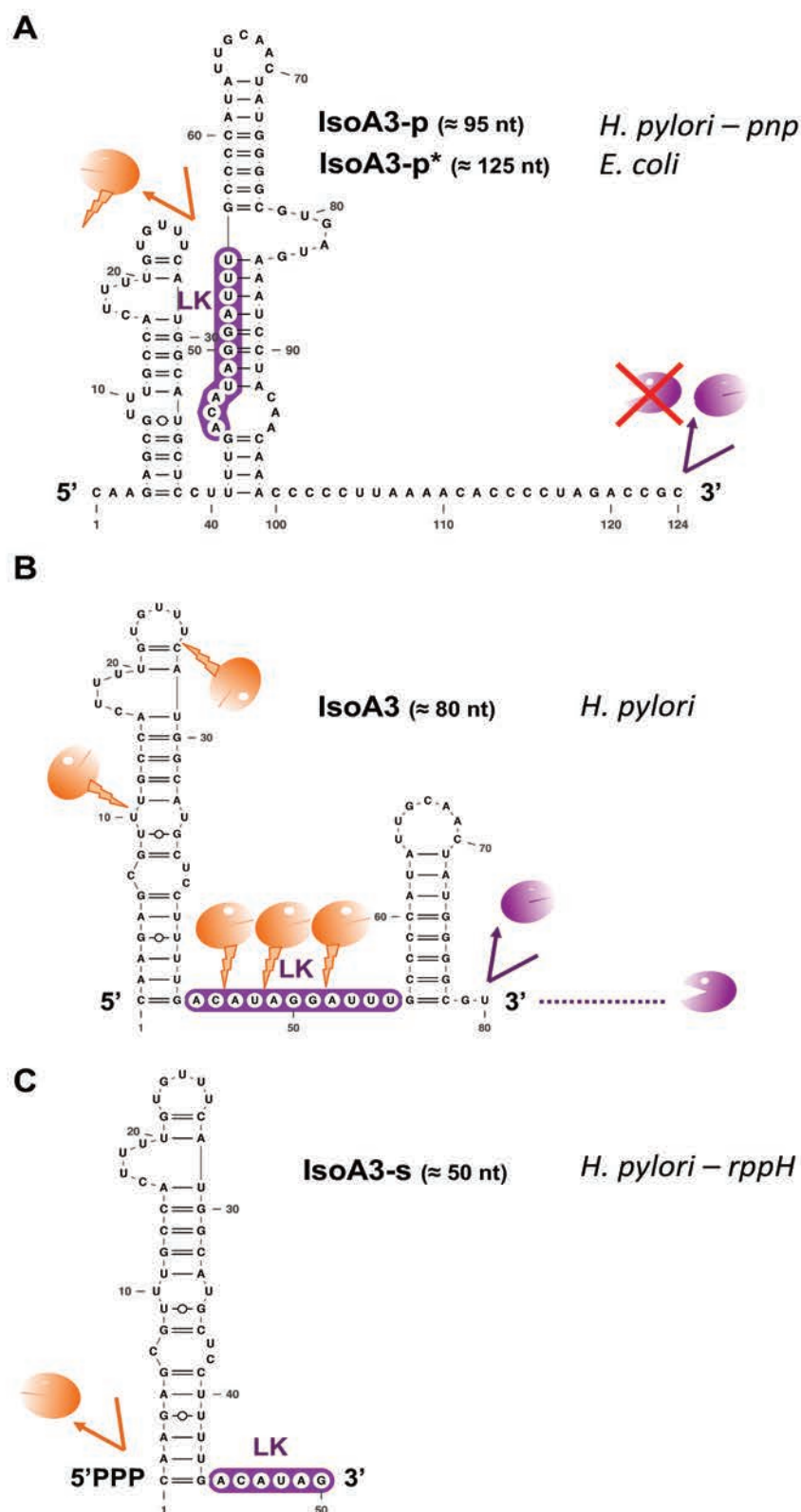
Most Gram-negative bacteria possess a so-called RNA degradosome, a multiprotein complex involved in mRNA degradation. This complex is principally composed by the 3'-5'-exonuclease PNPase, the RNA helicase RhlB, and an enolase, all assembled on RNase E ((Miczak et al., 1996; Py et al., 1996), for a review see (Carpousis, 2007)). However, the existence of an RNA degradosome in bacteria species lacking RNase E (*e.g.*, *H. pylori*) remains an open question. A recent study pointed at the existence of a minimal RNA degradosome in *H. pylori* composed by RNase J and the helicase RhpA (Redko et al., 2013). However, we did not observe any effect on AapA3 or IsoA3 RNA decay in a RhpA deletion strain (Figure 8.2.5 in the Appendix), indicating that the AapA3 mRNA 3'-exonucleolytic activation pathway, as well as the IsoA3 sRNA 5'-exonucleolytic decay, do not require helicase activity.

Interestingly, we show that PNPase takes part in the IsoA3 sRNA antitoxin maturation pathway, in a similar manner as previously shown for the AapA3 mRNA 3' end exonucleolytic activation. The adoption of this kind of parallel strategies is intriguing, and may represent a

biologically relevant mechanism helping to balance the pools of functional toxin mRNA and antitoxin sRNA in the cell. Furthermore, the observation of a PNPase-mediated IsoA3 maturation is interesting, as PNPase has been previously associated to the degradation of sRNAs that do not interact with Hfq, and which decay occurs in a target- and RNase III- independent manner (e.g., MicA, GlmY, RyhB and SgrS, [Andrade et al., 2012](#)), and this is the case for the IsoA3 sRNA, which *aapA3*-Tr-dependent RNase III-mediated degradation has little impact on its global levels (see [Chapter 4 section 4.2](#)). Nevertheless, other studies have reported a protective effect of PNPase on a Hfq-dependent manner against other RNases such as RNase E ([De Lay and Gottesman, 2011](#)), or even further, pointed to a reprogrammable PNPase activity, having a protective role when forming a ternary complex together with the RNA chaperone Hfq and degradative in absence of Hfq ([Bandyra et al., 2016](#)). Thus, it is possible that part of the reduction on the mature IsoA3 RNA levels in the absence of PNPase comes from the de-protection from other RNases, notably RNase J, which we know is involved in IsoA3 turn-over (see the [next discussion section](#)). Nevertheless, although the latter hypothesis cannot be discarded, the observation of the accumulation of a longer IsoA3 form (IsoA3-p) in absence of PNPase, clearly demonstrates that the major reason for the reduced levels of mature IsoA3 RNA in absence of PNPase is that PNPase is involved on its maturation. And even further, *H. pylori* lacking Hfq, as well as any other characterized RNA chaperone, it remains an open question whether a protective activity of PNPase could be observed independently of any ternary partner.

[Figure 4.4.3](#) summarizes the here obtained information about IsoA3 sRNA decay. Through the study of the deletion strain for the *pnp* gene, we uncovered the accumulation of a long IsoA3 species (IsoA3-p), which demonstrated that PNPase is involved in the 3' end maturation of the IsoA3 sRNA ([Figure 4.4.3A](#)). Secondary structure analysis of the IsoA3 RNA reflected the existence of a highly stable stem-loop on its 3' end. The fact that the PNPase stops degrading at that position indicates that the 3' stem-loop is a roadblock, acting as an energetic shield for the 3'-exonuclease activity of PNPase ([Figure 4.4.3B](#)). Additionally, *in vitro* RNase J cleavage assays revealed that the linker region (LK), located between the 5' and the 3' termini stem-loops of the IsoA3 sRNA, represents the major target for the RNase J endonucleolytic activity ([Figure 4.4.3B](#)). Furthermore, Northern Blot analysis of a *rppH* deletion strain revealed the accumulation of a ~50 nt-long IsoA3 cleavage product, corresponding to the major IsoA3 RNase J endonucleolytic product (cleaved in the LK). This result allowed us to demonstrate that RNase J senses the 5' phosphorylation state as part of the regulation of the IsoA3 decay pathway ([Figure 4.4.3C](#)).

Nevertheless, although the two stem-loop structures present in the IsoA3 sRNA may be reminiscent to the 5' and 3' protective stem-loops observed in the AapA3 mRNA, it is important to note that the IsoA3 sRNA half-life is about 15 minutes, while that of the AapA3 mRNA is around 10 times longer. This difference may be the result of several factors. However, taking a closer look to the stability of the termini stem-loops in both RNAs, while the 3' end stem-loops on both RNAs have similar stabilities (-13.80 kcal/mol and -13.70 kcal/mol, 3' stem-loop in the IsoA3 and AapA3 RNAs, respectively), the 5' stem-loop in the IsoA3 sRNA is more unstable (44 nt, -10.80 kcal/mol) than that of the active AapA3 mRNA (25 nt, -13.30 kcal/mol). Hence, it is tempting to think that this difference in the stability of the 5' end stem-loop may be behind the stronger RNase J activity in the IsoA3 sRNA. Nevertheless, the results of this work also indicate that IsoA3 decay is strongly determined by the RNase J endonucleolytic activity. Indeed, the observation of the IsoA3-p in the *H. pylori* -*pnp* strain and the IsoA3-p\* in *E. coli* (having an inactive PNPase against IsoA3 substrate), demonstrates that these long IsoA3 species are highly protected against RNase J endonuclease activity thanks to the occlusion of the LK region within the double-stranded region of the stem-loop (see LK in [Figure 4.4.3A](#)).

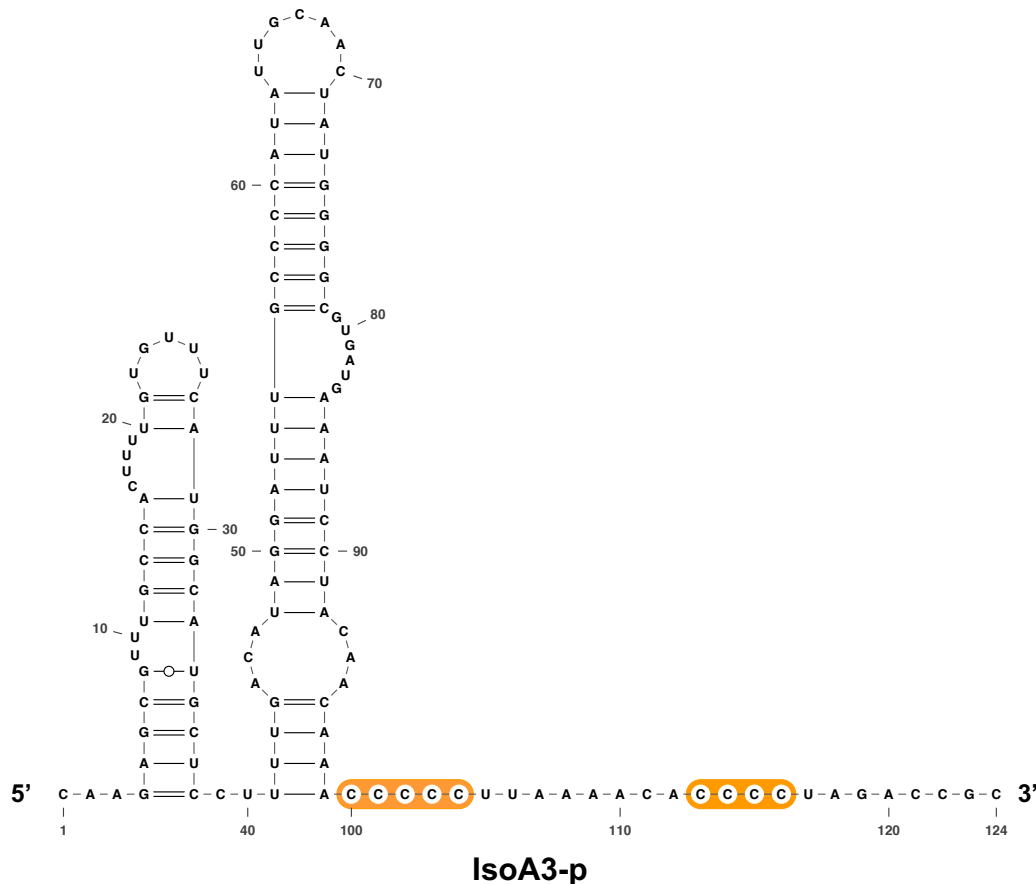


**Figure 4.4.3. Model of IsoA3 maturation and decay.** (A) In absence of PNPase, a long IsoA3 form is observed (IsoA3-p,  $\approx 95$  nt) indicating that PNPase is the major actor of IsoA3 RNA maturation. A similar pattern is observed when IsoA3 is expressed in *E. coli* (IsoA3-p\*,  $\approx 125$  nt), indicating that PNPase activity on IsoA3 substrate is weaker compared to that of *H. pylori*. (B) In wild-type *H. pylori* conditions, IsoA3-p is matured from its 3' end by the 3'-5' exonuclease activity of PNPase until it reaches the stable 3' stem-loop structure and generates the major and most abundant IsoA3 species (IsoA3,  $\approx 80$  nt). The secondary structure of this species is composed by two stem-loops connected by a linker (LK) sequence. As the stem-



loop structures protect IsoA3 RNA from RNase J 5'-3' exonucleolytic activity, the LK sequence is the major target of RNase J endonucleolytic activity on IsoA3 RNA and key on its decay pathway. Importantly, the LK sequence in the IsoA3-p species is protected within the structure (see A). (C) In the absence of RppH activity, a short IsoA3 form (IsoA3-s,  $\approx 50$  nt) is observed. Its size corresponds to the RNase J most abundant endonucleolytic product, demonstrating that RNaseJ 5'-exonuclease activity on IsoA3 substrate requires a monophosphorylated 5' end. RNAfold (Gruber et al., 2008) was used for secondary structure prediction and VARNA RNA (Darty et al., 2009) for drawing.

Additionally, as in the case of the AapA3 mRNA, the 3' end stem-loop structure does not act as a Rho-independent (*i.e.*, intrinsic) transcription terminator in the IsoA3 sRNA. However, a closer look at the 3' end sequence of the IsoA3-p sRNA revealed a potential Rho-binding site (C-rich sequence highlighted in orange, Figure 4.4.4), which is phylogenetically highly conserved between *H. pylori* strains (data not-shown). Rho is a homohexameric protein with RNA-dependent ATPase and helicase activities (Dombroski and Platt, 1988). Rho binds to poorly-structures C-rich sequences on ribosome-free RNAs, occupying a region of 70-80 nt (Richardson and Richardson, 1996; Zhu and von Hippel, 1998). Further experiments will be required to confirm or discard the role of Rho in IsoA3 transcription termination, and to understand the possible mechanistic consequences of it in the regulation of the *aapa3*/IsoA3 system. Interestingly, Rho has been proposed to be essential in *H. pylori* (Castillo et al., 2007). This phenomenon will make the experimental validation challenging, but, at the same time, it opens the exciting question of whether its essentiality could be linked to the proper transcription termination of the essential sRNA antitoxin IsoA3, or of any other of the highly conserved IsoA sRNAs in *H. pylori*.



**Figure 4.4.4. Putative single-stranded Rho-binding sites on the IsoA3-p sRNA.** RNAfold (Gruber et al., 2008) was used for secondary structure prediction and VARNA RNA (Darty et al., 2009) for drawing. Putative Rho-binding sites (C-rich sequences) are highlighted in orange.

## CHAPTER 5. Conclusion and outlook

### 5.1. Conclusion

The work presented in this thesis provided unpredictable insights into the regulation of the *aapA*/IsoA family of chromosomally-encoded type I TA systems in *H. pylori*. We developed a new approach, called FASTBAC-Seq, that uses lethality as selection pressure to identify loss-of-function mutations (toxicity) in toxin-encoding genes belonging to type I TA systems. Furthermore, as lethality is achieved by antitoxin promoter inactivation, FASTBAC-Seq allows the identification of gene expression determinants of the toxin gene component that are beyond their ‘primary regulator’, the RNA antitoxin. The coupling of such genetic selection of suppressors to Next-Generation Sequencing provides a high-throughput dimension to the approach, and allows the identification with nucleotide resolution of a great number suppressors, mainly limited by the homologous recombination rate of the bacterium. Thanks to the high selection pressure that lethality imposes, the use of mutagenesis techniques is not necessary for the highly-recombinant bacterium *H. pylori*. Indeed, this work demonstrated that the error rate of the DNA polymerase used to generate the PCR products for transformation provides sufficient mutations as to obtain a good coverage of the TA locus if recombination is not a limiting factor.

The use of FASTBAC-Seq to characterize the *aapA3*/IsoA3 system revealed a myriad of *cis*-encoded gene expression determinants in the AapA3 toxin-encoding mRNA. This study highlighted that type I RNA antitoxins do not act alone to regulate toxin expression. Instead, they act in coordination with a battery of regulatory elements embedded in the toxin-encoding mRNA. This observation makes sense, as contrary to the other antitoxin types (with the exception of type III), type I antitoxins can only act at the post-transcriptional level. There are several blind-spots for type I antitoxin action (*e.g.*, co-transcriptional translation and degradation, mRNA decay and diffusion rates determining mRNA translation) that are covered by *cis*-acting regulatory elements embedded in the toxin-encoding mRNA (*e.g.*, co-transcriptional folding, 5’ and 3’ end stabilizer structures and inert folding of primary transcript). The study of suppressor mutations in the AapA3 toxin-encoding mRNA showed that mRNA folding plays a central role in the regulation of the AapA3 toxin expression. Indeed, mRNA folding rules both, AapA3 translation and mRNA decay, two major determinants of gene expression in bacteria.

This thesis has highlighted the role of co-transcriptional folding to avoid both, co-transcriptional mRNA translation and degradation, in the context of absence of compartmentalization found in bacteria. AapA3 co-transcriptional translation is impeded by the formation of so-called metastable structures, which are transient SD-occluding hairpins that are essential to avoid premature toxin translation, and probably also premature antitoxin binding.

Moreover, the results of this work demonstrated that a functional mRNA stability is the major gene expression determinant of the AapA3 mRNA, as suppressors destabilizing the mRNA impede the achievement of enough AapA3 translation efficiency as to reach the threshold of toxicity. The active AapA3 mRNA possesses two highly stable termini stem-loop structures that act as energetic-shields against exonuclease activity and ensures its functional stability. The study of the destabilizing mutations located in such protective stem-loops shed light on the actors of mRNA decay in *H. pylori*, and how RNA secondary structure affects their activities. We demonstrated that the AapA3 translational activation is achieved by 3’-5’-exonucleases. Indeed, the 3’ protective stem-loop, previously thought to act as an intrinsic terminator, is instead essential for the 3’ end maturation of the

AapA3 mRNA. This work has demonstrated that the AapA3 mRNA is transcribed passed the 3' stem-loop, and it is then processively degraded from its 3' end back to the 3' stem-loop roadblock. We show that the 3'-5'-exonuclease PNPase is the major actor in the AapA3 activation pathway. However, the mRNA activation can also be achieved by one or more, so far unidentified, PNPase redundant activity(ies) (see section 5.2.3 of this Chapter for further discussion).

Importantly, spotting RNA folding in the focus of how this TA system is regulated, allowed the identification of similar regulatory mechanisms in other previously studied systems, reflecting the evolution of somehow similar strategies to overcome the high selection pressure of toxicity. Such mechanisms, although strongly dependent in the TA operon organization, follow clear patterns, that may help the future characterization of *de novo*-identified type I TA systems (see Masachis and Darfeuille (2018) in the section 1.4.1 of Chapter 1). TA systems appear to move horizontally and vertically. Type I TAs are believed to be inherited vertically by gene duplication in specific lineages, while type II and III are thought to move by horizontal gene transfer (HGT) (Fozo et al., 2010). A recent study showed that type I TAs seem to be represented in a narrower phylogenetic range than type II (*i.e.*, all toxins being represented in only one phylum, except SymE, which is in four) (Coray and Wheeler, 2017). The fact that the only type I TA having a larger representation is the only one using an antitoxin that is not a sRNA (the DNase RalA was not included in the study), is very informative and reflects the difficulty to computationally identify sRNAs, and by extension, type I TAs (see review (Coray and Wheeler, 2017)). Similarly, although both, type II and III are thought to be prone to HGT, type III present a much narrower distribution (*i.e.*, mainly found in Firmicutes and Fusobacteria) (Coray and Wheeler, 2017). This is probably not a coincidence since, as type I, type III TAs use RNA antitoxins. The results of this thesis may provide hints for the future optimization of computational prediction tools for type I TAs in bacterial genomes, by incorporating RNA folding constraints (*e.g.*, SD sequestration, double-stranded ends, etc.) in addition to the already used protein folding and conservation. Indeed, our group has recently developed a new type I TA database (T1tdb), using RNA structure based predictions that revealed the existence of an unanticipated massive number of type I TAs in several bacteria species, including *H. pylori* (visit our database in <https://d-lab.arns.cnrs.fr/t1tdb>).

Furthermore, this study has shown that a single point mutation in the AapA3 toxin-encoding gene is able to abolish toxicity. However, phylogenetic studies demonstrate that it happens very rarely, and when it does, the whole TA locus appears degenerated (data not-shown). These results, highlight the importance of the co-evolution of genes composing TA systems. Indeed, it seems intuitive to think that the existence (and conservation) of toxin genes rely on that of their antitoxin counterparts, as its loss would lead to cell death. However, it is just as important (though less intuitive) to understand that the conservation of antitoxin genes strongly depends on that of their toxin counterparts, as the loss of the toxin gene would immediately result in the loss of the selection pressure to maintain the antitoxin functionality. Thus, both genes are under an 'interdependent-existence' selection pressure.

Several studies have explored the use of TA systems as alternative antimicrobial strategies (due to their self-killing activity and their wide distribution in bacterial genomes) (Williams and Hergenrother, 2012). However, most of them target the toxin-antitoxin interaction or the antitoxin levels to induce the toxin production. The here identified *cis*-acting mRNA structural regulatory elements may represent new targets for the future development of alternative antimicrobial strategies. For instance, identifying small molecules able to impede the formation of the metastable SD-sequester structures in the toxin-encoding mRNA, would lead to premature toxin synthesis and cell death, in an antitoxin independent manner. Additionally, some of the traits of the toxin-encoding mRNA could be exploited with biotechnological purposes. For instance, the highly stable termini stem-loop structures could be used as portable mRNA stabilizing elements to enhance protein production, as recently

proposed upon the screening of a library of 5'UTRs based on the OmpA mRNA (Viegas et al., 2018) and six randomized 5'UTR libraries aiming at gene product synthesis optimization (Kang et al., 2018).

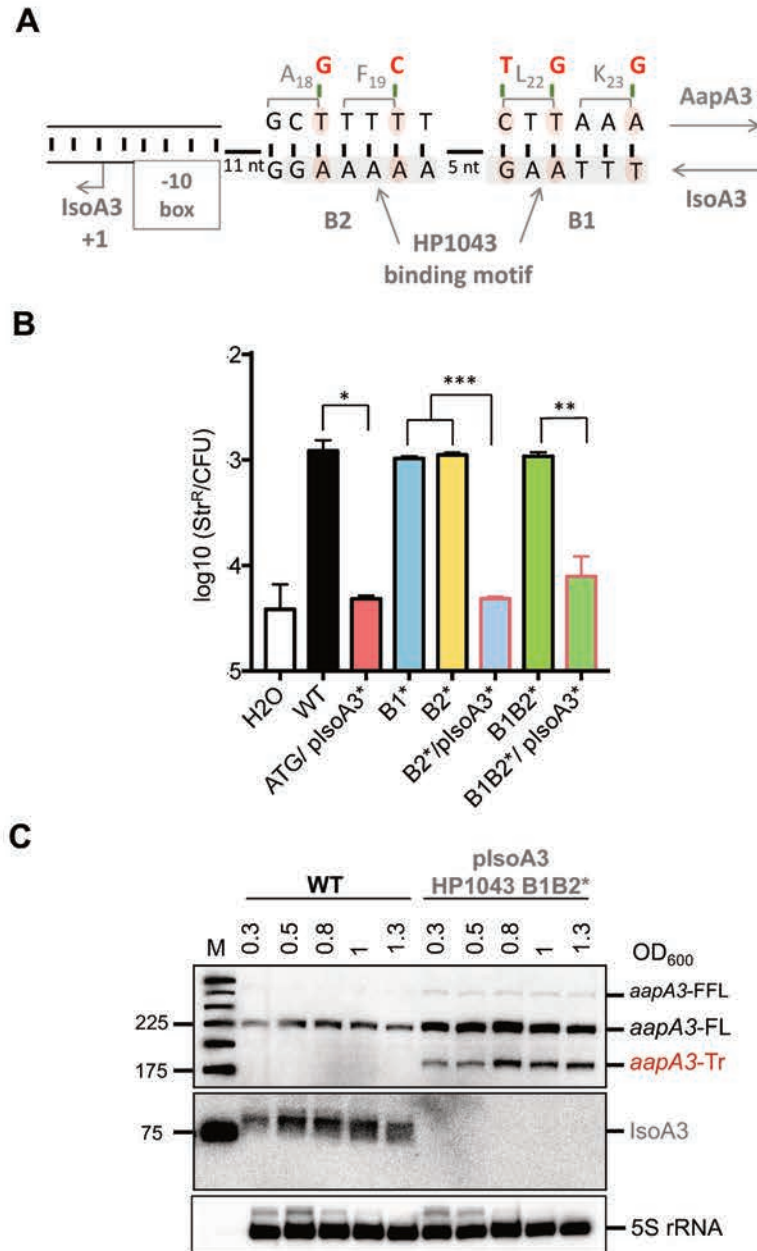
## 5.2. Open questions

### 5.2.1. Transcriptional regulation of the *IsoA3* sRNA?

The results presented in Chapter 4 section 4.2 (Paper III), demonstrated that the endogenous AapA3 toxin expression is constitutive and lethal in *H. pylori*. Thus, the IsoA3 RNA antitoxin is expected to constitutively repress its translation under normal conditions. However, a recent genome-wide characterization of DNA-binding sites of the essential orphan response regulator HP1043 (Pelliciari et al., 2017), revealed a binding motif in the IsoA3 promoter region (Figure 5.1A).

The HP1043 regulator belongs to the OmpR family and has a highly degenerated receiver sequence that cannot be phosphorylated. To date, only three HP1043 genomic binding sites have been characterized, 1) its own promoter, 2) the promoter region of the *tlpB* chemotaxis protein-encoding gene, and 3) the promoter region of the CncR1 sRNA, encoded within the Cag-PAI (Delany et al., 2002; Vannini et al., 2016). The characterization of the role of HP1043 is challenging as the gene has been proven to be essential (Delany et al., 2002). However, some studies relate it to cell cycle regulation (Delany et al., 2002) and homeostatic (Bauer et al., 2013) or oxidative stress control (Olekhovich et al., 2014). To investigate whether HP1043 was playing a role in the transcriptional regulation of IsoA3, we generated mutant strains on which one or both of the HP1043 binding motifs were mutated (playing with the genetic code to conserve the AapA3 ORF) (B1 and B2, Figure 5.1A). If HP1043 was a transcriptional activator of IsoA3, the mutation of the HP1043 binding sites on IsoA3 promoter region was expected to be lethal, as IsoA3 RNA would no-longer be produced to counteract AapA3 toxicity.

As shown in Figure 5.1B, the mutation of neither one or both HP1043 binding motifs on IsoA3 promoter region was lethal unless the -10 box was also inactivated (pIsoA3\*, as described in Chapter 4 section 4.1 and 4.2), and no suppressor mutations in the TA locus were observed in any of the several sequenced strains. This result indicated that HP1043 is not a transcriptional activator of IsoA3 expression, however, Northern Blot analysis revealed something puzzling. As shown in Figure 5.1C, IsoA3 RNA is absent in the strain mutated for both HP1043 binding motifs on IsoA3 promoter region (single mutants were not analyzed). Furthermore, the strain presents an OD-dependent accumulation of the translationally active *aapA3*-Tr mRNA species which, we know, is only present in absence of IsoA3 (as shown in Chapter 4 section 4.2), strongly suggesting that HP1043 is a transcriptional activator of the IsoA3 sRNA. However, this result was unexpected, and the reason behind the viability of the strain despite accumulating a translationally active toxin-encoding mRNA remains a mystery. One possibility could be the apparition of spontaneous suppressor mutations somewhere else in the genome, a future genome-wide sequencing of several of these HP1043 B1B2\* isolated strains will provide some answers. Another option could be that the HP1043 also acts as a transcriptional activator of one of the components on the AapA3 toxicity pathway. A future RNA-seq or proteomic comparative analysis could provide an answer, potentially shedding light into the molecular underpinnings of AapA3 toxicity.



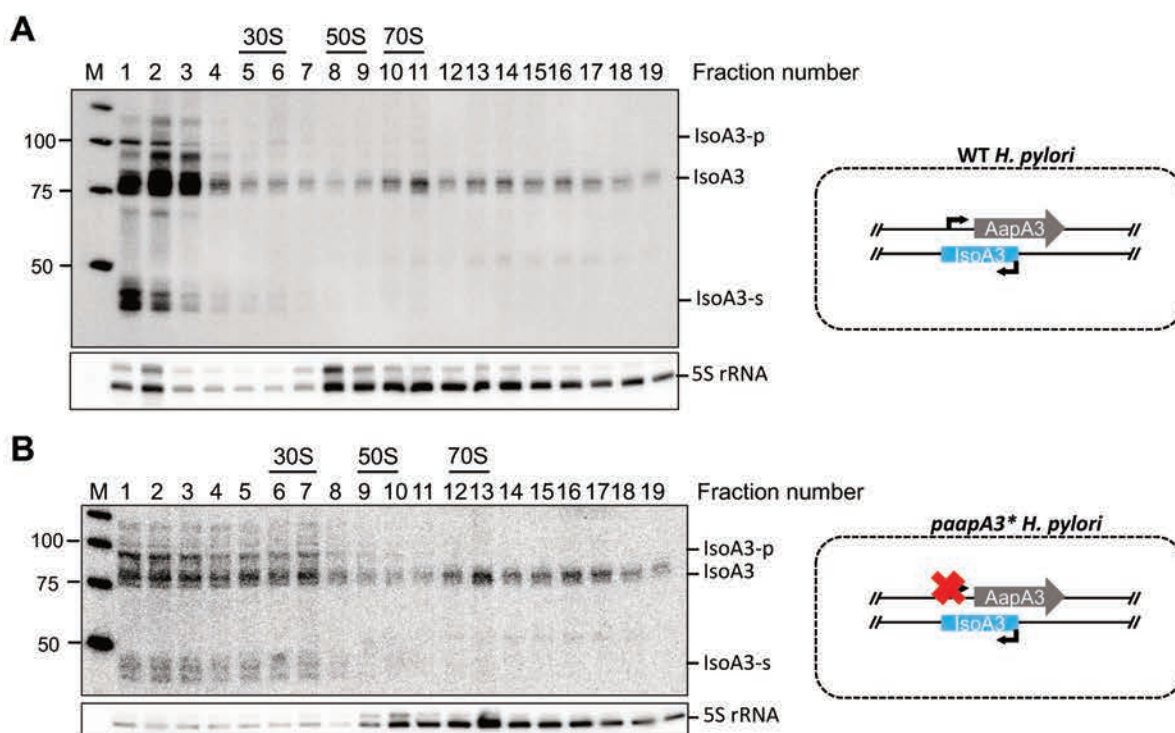
**Figure 5.1. The orphan response regulator HP1043 might act as a transcriptional activator of *IsoA3*.** (A) HP1043 binding motifs on *IsoA3* promoter region. Point mutations used to inactivate the binding motif are shown in red. Positions are shown relative to *aapA3* Transcriptional Start Site (TSS). (B) Number of StrR CFU/total CFU upon transformation with the indicated PCR-generated constructs or H<sub>2</sub>O (no-DNA control) was calculated. Error bars represent *s.d.*; *n*=3 biological replicates. (\* *p*=0.14; \*\* *p*=0.058; \*\*\* *p*=0.032, according to unpaired t-test). (C) Total RNA was isolated from cultures at the indicated optical densities and subjected to Northern Blot analysis.

Another interesting observation is that the levels of *aapA3*-FL and -FFL mRNA species are increased in the absence of HP1043 binding motifs on *IsoA3* promoter region (Figure 5.1C). This could be due to the lack of *IsoA3* transcription, suggesting that there might be a polar effect under normal conditions that could represent an extra element ensuring low AapA3 toxin expression. Another option could be that the binding motifs on *IsoA3* promoter region been mutated, HP1043 artificially binds more often to a 'binding motif-like' that is present in the *aapA3* promoter region (5' TGTTTTAAG 3', -32 to -24 from *aapA3* TSS). Further experiments will be required to understand and confirm, or not, this phenomenon.



### 5.2.2. Other IsoA3 targets?

In the Chapter 4, sections 4.2 and 4.3 (Papers III and IV), polysome fractionation experiments were performed to study the translatability of the different AapA3 mRNA species. Strikingly, Northern Blot analysis revealed the presence of IsoA3 sRNA in translating fractions in the wild type *H. pylori* 26695 strain (70S and polysomic (fractions 12 to 19) fractions, Figure 5.2A). This was unexpected as, the only so far known, IsoA3 target is the *aapA3*-Tr mRNA species, which cannot be visualized in the presence of IsoA3 due to its fast degradation by RNase III when in duplex (e.g., see Figure 8.1.4A in Appendix Paper III, which corresponds to the same membrane). To discard that the presence of IsoA3 in translating fractions was due to interactions with the AapA3 mRNA, a strain having inactivated the *aapA3* promoter (by a single point mutation in the last -10 box position, TAGGAT→TAGGC, fully inhibiting transcription, data not shown) was generated. Astonishing, the absence of the AapA3 mRNA did not alter the distribution of IsoA3 sRNA in the fractions (Figure 5.2B). This result is intriguing and could be explained in two ways, 1) IsoA3 interacts in *trans* with one or more other mRNAs (that are probably longer than AapA3 mRNA, as it is present in polysomic fractions and not only in monosomic and disomic), or 2) IsoA3 interacts, directly or indirectly, with a ribosomal component. Interestingly, only the ‘mature’ IsoA3 species (≈80 nt) is enriched in the translating fractions. The precursor IsoA3-p (≈100 nt) nor the RNA decay product IsoA3-s (≈50 nt) are present on those fractions, except for a very faint band of an intermediate size which is strongly enriched in polysomic fractions. This result suggests an RNA decay/function relationship for the IsoA3 sRNA, in a similar manner to what is observed for the AapA3 mRNA.



**Figure 5.2. IsoA3 sRNA is associated with translating ribosomes in a AapA3-independent manner.** Cell lysates of the wild-type (WT) (A) or AapA3 promoter inactivated (*paapA3\**) (B) *H. pylori* 26695 strains were subjected to ultracentrifugation through a sucrose gradient. A profile at OD<sub>254</sub> was recorded. RNA was extracted from each fraction and equal volumes of each extract were subjected to Northern Blot analysis. Fractions corresponding to the 30S and 50S subunits and 70S ribosomes are indicated are

indicated. The different transcripts are annotated as: IsoA3 (~80 nt), IsoA3-s (~50 nt), and 5S rRNA as loading control (5S rRNA).

The fact that a sRNA has multiple targets is not surprising. Indeed, several recent studies have placed sRNAs as central regulators of diverse physiological and adaptive responses (*e.g.*, bacterial communication, carbon metabolism or even phage-defense (Altuvia et al., 2018; Papenfort and Vogel, 2014; Svenningsen, 2018)). However, to date, it has never been reported in the case of type I antitoxins. Although an *in vivo* approach for the identification of IsoA3 sRNA targets (*e.g.*, RIL-Seq (Melamed et al., 2016), GRIL-Seq (Han et al., 2016), Ribo-Seq (Wang et al., 2015), MAPS (Lalaouna and Massé, 2015)) will have to be performed in the future, an *in silico* approach (TargetRNA2 software) predicted some potentially new IsoA3 targets. For instance, it is predicted to bind to the start codon region of a type I restriction enzyme M protein (hsdM) mRNA as well as to that of the glycyl-tRNA ligase subunit alpha (glyQ) mRNA. Other targets are components of the central metabolism and ammonia assimilation pathways, such as the acetyl-CoA synthetase (*acsA*, HP1045) and the aspartate ammonia-lyase (*aspA*, HP0649), predicted to be targeted by IsoA3 at the Shine-Dalgarno (SD) region (Figure 5.3, see Figure 5.5 for location of interaction region in IsoA3 sRNA). The identification of the two latter targets is intriguing. A reduced acetyl-CoA activity would lead to reduced levels of fumarate coming from the TCA (tricarboxylic acid) cycle. At the same time, reduced AspA activity would also result in reduced levels of fumarate coming from the ammonia assimilation pathways. Fumarate is an endogenous substrate for accepting electrons from NADH in the *H. pylori* anaerobic fumarate-respiration pathway and it is been proposed to reduce H<sub>2</sub>O<sub>2</sub> production from NADH (Chen et al., 1999). *H. pylori* possess a functional NADH-fumarate reductase (FRD), which has been shown to be essential for colonization of the mouse stomach (Ge et al., 2000) and proposed as a therapeutic target to treat *H. pylori* infections (Ge, 2002). In this context, if confirmed, IsoA3 interaction with the acetyl-CoA and AspA mRNAs would result in increased oxidative stress, reduced activity of the respiratory chain and reduced colonization potential.

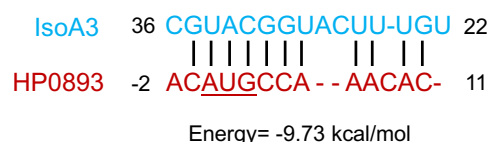


**Figure 5.3. IsoA3 *in silico* predicted interaction with the *aspA* and *acsA* mRNAs.** (A) Predicted interaction between the IsoA3 and the *aspA* and *acsA* mRNAs is shown. TargetRNA2 tool was used for prediction. AspA and AcsA Shine-Dalgarno (SD) region is underlined. Positions in the mRNAs are relative to the translation start codon.

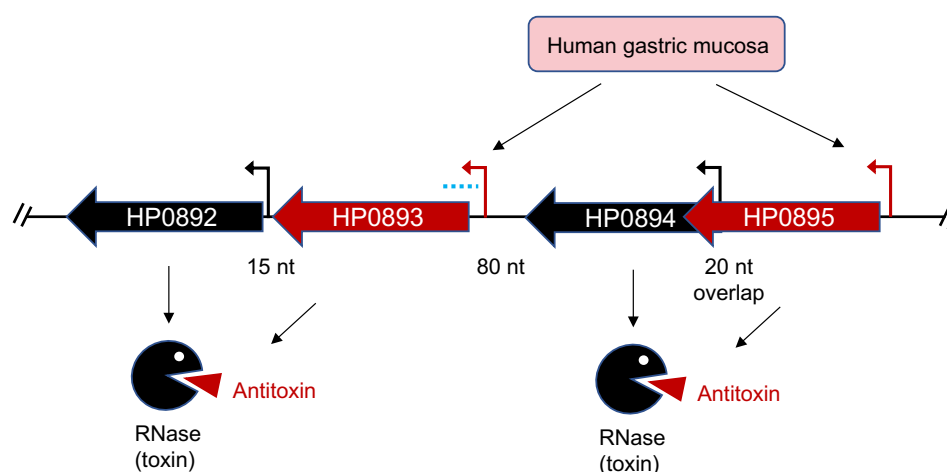
Another IsoA3 predicted target is the HP0893 gene. As shown in Figure 5.4A, the translation initiation region (TIR) of the HP0893 mRNA was predicted to be a binding site for the IsoA3 RNA (see Figure 5.5 for a location of the interaction region in the IsoA3 sRNA). The HP0893 gene belongs to a type II TA system of *H. pylori*. To date, only four type II TA systems have been characterized in the *H. pylori* 26695 strain: HP0315 (VapD-homologue)-HP0316 (Kwon et al., 2012), HP0894 (YafQ-homologue)-HP0895 (Pathak et al., 2013), HP0967-HP0968 (Cárdenas-Mondragón et al., 2016) and HP0892-HP0893 (Han et al., 2013; Im et al., 2014). All of them belong to the Vap (virulence-

associated protein) family, which is widely found in Bacteria and Archaea (Makarova et al., 2006). Interestingly, the HP0892-HP0893 TA module is found in close proximity (50 nucleotides downstream) to the HP0894-HP0895 module (Figure 5.4B), and the antitoxin transcription of both TA modules (HP0893 and HP0895) was shown to be induced upon interaction with the human gastric mucosa during infection (Graham et al., 2002; Han et al., 2011). Furthermore, the HP0892-HP0893 module is considered a marker of virulent strains, as its absence is associated with *cag*-PAI-negative strains (Terry et al., 2005).

A



B



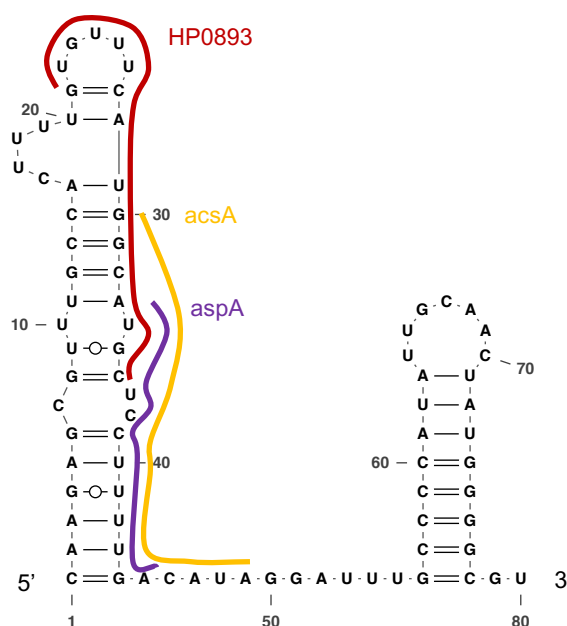
**Figure 5.4. IsoA3 *in silico* predicted interaction with the HP0893 mRNA.** (A) Predicted interaction between the IsoA3 and the HP0893 mRNA is shown. TargetRNA2 tool was used for prediction. HP0893 translation start codon is underlined. Positions in the HP0893 are relative to the translation start codon. (B) Schematic representation of the chromosomal organization of the HP0892-HP0893 and HP0894-HP0895 type II TYA loci of the *H. pylori* 26695 strain. IsoA3 predicted binding region in the HP0893 mRNA is indicated by a blue dotted line. The antitoxin promoter of both TA modules have been described to be activated upon interaction with the human gastric mucosa (Graham et al., 2002; Han et al., 2011).

The binding of IsoA3 to the TIR region of the HP0893 antitoxin gene would result in translation inhibition. Considering the constitutive IsoA3 expression, this would imply a constitutive repression of the HP0893 antitoxin. Although the constitutive inhibition of an antitoxin seem as a counterintuitive and unlikely role, regulation could be complex, involving, for instance, ternary actors (*e.g.*, chaperons) or specific conditions that modulate the ratio of IsoA3::HP0893 RNA molecules (*e.g.*, stress conditions found during infection). If confirmed, the role of IsoA3 as a translation inhibitor of the HP0893 antitoxin could be of allowing HP0892 ribonuclease activity during infection (when higher HP0893 transcription levels are expected (Graham et al., 2002; Han et al., 2011)), which might be directly or indirectly linked with *H. pylori* virulence. The IsoA3-mediated inhibition of the HP0893 translation could also serve to uncouple the action of the HP0894 and the HP0892 toxins



during infection. Additionally, if confirmed, it would represent the first example of TA ‘cross-type’ non-cognate interaction.

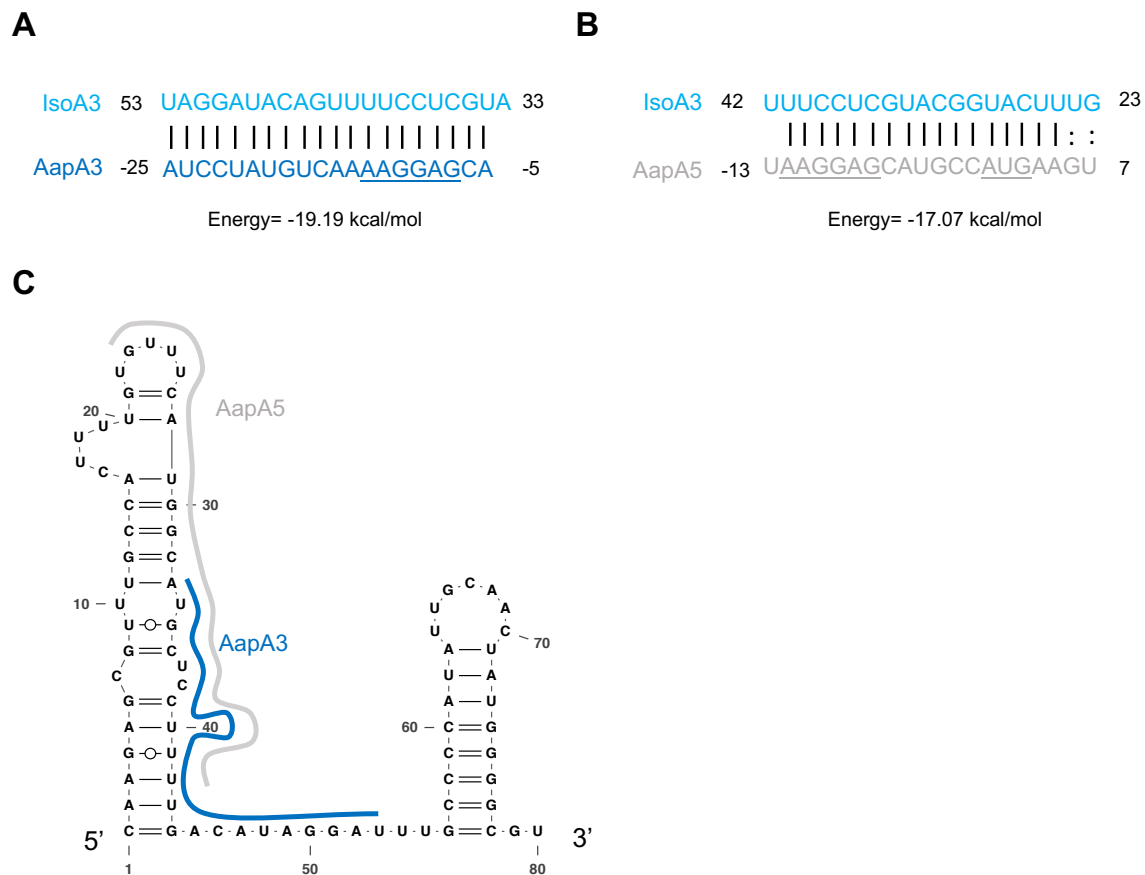
Interestingly, the predicted regions involved in the IsoA3 binding to the three potentially new IsoA3 targets are located in close proximity to each other, having overlapping regions mainly corresponding to the anti-start codon and the anti-SD sequences. (Figure 5.5). The binding region to the *aspA* mRNA is fully overlapping with that of the *acsA* mRNA (which has 58% overlap with that of the *aspA* mRNA). This is intriguing as it may represent a mechanism of competition between both mRNAs to bind to IsoA3, which may balance or keep IsoA3 inhibition effects within specific biologically relevant limits.



**Figure 5.5. Regions of IsoA3 sRNA interaction with its *in silico* predicted targets.** RNAfold (Gruber et al., 2008) was used for secondary structure prediction and VARNA RNA (Darty et al., 2009) for drawing.. Region involved in IsoA3 binding to the HP0893, *acsA* and *aspA* mRNAs are indicated by a red, an orange and a purple line, respectively.

Nevertheless, it is important to keep in my that these are only computational predictions and they can be vague and misleading when studying such highly-structured RNAs. For instance, the TargetRNA2 also predicts binding of the IsoA3 RNA to the TIR region of the AapA3 homolog encoded at the locus V (AapA5 toxin-encoding mRNA) (Figure 5.6). However, as this work has shown, the interaction of IsoA3 with the AapA3 mRNA is highly specific and strongly dependent on RNA structure, being even able to discriminate between the different mRNA species and exclusively target the 3' end processed form. More importantly, and in line with previous studies (Sharma et al., 2010), this thesis has demonstrated that despite the presence of six IsoA homologs in the *H. pylori* chromosome, none of them is able to avoid AapA3 lethality in the absence of IsoA3, reflecting a complete lack of cross-interaction between modules (see Chapter 4 section 4.2). Therefore, even though the interaction with the AapA5 mRNA is predicted, it does not occur *in vivo*. One possible hint can come from the observation of where the predicted interaction region locates in the IsoA3 RNA secondary structure. For instance, the interaction with the AapA3 mRNA comprises a large single-stranded region (the linker region), whereas that of the AapA5 mRNA mostly locates within a double-stranded region (stem-loop). Such intramolecular interaction (stem-loop) masking the binding sequence, would be unlikely replaced by the interaction with the AapA5 mRNA *in trans*. Considering this, the above-described predicted interactions, especially the one with the HP0893 antitoxin and the

aspA mRNAs are highly unlikely to take place *in vivo* and will need to be considered carefully and experimentally validated.



**Figure 5.6. IsoA3 predicted interaction with the AapA3 and AapA5 mRNAs.** Predicted interaction regions between IsoA3 and the AapA3 (A) and AapA5 (B) mRNAs is shown. TargetRNA2 tool was used for prediction. Translation start codon of AapA5 and Shine-Dalgarno sequences are underlined. Positions in the mRNAs are relative to the translation start codon. (C) Regions of interaction in the IsoA3 secondary structure prediction are shown. RNAfold (Gruber et al., 2008) structure prediction and VARNA RNA (Darty et al., 2009) for drawing. Region involved in the IsoA3 binding to the AapA3 and AapA5 mRNAs are indicated by a dark blue and a grey line, respectively.

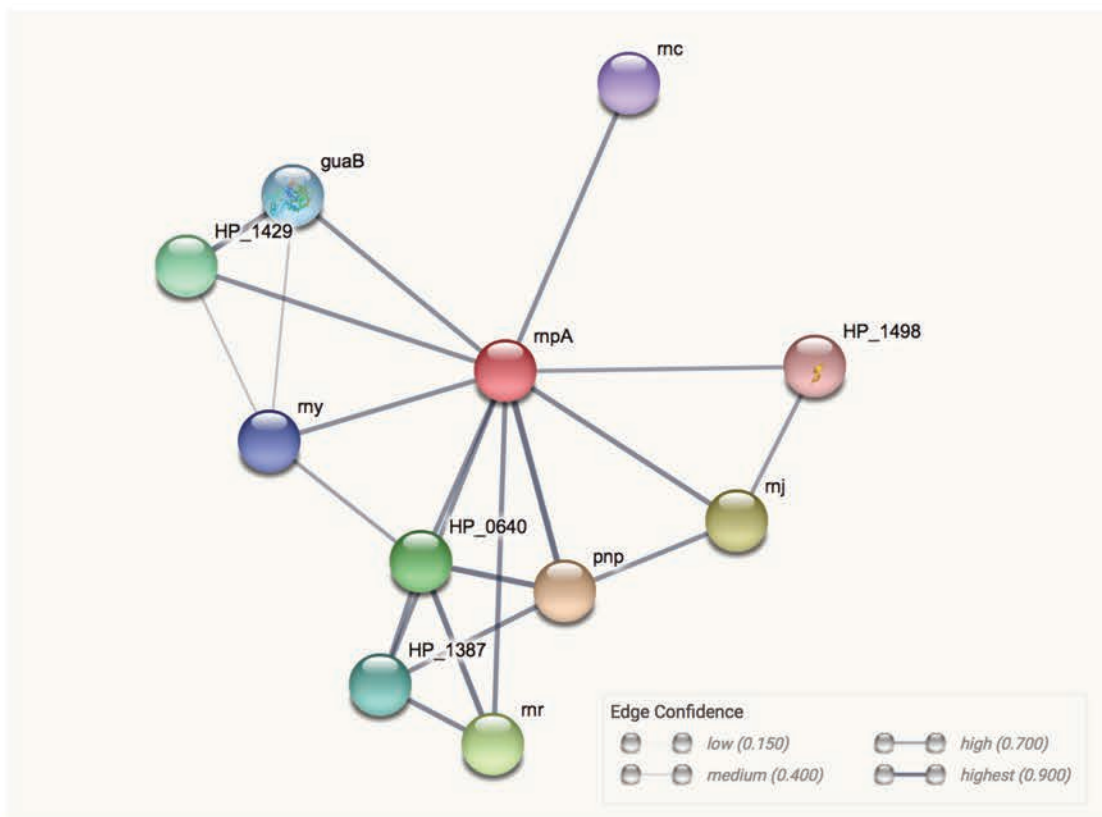
### 5.2.3. Missing actors in the AapA3 mRNA activation?

In Chapter 4 section 4.3 (paper IV), a 3' end nucleolytic activation of the AapA3 toxin-encoding mRNA was described. Our results demonstrated that the 3'-5'-exonuclease PNPase (HP1213) plays a major role, together with the endonuclease RNase Y (HP0760), and one or more, so far unidentified, redundant activities. We also demonstrated that this activation pathway is independent on the RhpA (HP0249) helicase activity as well as on the 5' end phosphorylation state. In our attempts to identify the redundant activities, we analysed AapA3 mRNA expression in different deletion backgrounds. Thereby, we discarded a potential role of the 3'-5'-exonuclease RNase R (HP1242) as well as of the membrane-associated nuclease NucT (HP0323) (data not-shown). An *in-silico* prediction using the STRING online tool revealed a possible interaction between the PNPase and the protein component of the endonuclease RNase P (*rnpA*, HP1448). The RNase P catalyzes the removal of the 5'-leader sequence from the pre-tRNA to produce the mature 5'-terminus (Li and Altman, 2003). Interestingly, it has been shown that either PNPase or RNase P is able to remove the Rho-independent transcription



terminator of the tRNA<sup>Leu5</sup>, by cleaving up to 1-3 or at 4-7 nucleotides, respectively, downstream of a CCA determinant (Mohanty and Kushner, 2010; Mörl and Marchfelder, 2001). Given the strong structural similarities between the AapA3 mRNA and tRNAs (cloverleaf structure), it seems tempting to think that indeed, RNase P may be the elusive activity. Even more, the STRING interaction network of the protein component of RNase P (*rnpA*, Figure 5.7) places it in a central position, being putatively involved in the interaction with several RNases, including the 3'-5'-exonuclease PNPase, the 5'-3'-exonuclease RNase J, the endonucleases RNase Y and RNase R, and the double-strand-specific endonuclease RNase III. RNase P represents a promising putative actor on the AapA3 mRNA activation pathway, therefore, further studies will be useful to confirm (or not) its role and understand the exact role of the putative interactions with the other RNases. However, as in the case of the RNase J, the RNase P is an essential protein for *H. pylori* (Chalker et al., 2001), making its *in vivo* characterization challenging. Nevertheless, assays such as co-immunoprecipitation may provide insights into the understanding of the interaction network.

Another interesting STRING association prediction with the *rnpA* is the poly(A) polymerase (HP0640, Figure 5.7). Most exonucleases are sensitive to RNA secondary structures, however, they can often overcome this obstacle in the presence of a single-stranded region, such a poly(A) tail, to which they can bind. For instance, it has been recently shown that the poly(A) polymerase (PAI I) plays a key role in ensuring sufficient decay of the 3' external transcribed spacer of the *glyW-cysT-leuZ* transcript (LeuZ<sup>3'ETS</sup>) in *E. coli*, which on its turn, ensures functional stability of the RhyB sRNA (Sinha et al., 2018). However, previous studies in our group have discarded a role of the poly(A) polymerase HP0640 in AapA3 mRNA 3' exonucleolytic maturation in *H. pylori* (data not-shown). Further studies will be required to uncover the identity of the elusive missing actors on the AapA3 mRNA activation pathway.



**Figure 5.7. STRING *rnpA* functional association network.** The online STRING tool (Szklarczyk et al., 2017) was used to generate the *rnpA* (RNase P) protein-protein interaction network in the *H. pylori* 26695

strain. Connection lines thickness indicates confidence as indicated at the bottom right. The following proteins are represented: RNase P protein component (*rnpA*, HP1448), RNase III (*rnc*, HP0662), integral membrane protein (*lptG*, HP1498), RNase J (*rnj*, HP1430), PNPase (*pnp*, HP1213), RNase R (*rnr*, HP1248), poly(A)polymerase (HP0640), DNA polymerase III subunit epsilon (HP1387), RNase Y (*rny*, HP0760), polysialic acid capsule expression protein (*kpsF*, HP1429), inosine 5'-monophosphate dehydrogenase (*guaB*, HP0829).

---

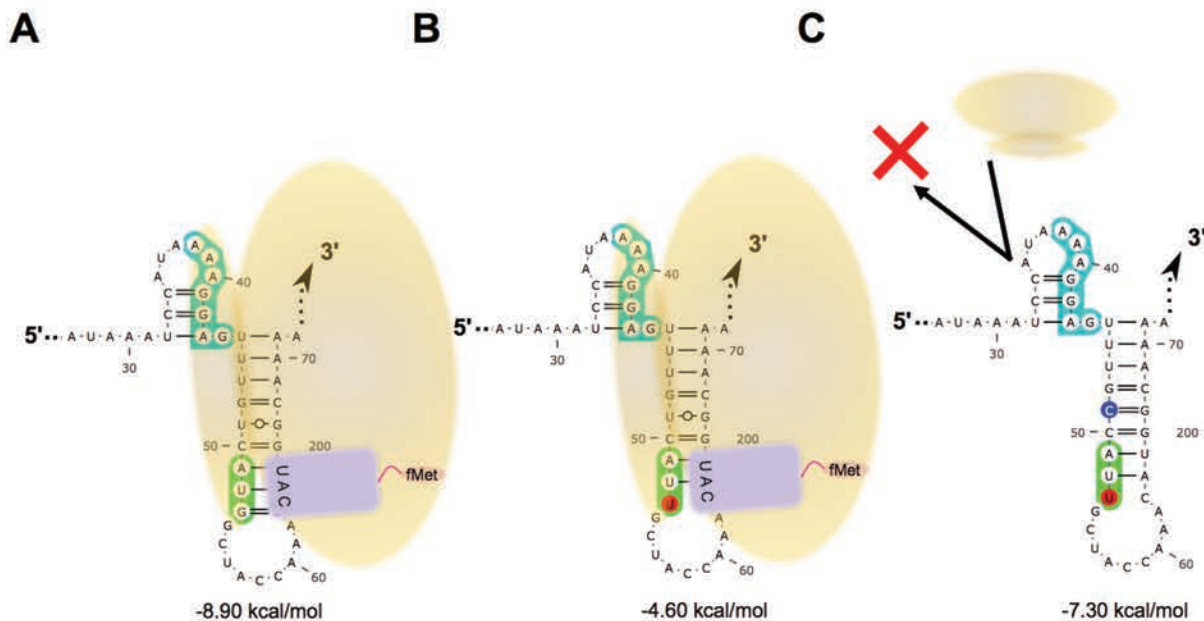
#### 5.2.4. Start codon accessibility, one more regulatory layer?

In Chapter 4 section 4.2 (paper III), the role of the co- and post-transcriptional Shine-Dalgarno (SD) sequence accessibility in translation initiation of the AapA3 toxin-encoding mRNA was discussed. Intriguingly, while studying the *aapA1*/IsoA1 system (*aapA*/IsoA module encoded in the chromosomal locus I of the *H. pylori* 26695 strain), we observed that, when trying to inactivate the AapA1 translation start codon as previously done for the AapA3 peptide (G to T single point mutation at the third codon position, AUG to AUU), it would still be toxic in absence of IsoA1. As shown in Figure 5.8A, the AapA1 translation start codon is fully base-paired within a highly stable stem-loop structure (energy = -8.90 kcal/mol). Interestingly, the start codon inactivation by the point mutation G53T (guanine at position 53 from the AapA1 TSS mutated to thymine) is predicted to cause a dramatic destabilization of such stem-loop structure (energy = -4.60 kcal/mol, Figure 5.8B). Furthermore, the analysis of suppressor strains able to compensate the toxicity of the start codon mutated AapA1, revealed the existence of a suppressor mutation, T49C (thymine at position 49 from AapA1 TSS mutated to cytosine; changing a GU Wobble pair into a GC pair), partially restoring the stem-loop stability (-7.30 kcal/mol, Figure 5.8C). These results are striking as they strongly suggest that the increased accessibility to the start codon thermodynamically prompts the efficient recognition of the AUU non-canonical start codon by the fMet-tRNA<sup>fMet</sup> despite the pairing of only 2/3 codon positions. This recognition triggers the subsequent 50S ribosomal subunit assembly to the 30S subunit and leads to translation initiation and the generation of a wild-type peptide (Figure 5.8B).

Nevertheless, although these results were surprising, several studies have highlighted that translation from non-canonical codons occurs more often than previously anticipated (Hecht et al., 2017; Romero and García, 1991), in many cases having important regulatory roles. For instance, the *tufA* (coding for the EF-TuA) and the *hupB* (coding for the EF-TuB) are among the most highly expressed genes in *E. coli* (one cell is estimated to contain 74,981 EF-TuA and 87,672 EF-TuB molecules) despite the EF-TuA being translated from a non-canonical GUG initiation codon (Lu et al., 2007). Another example is the *infC* gene (coding for the IF3), which is translated from the non-canonical AUU or AUC initiation codons and is also expressed at fairly high levels (5,488 molecules/cell) in *E. coli* (Lu et al., 2007). Indeed, it has been demonstrated that *infC* translation from AUU or AUC initiation codons represents an essential *cis*-acting element in the autogenous translational control of translation initiation by the initiation factor IF3 (Brombach and Pon, 1987; Butler et al., 1987; Sacerdot et al., 1996). The second example of chromosomal translation initiated at AUU is the *E. coli* *pcnB* gene, coding for the poly (A) polymerase (PAP), which is toxic when overproduced. Interestingly, it has been shown that initiation at AUU is a mechanism to limit the translation of the potentially toxic *pcnB* gene product (Binns and Masters, 2002). In line with this observation, a further example can be found within a type I TA system, as is the case of the YonT toxin-encoding mRNA. Surprisingly, the SD sequence is not sequestered in the YonT mRNA, but instead, it uses the non-canonical GUG initiation codon to strongly limit the translation initiation rate (Durand et al., 2012; Reif et al., 2018).



Another important aspect that our results have highlighted is that, the effect of non-canonical initiation codons in translation strongly depends on the structural context. Thus, mRNAs sharing the same initiation codon can present different translational efficiencies (e.g., the AUU AapA1 mutant stops being lethal when the stem-loop structure where it is located is stabilized by the T49C suppressor mutation, see Figure 5.8B and C). Previous studies have evidenced this phenomenon. For instance, the nature of the downstream region (DR), including the +2 codon immediately following the start codon, was shown to affect the initiation together with the SD sequence in a synergic manner, leading to a high translation efficiency of UUG initiation codons (Stenström et al., 2001). Similarly, a change of the AUG initiation codon to AUA in the phage T7 0.3 gene did not affect ribosome binding (Dunn et al., 1978), however, it drastically did to 30% in the Q $\beta$  coat protein (Taniguchi T et al., 1978).

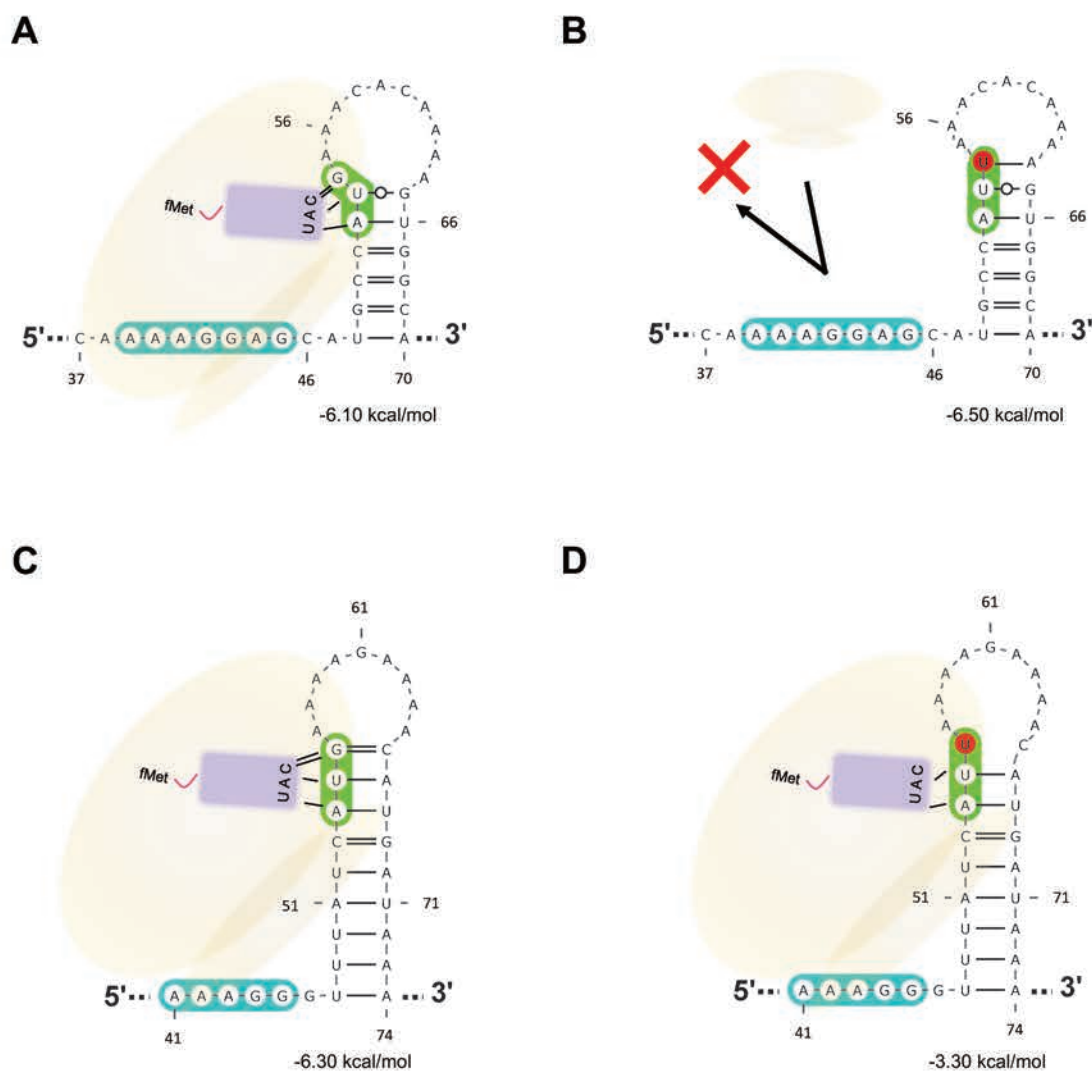


**Figure 5.8. Start codon accessibility determines the visibility of AapA1 translation initiation from non-canonical initiation codons.** (A) The translation start codon in the AapA1 toxin-encoding mRNA is sequestered within a stable stem-loop structure (energy = -8.90 kcal/mol). In this conformation, the mRNA is translatable and lead to lethality in absence of antitoxin. (B) When the start codon is mutated in the third codon position by changing the guanine at position 53 from AapA1 Transcriptional Start Site (TSS) to thymine (G53T), the stem-loop is destabilized (energy = -4.60 kcal/mol). Unexpectedly, the G53T *aapA1* mutant mRNA is translatable and induces lethality in absence of antitoxin. (C) Lethality of the G53A mutant mRNA translation can be abolished by the single point suppressor mutation T49C (thymine at position 49 from AapA1 TSS changed to cytosine), which partially restores the stability of the stem-loop (energy = -7.30 kcal/mol). Shine-Dalgarno (SD) sequence is highlighted in turquoise, translation start codon is highlighted in green, start codon mutation is highlighted in red, suppressor mutation is highlighted in blue. Ribosome is represented in yellow. fMet-tRNA<sup>fMet</sup> is represented in purple. RNAfold Web Server (Gruber et al., 2008) was used to predict the structures and minimum free energy values, VARNA RNA (Darty et al., 2009) was used for drawing.

Interestingly, when we look at the AapA3 toxin-encoding mRNA, we can see that in contrast to the AapA1, only the two first codon positions of the AUG translation initiation codon are basepairing within a stem-loop structure (energy = -6.10 kcal/mol, Figure 5.9A). Considering what the AUU start codon mutation in the AapA1 mRNA has taught us, we can understand that the AUU mutation in the context of the AapA3 mRNA (G54T, guanine at position 54 from the AapA3 TSS mutated to thymine) does not result in toxicity, as the stem-loop stability is unchanged, and thus, either



is its translation efficiency (energy = -6.50 kcal/mol, Figure 5.9B). However, it is also interesting to see, that for a given AapA mRNA, the context of the translation initiation codon changes depending on the strain. For instance, in the *H. pylori* G27 strain, the AapA3 AUG translation initiation codon is found in a context that is more similar to that of the AapA1 than to that of the AapA3 mRNA in the *H. pylori* 26695 strain (energy = -6.30, Figure 5.9C). Thus, it is tempting to assume, that the AUG to AUU mutation of the AapA3 translation initiation codon in the G27 strain may be toxic, as it strongly destabilizes the stem-loop structure and thereby increases the accessibility of the initiation codon (energy = -3.30, Figure 5.9D).

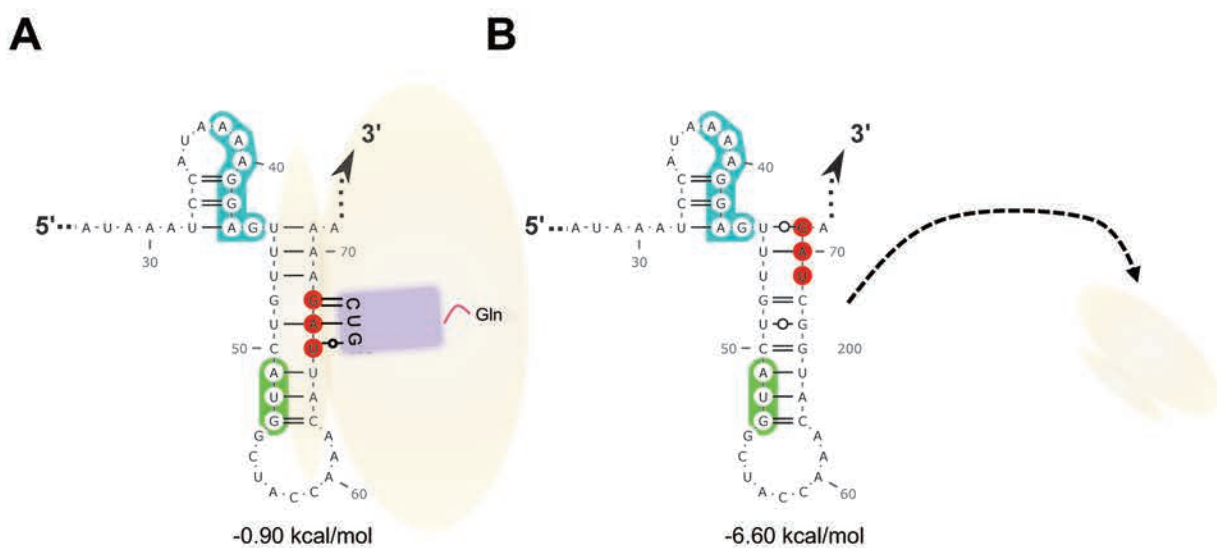


**Figure 5.9. Consequences of the stability of the start-codon harboring stem-loop in the AapA3 mRNA of two different *H. pylori* strains.** (A and B) Structure and stability of the wild-type (A) or mutated (B) AapA3 translation start codon harboring stem-loop in the *H. pylori* 26695 strain. (A) In the wild-type situation, only the two first codon positions are base pairing within a stem-loop with a minimum free energy of -6.10 kcal/mol. This mRNA is translatable and leads to toxicity in absence of antitoxin. (B) When the AapA3 start codon is inactivated by a single point mutation on the third codon position G54T (guanine at positions 54 from AapA3 TSS mutated to thymine), the stability of the stem-loop is not decreased, but rather slightly increased (energy = -6.50 kcal/mol). This mutated mRNA is not efficiently translated and does not lead to lethality in absence of antitoxin. (C and D) Structure and stability of the wild-type (C) or mutated (D) AapA3 translation start codon harboring stem-loop in the *H. pylori* G27 strain. (C) The stem-loop structure in the G27 strain resembles that of the AapA1 mRNA in the 26695 strain, thus, having all three bases of the start codon basepairing within a stable stem-loop structure. By



homology to the AapA1 mRNA structure, the AapA3 mRNA in G27 is expected to be translated. **(D)** On the contrary, the inactivation of the AapA3 translation start codon by the single point mutation G54T causes a strong destabilization of the stem-loop in the G27 strain, and by homology to what was observed in the AapA1 mRNA in the 26695 strain, the reduced stem-loop stability is expected to reveal start codon misreading events that would result in lethality in absence of antitoxin. Shine-Dalgarno (SD) sequence is highlighted in turquoise, translation start codon is highlighted in green, start codon mutation is highlighted in red, suppressor mutation is highlighted in blue. Ribosome is represented in yellow. fMet-tRNA<sup>fMet</sup> is represented in purple. RNAfold Web Server (Gruber et al., 2008) was used to predict the structures and minimum free energy values, VARNA RNA (Darty et al., 2009) for drawing.

Soon after the discovery of the *aapA*/IsoA family of type I TA systems in *H. pylori* (Sharma et al., 2010), studies aiming at the characterization of the different AapA small proteins identified an, at that time, enigmatic result (Cynthia Sharma, unpublished work). When the AapA1 toxin protein of the *H. pylori* 26695 strain was cloned into an *E. coli* arabinose-inducible overexpression vector and cells were grown in presence of arabinose, it expectedly induced toxicity. However, when a stop codon was introduced at the 6<sup>th</sup> AapA1 codon, toxicity was strikingly still observed, and this effect that was lost when the stop codon was placed at the 7<sup>th</sup> AapA1 codon. Now, taking into consideration the knowledge acquired from the study of the AUU start codon mutant and the T49C suppressor mutation, we can hypothesize that the translational stop codon misreading events (reflected in the form of a toxic phenotype) observed when the stop codon was introduced at the 6<sup>th</sup> position, were favored by the strong destabilization of the start codon-harboring stem-loop structure (energy = -0.90 kcal/mol, Figure 5.10A), which does not occur when the stop codon is introduced at the 7<sup>th</sup> position (energy = -6.60 kcal/mol, Figure 5.10B). Such misreading events were probably achieved by the UAG stop codon misrecognition by the CUG anticodon in the tRNA(Gln). Thus, despite involving a GU Wobble base pair, the translation initiation rate was sufficiently high (thanks to the high start codon accessibility) as to lead to toxicity through stop codon misreading events. Importantly, it was possible to observe toxicity despite the AapA1 peptide carrying the amino acid substitution Glycine at position 6 to Glutamine (Gly6Gln) because, as shown by Korkut et al., in preparation, and similarly to the AapA3 peptide (see case study in Chapter 4 section 4.1), the first nine amino acid residues in the N-terminal domain of the AapA1 peptide are not determinant for toxicity.



**Figure 5.10. Translation start codon accessibility determines the visibility of AapA1 translation stop codon misreading events.** (A) The introduction of a translation stop codon at the 6<sup>th</sup> AapA1 codon position strongly destabilizes the start codon-sequester stem-loop structure (energy = -0.90 kcal/mol) and



does not efficiently avoid AapA1 peptide synthesis, probably due to the stop codon recognition by a Glutamine tRNA through a GU Wobble base pair. **(B)** When a stop codon is introduced at the 7<sup>th</sup> AapA1 codon position, the stability of the stem-loop is not strongly affected (energy = -6.60 kcal/mol) and thus, possible stop codon misreading events are not reflected as toxicity in absence of antitoxin. Shine-Dalgarno (SD) sequence is highlighted in turquoise, translation start codon is highlighted in green, the newly introduced stop codon is highlighted in red. Ribosome is represented in yellow. Gln-tRNA<sup>Gln</sup> is represented in purple. RNAfold Web Server (Gruber et al., 2008) was used to predict the structures and minimum free energy values, VARNA RNA (Darty et al., 2009) for drawing.

Altogether, these results suggest that the AapA1 start codon-sequestering stem-loop structure is one more layer of regulation ensuring a limited AapA1 translation. From a physiological point of view, the use of start codon accessibility to dictate the translation efficiency of a toxic gene could represent a sensor mechanism. For instance, it could provide a fast response (lethality) to changes in ribosome concentration or temperature (as shown for other RNA thermometers, Kortmann and Narberhaus, 2012)). Apart from the toxicity observed in the transformation assays, we confirmed by *in vitro* translation assays and toeprint experiments the translation of the AUU mutant and the limited translation of the T49C suppressor. Additionally, polysome fractionation assays showed a slight increase of translation of the AUU mutant compared to that of the double mutant AUU T49C *in vivo* (unpublished data). However, in the future, more quantitative approaches such as ribosome profiling or proteomics will allow the further characterization of this exciting hypothesis.

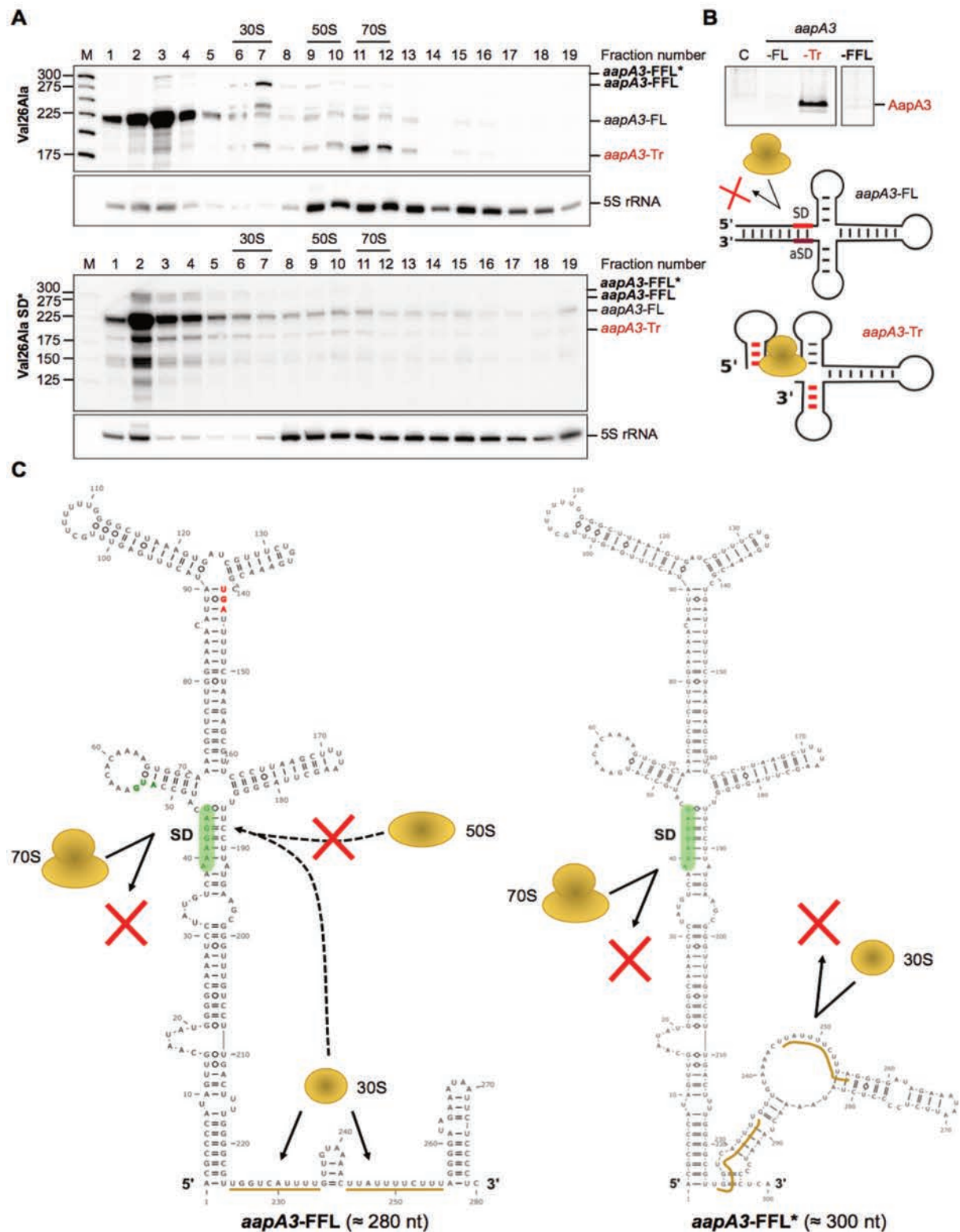
### 5.2.5. A 3' end standby site in a AapA3 mRNA species?

In the Chapter 4 section 4.3, we described how Northern Blot analysis revealed the existence of a AapA3 transcript having a 3' extension of approximately 55 nt compared to the full-length species, the so-called *aapA3* full-full length (*aapA3*-FFL,  $\approx$  280 nt). Interestingly, this longer transcript appears to be more abundant in a suppressor strain carrying a mutation in the AapA3 ORF, the Val26Ala (Valine at position 26 mutated to Alanine, previously identified by FASTBAC-Seq, see Chapter 4 section 4.1) strain, allowing us to study this mRNA species in more depth. In previous studies (see Chapter 4 section 4.2), we showed that the only translationally active AapA3 mRNA species is the *aap3*-Tr ( $\approx$  190 nt). However, with the discovery of the *aapA3*-FFL, we asked whether it could be translated. To test *aapA3*-FFL translation *in vivo*, we performed polysome fractionation in sucrose gradient coupled to Northern Blot analysis of the isolated fractions (as described in Chapter 4 section 4.2). As controls (apart of the 5S rRNA location), we used the *aapA3*-FL species ( $\approx$  225 nt), which is located in the first fractions of the gradient, indicative of being in 'free state' (non-translated), and the *aapA3*-Tr species ( $\approx$  190 nt), which is located in monosome and disomes fractions reflecting that it is being translated *in vivo*, as described in our recent paper (see Chapter 4 section 4.2). Strikingly, polysome fractionation analysis showed that the *aapA3*-FFL species mainly locates in the fractions corresponding to the 30S ribosomal subunit (Figure 5.11A, upper panel), and surprisingly, RNA secondary structure prediction revealed the presence of a putative ribosome standby site located on the 3' end of this mRNA species (within the 55 nt 3' end extension, underlined in yellow, left panel, Figure 5.11C).

Standby sites, originally discovered in the MS2 phage coat protein, are nucleotide unspecific single-stranded regions that allow translation of mRNAs on which the Shine-Dalgarno (SD) sequence is sequestered (de Smit and van Duin, 1990, 2003; Sterk et al., 2018). They have previously been identified at the 5' end of other type I TA toxin-encoding mRNAs such as the TisB (Darfeuille et al., 2007) and ZorO (Wen et al., 2017). However, the standby site of the *aapA3*-FFL mRNA would

represent the first identified case of a 3' end standby site. To test whether the *aapA3*-FFL presence in 30S fractions was due to the predicted 3' standby site, we introduced a mutation inactivating the SD sequence (the single point mutation guanine at position 43 from AapA3 TSS to adenine, G43A, suppressor mutation identified by FASTBAC-Seq, see [Chapter 4 section 4.1](#)) into the Val26Ala genetic background, generating the strain Val26Ala SD\* (Val26Ala/G43A/pIsoA3\*).

As shown in [Figure 5.11A](#) (lower panel), the *aapA3*-FFL mRNA species in this SD inactivated strain is no longer in the 30S fractions but on the 'free RNA' ones, together with the *aapA3*-FL mRNA. This result clearly demonstrated that the *aapA3*-FFL mRNA species was able to interact with 30S ribosomal subunits by using a standby site mechanism. However, the low presence of the *aapA3*-FFL species in the 70S fractions was sticking, as usually standby sites allow the translation of the mRNAs they are harbored in. To better understand this phenomenon, we performed *in vitro* translation assays using *E. coli* 30S extracts. The *aapA3*-FFL RNA, in line with the polysome fractionation results, was proved to be non-translatable ([Figure 5.11B](#)), obtaining a slightly higher translation than when the *aapA3*-FL RNA was used as substrate (-FL lane, [Figure 5.11B](#)), but being very close to background levels obtained in absence of exogenous RNA (control lane 'C', [Figure 5.11B](#)). Furthermore, a so far unobserved band of approximately 300 nt-long (*aapA3*-FFL\*) appeared in the 'free RNA' fractions. Interestingly, RNA secondary structure predictions of the *aapA3*-FFL\* mRNA located the standby site regions stably sequestered within stems (underlined in yellow, right panel, [Figure 5.11C](#)). This situation is in conformation with the polysome fractionation results (lower panel, [Figure 5.11A](#)) and explains the absence of the *aapA3*-FF\* mRNA in 30S fractions in contrast to the *aapA3*-FFL where the standby site regions are not sequestered. Put together, these results demonstrated that the *aapA3*-FFL mRNA is able to interact with 30S ribosomal subunits by using the single-stranded standby site region. However, due to the stable sequestration of the SD sequence within the long-distance interaction between the 5' and the 3' ends of the mRNA, the 50S ribosomal subunit cannot be assembled, explaining the low presence of the *aapA3*-FFL in the 70S fractions (upper panel, [Figure 5.11A](#)). Nevertheless, although the use of polysome fractionation was useful for us the future, the performance of *in vitro* toe-prints assays would be interesting and may reveal further mechanistic details.



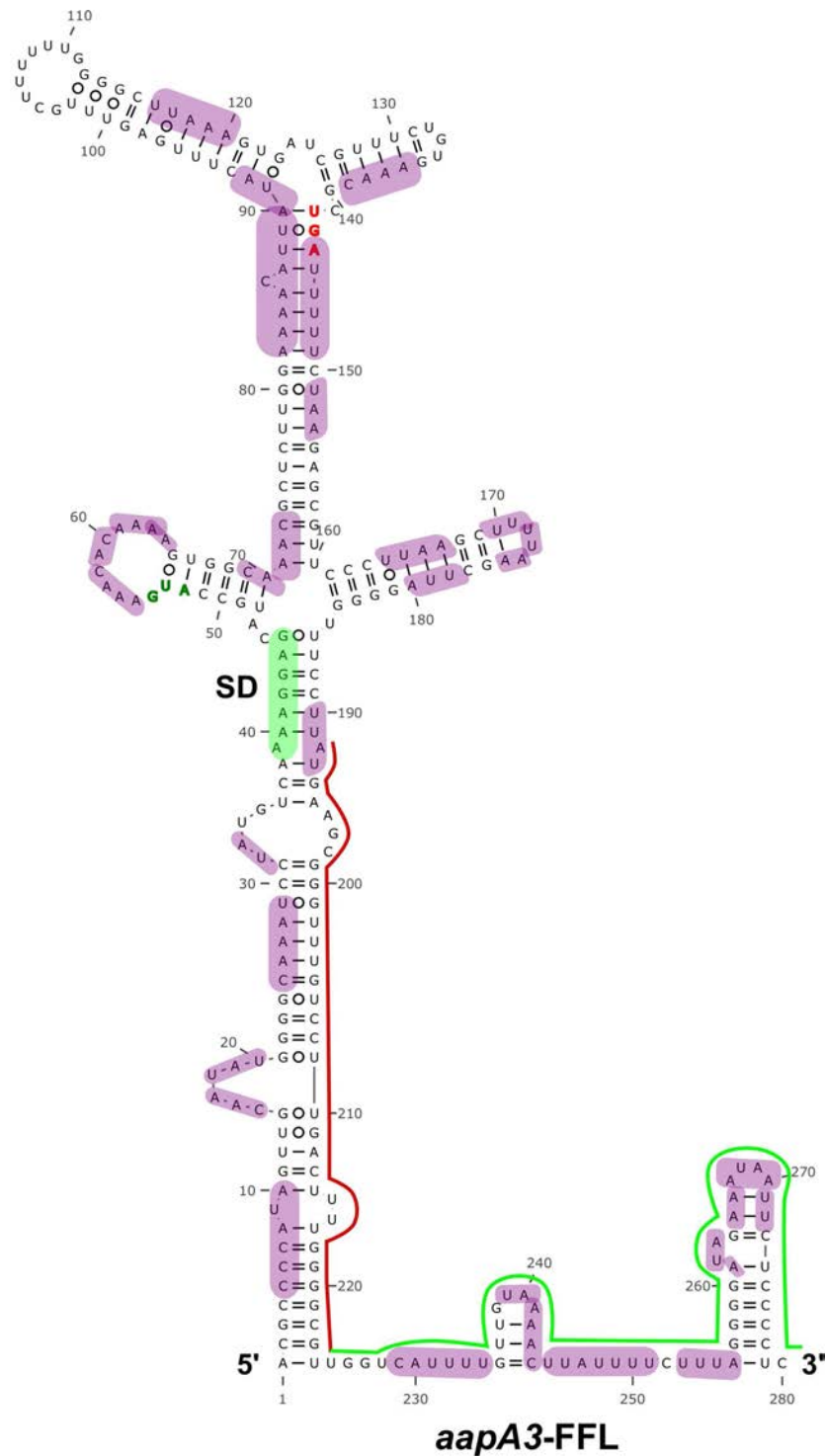
**Figure 5.11. The *aapA3*-FFL species harbors a 3' end standby site but cannot be translated.** (A) Cell lysate of the *H. pylori* 26695 carrying a toxicity suppressor mutation in AapA3 ORF (Val26Ala/pIsoA3\*) alone or together with an additional mutation in the Shine-Dalgarno sequence (SD\*, Val26Ala/G43A/pIsoA3\*) strains were subjected to ultracentrifugation through a sucrose gradient. A profile at OD<sub>254</sub> was recorded. RNA was extracted from each fraction and equal volumes of each extract were subjected to Northern Blot analysis. Fractions corresponding to the 30S and 50S subunits and 70S ribosomes are indicated. The different transcripts are annotated as: *aapA3*-FFL\* (full-full-length\*, ≈ 300 nt), *aapA3*-FFL (full-full-length, ≈ 280 nt), *aapA3*-FL (full-length, ≈ 225 nt), *aapA3*-Tr

(truncated,  $\approx 190$  nt) and 5S rRNA as loading control (5S rRNA). **(B)** *In vitro* translation assays were run with 0.5  $\mu\text{g}$  *aapA3* RNAs using [ $^{35}\text{S}$ ]-Met for labeling. Control lane (C) shows the translation background obtained in absence of exogenous RNA. Band corresponding to AapA3 peptide production is indicated (AapA3). **(C)** Secondary structure folding of the *aapA3*-FFL (full-length,  $\approx 225$  nt) species reveal a single-stranded region that serves as standby site for 30S ribosomal subunits. However, due to the tight Shine-Dalgarno (SD) sequestration, assembly of the 50S ribosomal subunit cannot take place. On the opposite, in the longer form, *aapA3*-FFL\* (full-full-length\*,  $\approx 300$  nt), the standby site is trapped within secondary structures, rendering it non-functional. RNAfold Web Server (Gruber et al., 2008) was used to predict the *aapA3*-FFL\* and the *aapA3*-FFL (full-length,  $\approx 225$  nt) structures and VARNA RNA (Darty et al., 2009) was used for drawing. SD sequence is shown in green, 30S and 50S ribosomal subunits and 70S ribosomes are represented in yellow, standby site regions are underlined in yellow.

---

Interestingly, a recent study has highlighted the role of AU-CA-repeats (Sterk et al., 2018), known to enhance translation of highly structured mRNAs (Komarova et al., 2005; Sharma et al., 2007; Yang et al., 2014), and be potentially recognized by the S1/S2 ribosomal proteins (as recently described for the *lpp* mRNA (Andreeva et al., 2018)), as enhancers of the standby site mechanism. As shown in Figure 5.12, several AU-CA rich motifs can be observed in the *aapA3* mRNA. Strikingly, the distribution of the AU-CA rich motifs in the AapA3 mRNA is not random. Despite being seemingly homogeneously decorated with AU-CA rich motifs, none of them is located in the 3' processed region that differentiates the primary transcript (*aapA3*-FL) from the active message (*aapA3*-Tr) (region indicated by a red line in Figure 5.12). This observation is intriguing and suggests that the AU-CA rich motifs in the *aapA3* mRNA may be functional and enhance the translation efficiency of this unusually highly structured mRNA. Furthermore, the AU-CA rich motifs re-appear in the 3' extended region that is unique to the *aapA3*-FFL species (indicated by a green line in Figure 5.12). This suggests that the *aapA3*-FFL species may be translationally active, however, further studies (e.g., point-mutation analysis, toeprints in presence or absence of S1 ribosomal protein, etc.) will be required to confirm and understand this mystery. The biological relevance for the generation of such a long (and energetically costly) transcript is enigmatic. However, one can think that it may provide a biological benefit under circumstances where a low toxin production is advantageous (e.g., stress conditions). Or even further, it may act as a sensor of intracellular ribosome concentration (i.e., high ribosome density would kinetically compensate the tight SD sequestration) and for instance, promote growth arrest or death of cells overproducing ribosomes (which may have deleterious effects, e.g., increased aberrant translation, or TA system lost due to increased cell division rates).





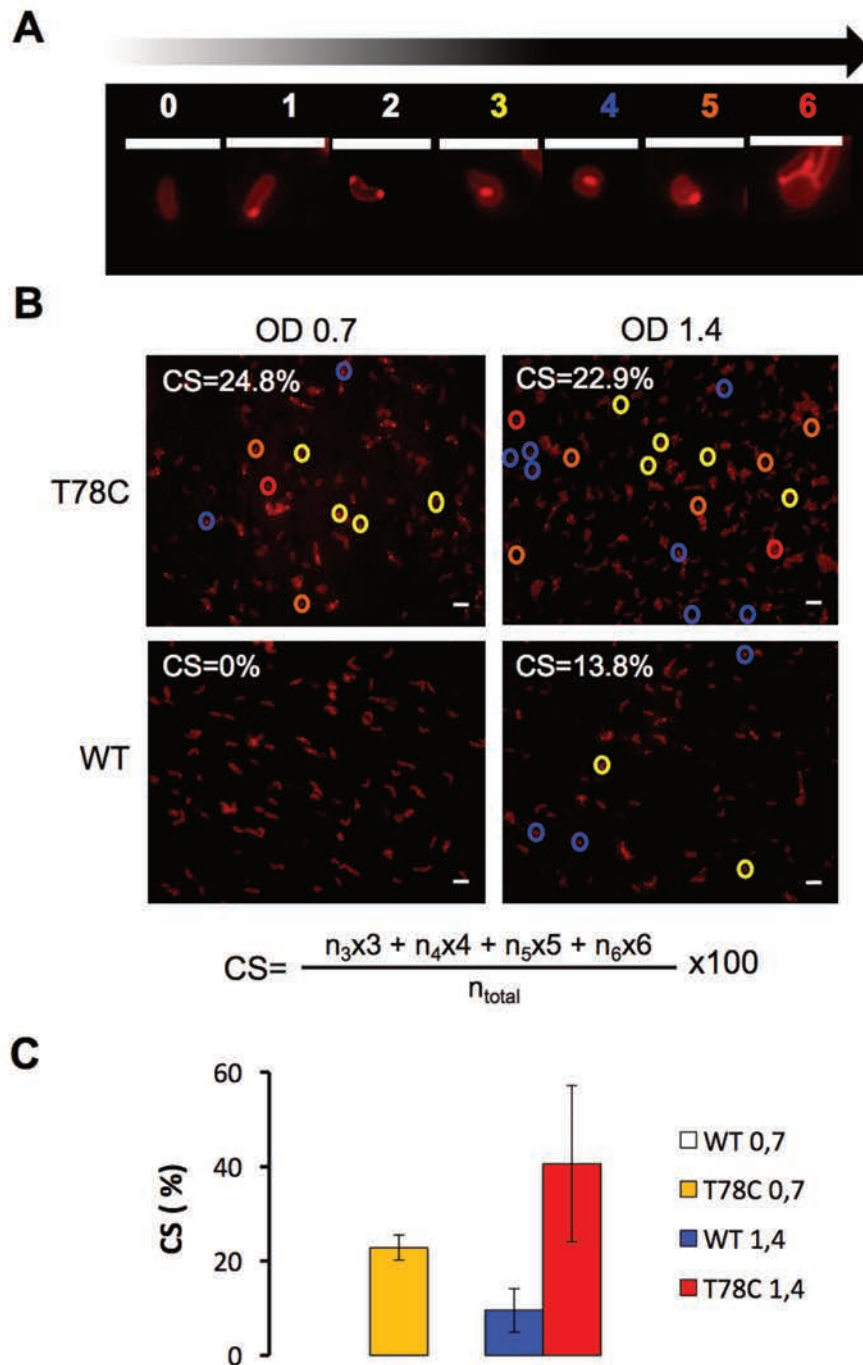
**Figure 5.12. Distribution of AU-CA-rich motifs in the *aapA3* mRNA.** RNAfold Web Server ([Gruber et al., 2008](#)) was used to predict the structures and VARNA RNA ([Darty et al., 2009](#)) for drawing. Translation start and stop codons are shown in green and red, respectively; Shine-Dalgarno (SD) sequence is highlighted in green; AU-CA-rich motifs are highlighted in purple; the 3' processed region in the mature *aapA3* mRNA species is indicated by a red line. The 3' extended region that is unique to the *aapA3*-FFL is indicated by a green line.

### 5.2.6. The inner membrane, the AapA3 target?

In Chapter 4 section 4.2 (paper III), we identified and characterized toxicity suppressor mutations able to inhibit the translation of the toxin-encoding AapA3 mRNA. Between them, there was the synonymous mutation T78C, which acts as an energetic trap stabilizing the active AapA3 mRNA species into a translationally inert conformation. Because this suppressor strain still codes for a wild-type AapA3 peptide, possible growth defects were studied. Interestingly, although no significant doubling time defects were observed (just a very slight growth retardation, unpublished data), microscopic observation of the cell morphology revealed something striking. *H. pylori* is known for being a ‘chameleonic’ bacteria, which undergoes several cell shape changes under different growth phases and environmental conditions, having a great impact in the colonization and virulence processes. During the exponential growth phase, *H. pylori* cells have helicoidal shape, are actively dividing, and are motile, being able to rapidly colonize the host. Towards the stationary phase, cells progressively become coccoïdal, stop dividing, and lose the motility, playing key roles in biofilm formation (see section 1.1.1.1 in Chapter 1, Bonis et al., 2010; Carron et al., 2006; García et al., 2014; Sycuro et al., 2012).

As shown in Figure 5.13, an unexpectedly high proportion of cells in the T78C/pisoA3\* strain present signs of transitioning to coccoïd shape, even at early exponential growth phases (OD ~0.7). Interestingly, we were able to identify all of the previously described growth phase-dependent transition stages (see Figure 1.1 in Chapter 1) (Cellini, 2014; Keshavarz et al., 1999), in one sole time-point in the T78C/pisoA3\* strain, and built a ‘coccoïd score (CS)’ ranging from 0 (helicoidal shape) to 6 (coccoïd shape) (Figure 5.13A). As shown in Figure 5.13B and C, meanwhile the CS in the wild-type strain is, as expected, growth phase-dependent ( $0 \pm 0$  at OD 0.7 and  $9.5 \pm 4.5$  at OD 1.4,  $n=3$  images per OD, with ~180 cells each), the CS observed in the T78C/pisoA3\* strain seems to remain stable over growth ( $22.8 \pm 2.6$  at OD 0.7, and  $40.6 \pm 22.6$ ,  $n=3$  images per OD, with ~180 cells each), and is in all cases significantly higher than the one found in the wild-type strain. Because despite stabilizing the AapA3 mRNA into a conformation on which the SD sequence is sequestered, mRNA folding is dynamic, it is likely that the observed phenotype in the T78C/pisoA3\* strain is produced by the leaking AapA3 expression. Thus, this result suggests that the AapA3 peptide can directly or indirectly promote cell transition to coccoïd shape.

To date, how *H. pylori* cells become coccoïd is still enigmatic. Electron microscopy and proteomic studies have pointed out that there are two types of coccoïd cells: 1) viable, having intact cytoplasm (e.g., unchanged protein and DNA content) and membrane structure, and 2) non-viable, having degenerated organelles and cell membrane and usually bigger than the viable ones (Bumann et al., 2004; Hua and Ho, 1996; Sisto et al., 2000; Young et al., 2001). In our case, and in line with the no-lytic phenotype observed during AapA3 peptide overexpression in *E. coli* (unpublished data), FM4-64 staining showed apparently intact membranes in the coccoïd cells, suggesting that they are viable (Figure 5.13).



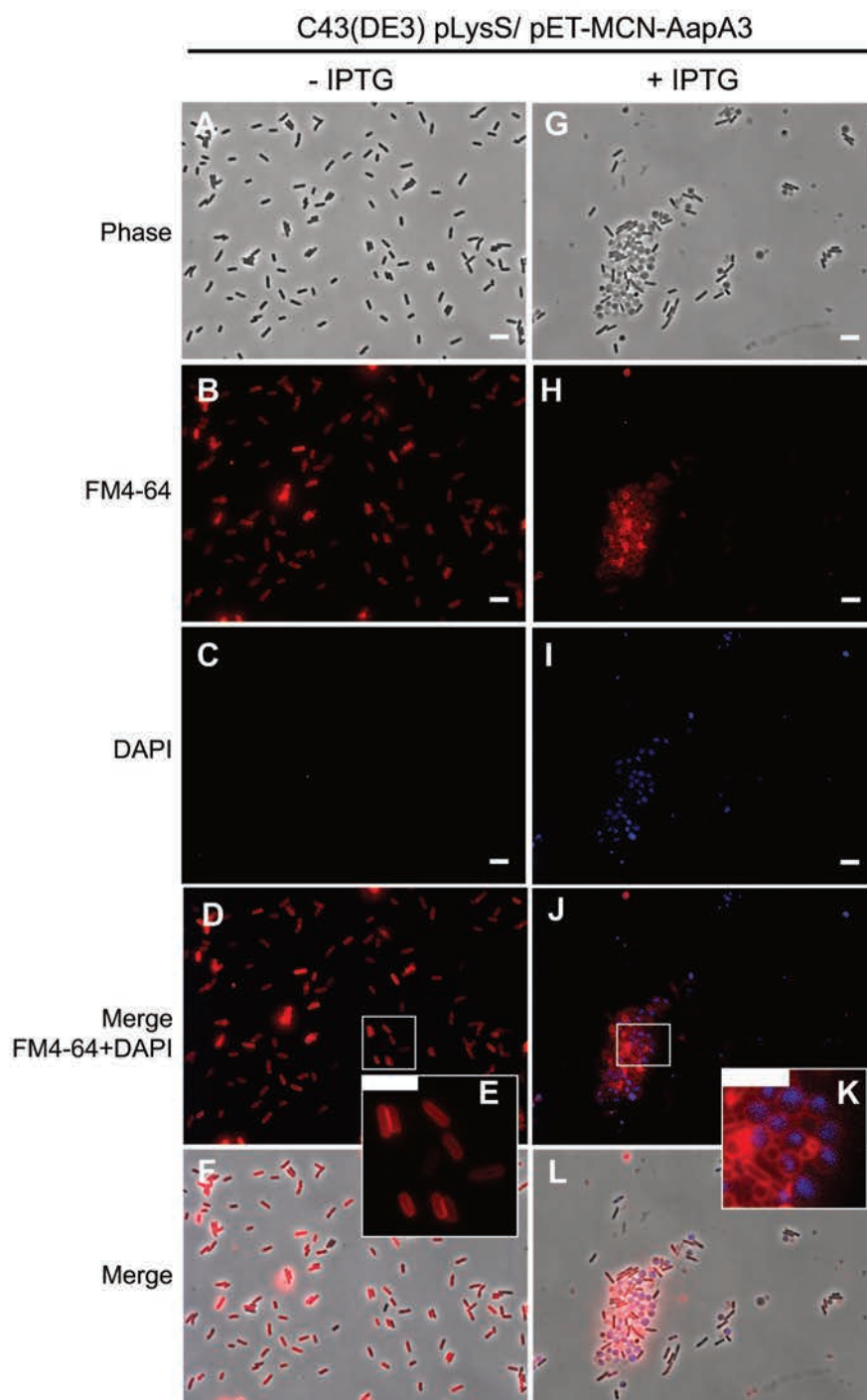
**Figure 5.13. AapA3 leaking chromosomal expression promotes the formation of coccoid cells at early growth phase stages in *H. pylori*.** (A) Phases of coccoid cell formation in the *H. pylori* 26695 strain. Coccoïd score (CS) ranging from 0 (helical shape) to 6 (coccoid shape). Images correspond to the T78C/pisoA3\* suppressor strain. FM4-64 was used for membrane staining. Scale bar, 5  $\mu$ m. (B) Determination of the membrane damage score (CS) by quantifying the number of cells with each shape score over the total number of cells using microscopy images. The images shown are representative of at least 30 images. The CS was calculated at two different growth phases (early exponential, OD 0.7 and late stationary phase OD 1.4) in the wild-type (WT) and T78C/pisoA3\* suppressor strain. FM4-64 was used for membrane staining. Scale bar, 5  $\mu$ m. (C) The CS at an early exponential (OD 0.7) and late stationary (OD 1.4) growth phase in the WT and T78C/pisoA3\* strains was calculated. Bars represent s.d; n=3 images per strain and growth phase with ~180 cells each on average.

Interestingly, previous studies on the AapA1 toxic protein of *H. pylori* have demonstrated that it targets the cell membrane (Korkut et al., in preparation). Traditionally, the highly specific bacteriophage T7 polymerase has been used for protein overexpression in *E. coli*. However, the overexpression of membrane proteins is often challenging. A study aiming at the optimization of membrane protein production from a phage T7 polymerase expression system in *E. coli* led to the selection of the mutant strain C43 (DE3). Notably, in the C43 (DE3) strain the toxic effects associated with overexpression are strongly reduced (Miroux and Walker, 1996). Another strategy that has been developed to reduce the overexpression-associated toxicity is the co-transformation of the plasmid containing the protein of interest with a plasmid containing the phage T7 lysozyme (Studier et al., 1990). The T7 lysozyme is a protein of 150 amino acids (~17 kDa) (Dunn and Studier, 1983) with N-acetylmuramyl-L-alanine amidase activity able to cut the peptidoglycan (PG) layer of the *E. coli* cell wall (Inouye et al., 1973). However, the reason behind the use of the T7 lysozyme for the expression of toxic proteins is its ability to bind to the T7 RNA polymerase and inhibit transcription (Moffatt and Studier, 1987).

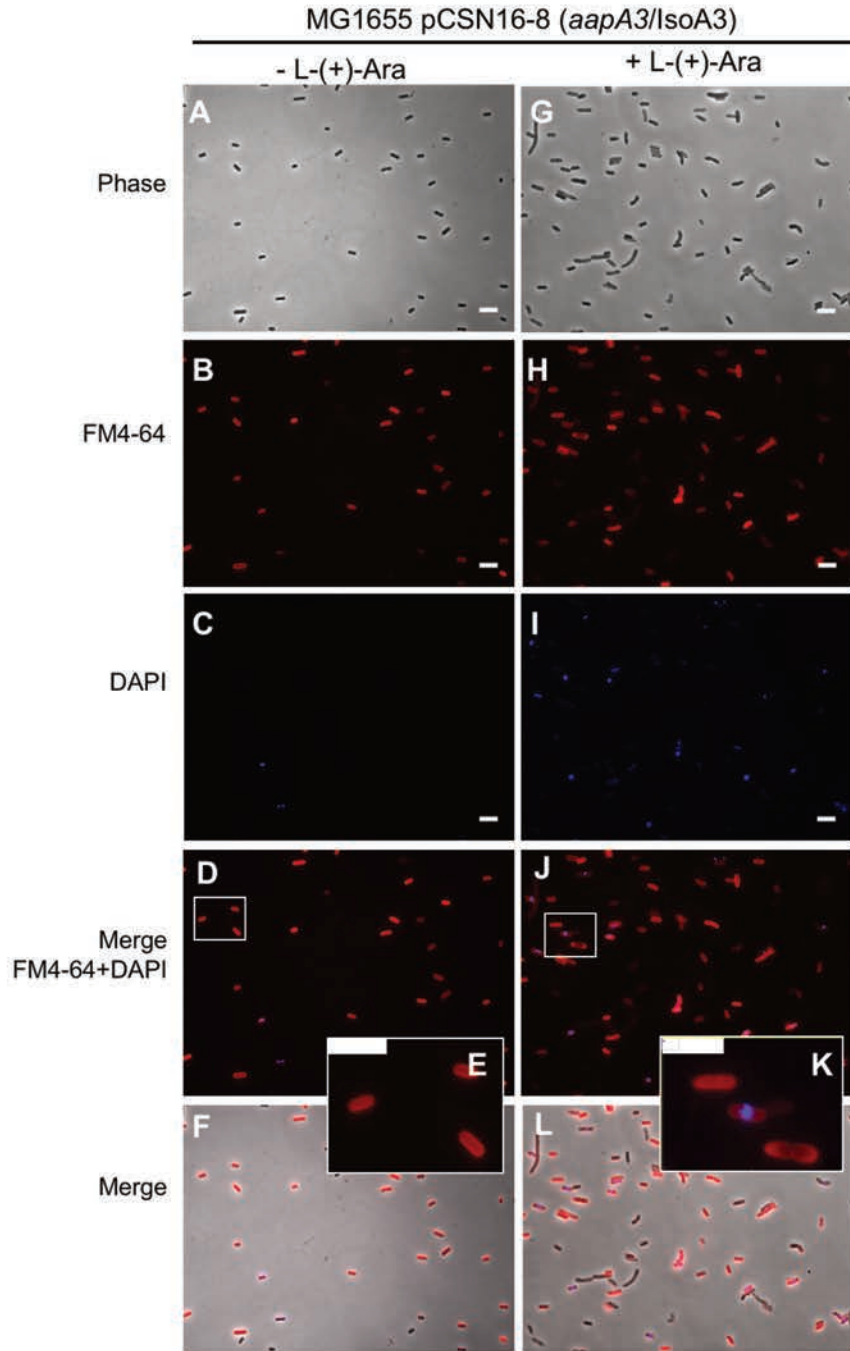
Preliminary attempts to clone the AapA3 ORF into an inducible vector repeatedly failed, leading to the apparitions of spontaneous mutations even under non-induced conditions. Considering this, we tested the use of the BL21 C43 (DE) strain and the pLysS plasmid containing the T7 lysozyme to express the AapA3 peptide from a phage T7 promoter (pET-MCN vector). In this case, the cloning of the AapA3 ORF was successful, and we were able to observe toxicity (bacteriostatic effect) upon induction (data not shown). However, once again, the microscopic analysis of the cells upon AapA3 protein induction revealed something striking. As shown in Figure 5.14, when AapA3 expression from the pET-MCN T7-derived expression vector in the C43 (DE3) *E. coli* strain co-transformed with the pLysS vector (containing the T7 lysozyme) was induced, strikingly round cells appeared visible (compare Figure 5.14E and K for detail). Furthermore, such round cells showed increased permeability to DAPI staining, suggesting an increased inner membrane permeability (compare Figure 5.14E and K for detail).

Interestingly, previous studies performed by C. Sharma soon after the identification of the *aapA*/IsoA family of type I TA systems in *H. pylori* (Sharma et al., 2010), demonstrated that it is possible to clone the AapA3 peptide into an overexpression vector (the arabinose-inducible pCSN16-8), and in a standard *E. coli* strain (e.g., MG1655), if it is in the context of the whole locus (C. Sharma, unpublished). Now, with the results generated in this thesis, we can understand the absence of leaking-toxicity in this case, as the AapA3 mRNA folding (involving the 5' and 3' UTRs, absent in the ORF-cloning strategy) is a key regulator of the AapA3 expression. Furthermore, in this case, the IsoA3 sRNA was also cloned, thus, it may also contribute to the lack of leaking AapA3 expression, and its presence does not represent a problem to study AapA3 toxicity as its low production from its endogenous promoter cannot compete with the high AapA3 expression levels obtained upon induction. Surprisingly, microscopy analysis of the MG1655/pCSN16-8 cells upon AapA3 induction showed no round phenotype (Figure 5.15, compare E and K for detail). Thus, these results demonstrated that the round phenotype observed in the BL21 C43 (DE3)/pLysS cells (Figure 5.14) was not a direct effect of AapA3 expression.





**Figure 5.14. Using T7 phage lysozyme as a reporter of inner-membrane damage in *E. coli*.** *E. coli* BL21 C43 (DE3) pLysS cells carrying the pET-MCN-AapA3 IPTG-inducible vector under non-induced conditions (**A-F**) or upon 1 h induction with 0.25 mM IPTG (**G-L**) were observed. FM4-64 was used for membrane staining. DAPI was used for inner membrane permeability assessment. Scale bar, 5  $\mu$ m.



**Figure 5.15. In absence of LysS no coccoïd-like cells are formed upon toxin induction in *E. coli*.** *E. coli* MG1655 cells carrying the arabinose-inducible AapA3 overexpression vector pCSN16-8 grown under non-induced conditions (**A-F**) or 1h upon induction with of 13.3 mM L-(+)-arabinose (**G-K**) were observed. FM4-64 was used for membrane staining. DAPI was used for inner membrane permeability assessment. Scale bar, 5  $\mu$ m.

Comparing the context of AapA3 expression in both strains, we spotted the pLysS (encoding the phage T7 lysozyme) as the responsible of the round phenotype. Remarkably, the presence of the phage T7 lysozyme turned out to be an excellent reporter of the AapA3 activity. Indeed, due to its high molecular weight (~17 kDa), the phage T7 lysozyme alone is incapable of crossing the *E. coli* inner membrane (lysis requires the action of several phage proteins (Bartel et al., 1996; Rajaure et al., 2015)). In our experiments, the T7 lysozyme could access and degrade the PG layer, leading to a round phenotype, only when the AapA3 peptide was produced. Therefore, these results suggest that the AapA3 peptide may interact with the inner membrane (IM) in a way that allows the T7 lysozyme to cross (e.g., behaving in a holing-like way, forming pores, as previously described allowing endolysin action (Wang et al., 2000; Young, 2014), or as described for the type I toxins TisB (Berghoff et al., 2017; Gurnev et al., 2012; Steinbrecher et al., 2012) and Hok (Wilmaerts et al., 2018) in *E. coli*).

Going a step further, in *H. pylori*, a similar mechanism involving *H. pylori* lysozyme-like autolytic enzymes (such as *lys* (Marsich et al., 2002)), may be responsible for the transition to coccoïd shape observed in the T78C/pIsoA3\* strain, or even in general matters. Importantly, we do not observe cell lysis or membrane fusion. This phenotype is reminiscent to the one found in the spanin (phage-encoded proteins that connect the IM with the OM during phage lysis) mutants (Rajaure et al., 2015), indicating that the outer membrane (OM) remains intact (and can be observed by FM4-64 staining, see Figures 5.13 and 5.14). Overall, the phage T7 lysozyme turned out to be an unexpectedly efficient *in vivo* reporter system of the AapA3 expression and activity in the inner membrane, which could likely be expanded to the study of other toxic membrane proteins. Interestingly, these results strongly suggest that the AapA3 peptide may act by forming pores in the IM of the bacterium and promote the formation of coccoïd cells. As described in the section 1.1.1.1 of Chapter 1, *H. pylori* coccoïd cells are often associated with biofilm formation and persistent infections. Pore formation is often linked to the entry to the so-called persister state (see section 1.4.2. in Chapter 1). Thus, these results suggest that coccoïd cells may represent a variant of the persister cells previously described in other bacteria, suggesting a role of AapA3 in *H. pylori* virulence.

However, further characterization will be required to confirm AapA3 peptide location at the IM (as previously described for the AapA1 toxin, Korkut et al., in preparation) and pore formation. Additionally, further studies analyzing the role of pore formation in the transition to coccoïd cells (e.g., direct or indirect effect) will be necessary. Importantly, the confirmation of the reversibility of the coccoïd phenotype and the restoration of growth, will be a key factor to confirm or discard a possible role of coccoïd cells in persistence. Finally, the role of AapA3 in *H. pylori* biofilm formation and virulence will need to be assessed, e.g., under which conditions AapA3 peptide is produced at the sufficient levels as to induce cell shape transition? Does IsoA3 transcriptional regulation play a role? If not, may stochastic expression of AapA3 increase population heterogeneity, as a mechanism of ensuring the presence of some coccoïd cells (which may be advantageous to respond to stress conditions, e.g., during infection)?

## CHAPTER 6. Material and Methods

### 6.1. Material

#### 6.1.1. Equipment

Equipment and instrument	Manufacturer
Anaerobic jar system	Oxoid
Analytical balances,	Sartorius
Anoxomat® Mark II	Advanced Instruments
Benchtop centrifuge	Eppendorf
Cell culture hood,	Thermo Scientific
Cell density meter	Amersham Biosciences
Gel dryer	Biorad
Heating block	GRANT
Horizontal electrophoresis systems	Emoi-Tec
Hybridization oven	TECHNE
Imaging plate cassettes	Biorad
Incubator for bacterial plates	Thermo Scientific
Incubator shaker (Innova® 44)	New Brunswick Scientific
K-screen, gel documentation system	Biorad
Microvolume UV-Vis spectrophotometer	Denovix
Orbital Shaker (Thermoshake)	Gerhardt
Pharos FX phosphorimager	Biorad
Precellys homogenizer	Bertin
Screen-K eraser	Biorad
SW41 Ti Swinging-Bucket rotor	Beckman Coulter
Swing-out rotor, refrigerated centrifuge	Heraeus
Thermocycler	Sensoquest
Trans-Blot® Cell	Biorad
UA-6 detector	ISCO
Ultracentrifuge	Beckman Coulter
UV-Crosslinker,	UVP
Vertical electrophoresis systems	APELEX
Vertical sequencing gel system CBS SG-400-20	C.B.S. Scientific
Water-bath (polystat®)	Bioblock Scientific

#### 6.1.2. Labware

Labware	Manufacturer
Autoradiography films	GE Healthcare
Cell culture flasks 25 cm <sup>3</sup>	Thermo Scientific
Cell culture flasks 75 cm <sup>3</sup>	Thermo Scientific
Cover slips	KNITTEL GALSS
Erlenmeyer glass flasks 250 ml	DURAN, SIMAX
Falcon tubes (plastic) 25 ml, 50 ml	SARSTEDT
Glass beads (2.85-3.35 mm) for plating of CFU/ml	SARSTEDT
Glass bottles	VWR
Inoculating loops (plastic) 1, 10 µl	VWR
Membrane for nucleic acid transfer, Hybond™-N	GE Healthcare
MicroSpin™ G-25 columns	GE Healthcare
Objective slides	KNITTEL GALSS
Petri dishes	SARSTEDT
Phase Lock Gel (PLG)-tubes, 2 ml	5PRIME
Pipet tips	SARSTEDT



Pipets 10 µl, 20 µl, 200 µl, 1000 µl	GILSON
Q-tips (CLASSIQSwabs™)	COPAN
Safe-lock tubes 0.2 ml, 1.5 ml, 2 ml	SARSTEDT
Serological pipets (plastic) 5 ml, 10 ml, 25 ml, 50 ml	SARSTEDT
Spectrophotometer cuvettes	SARSTEDT
Sterile filters (0.20 µm pore size)	SARSTEDT
Syringes (plastic) 10 ml, 50 ml	Terumo

### 6.1.3. Reagents and chemicals

Product	Manufacturer
1X alkaline hydrolysis buffer	Ambion
10X RNA structure buffer	Ambion
2-mercaptoethanol	Sigma
20X SSC (saline sodium citrate buffer)	Sigma
Acetic acid (100 %)	VWR
Acetone	VWR
Acrylamide	EUROMEDEX
Agarose	Thermo Scientific
Ammonium acetate	Sigma
Ampicillin sodium salt	Merk
Bacteriological agar	Thermo Scientific, Oxoid
Chloramphenicol	Sigma
Chloroform:isoamylalcohol (24:1)	Fluka
Dimethyl sulfoxide (DMSO)	EUROMEDEX
Dithiothreitol (DTT)	Sigma
ddNTPs/dNTPs	Sigma
DPBS (1X)	GIBCO
Erythromycin	Sigma
Ethanol	VWR
Ethanol absolute	VWR
Ethidium Bromide	Sigma
Ethylenediaminetetraacetic acid disodium salt dehydrate (EDTA)	Sigma
Formaldehyde	VWR
Formamide (99.5 %)	Sigma
Glycerol (99 %)	Sigma
Hydrochloric acid (HCl, 32 %)	Sigma
Isopropanol	Sigma
Kanamycin sulfate	Sigma
Lead (II)-acetate	Fluka
Magnesium chloride	Fluka
Methanol	VWR
Methylene Blue	Sigma
Molecular Biology Water	Eurobio
PEG4000 (50 %)	Thermo Scientific
Phenol	Sigma
Potassium chloride	Sigma
Rifampicin	Sigma
Sephadex G25	GE Healthcare
Sodium acetate	Sigma
Sodium chloride	Sigma
Sodium dodecyl sulphate (SDS)	Bisolve-Prolabo
Streptomycin	Sigma
Sucrose	Sigma
TRIS-Base	EUROMEDEX
TRIS-Glycine Buffer 10X	EUROMEDEX
Triton-X100	Sigma

Tween <sup>20</sup>	Amresco
Urea	Life-technologies
Yeast tRNA	Ambion
$\alpha$ - <sup>32</sup> P-UTP ( <sup>32</sup> P; 250 $\mu$ Ci (9.25 MBq))	Perkin-Elmer
$\beta$ - <sup>35</sup> S-Met ( <sup>35</sup> S; 1 mCi (37 MBq))	Perkin-Elmer
$\gamma$ - <sup>32</sup> P-ATP ( <sup>32</sup> P; 250 $\mu$ Ci (9.25 MBq))	Perkin-Elmer

#### 6.1.4. Enzymes and size markers

Product	Manufacturer
Calf Intestinal Phosphatase (CIP, 10 u / $\mu$ l)	New England Biolabs
Deoxyribonuclease I (DNase I, 1 u / $\mu$ l)	Thermo Scientific
<i>DpnI</i> (20 u / $\mu$ l)	Promega
DreamTaq DNA polymerase (5 u / $\mu$ l)	Thermo Scientific
Gene Ruler 1 kb plus DNA ladder	Thermo Scientific
Low-Range Rainbow molecular weight marker	GE Healthcare
Pfu Ultra High Fidelity DNA polymerase	Agilent
Phusion High-Fidelity DNA polymerase (2 u / $\mu$ l)	Thermo Scientific
Restriction enzymes	Promega/ Thermo Scientific
Ribonuclease A (RNase A)	Clinisciences
Ribonuclease H1 (RNase H1, 2 u / $\mu$ l)	Ambion
Ribonuclease III (RNase III, 2 u / $\mu$ l)	New England Biolabs
Ribonuclease T1 (RNase T1 0.01 u / $\mu$ l)	Ambion
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Thermo Scientific
RNasin® Ribonuclease Inhibitor	Promega
Shrimp Alkaline Phosphatase (SAP, 1 u / $\mu$ l)	Thermo Scientific
SuperScriptII Reverse Transcriptase	Invitrogen
SuperScriptIII Reverse Transcriptase	Invitrogen
T4 DNA Ligase	Promega
T4 Polynucleotide Kinase (PNK)	New England Biolabs
T4 RNA Ligase	Thermo Scientific

#### 6.1.5. Commercially available systems

System	Manufacturer
2X Gel Loading Buffer II (RNA)	Ambion
6X Sample Loading Buffer (DNA)	Fermentas
Cell Lysing Kit (2 ml tubes)	Precellys
<i>E. coli</i> S30 Extract System for Linear Templates Kit	Promega
GlycoBlue™	Ambion
High Purity Plasmid Miniprep Kit	Neo Biothech
HiTrap™ Heparin HP affinity columns	GE Healthcare
MAXIscript T7 <i>in vitro</i> transcription	Ambion
MEGAscript T7 <i>in vitro</i> transcription	Ambion
PCR and DNA fragment purification Kit	Neo Biothech
pGEM®-T vector system	Promega
Quick Bacteria Genomic DNA extraction Kit	Neo Biothech
Spin-X®-centrifuge tubes filter	Sigma
Stains-All	Sigma-Aldrich
Tricine Sample Buffer	Biorad

#### 6.1.6. Strains, oligonucleotides and plasmids

*Helicobacter pylori* strains used in this study are compiled in the [Tables 8.1.1](#) and [8.2.3](#) in the Appendix. Plasmids are listed in the [Tables 8.1.2](#) and [8.2.5](#) in the Appendix. Oligonucleotides are listed in the [Tables 8.1.4](#) and [8.2.6](#) in the Appendix.

### 6.1.7. Media and supplements

#### *H. pylori*-specific media:

##### **BHI medium** (Oxoid)

36 g                      Brain Heart Infusion: 0.5% (w/v) beef heart, 1.25% (w/v) calf brains, 0.25% (w/v) disodium hydrogen phosphate, 0.2% (w/v) D (+)-glucose, 1% (w/v) peptone  
ad 900 ml              H<sub>2</sub>O  
add after autoclaving:  
10% (w/v)              heat-inactivated fetal bovine serum

##### **CAB medium** (Oxoid)

36 g                      Columbia Blood Agar Base (CAB): 0.5% (w/v) enzymatic digest of casein, 0.8% (w/v) enzymatic digest of animal tissue, 1% (w/v) yeast enriched peptone, 0.1% (w/v) corn starch, 85.6 mM sodium chloride, 1.4% (w/v) agar  
ad 900 ml              H<sub>2</sub>O  
add after autoclaving:  
7% (w/v)              Laked Horse Blood (Thermo Scientific)  
2 vials                  *H. pylori* selective supplement (Dent, Oxoid, Basingstoke, UK)

#### *E. coli*-specific media:

##### **Lennox Broth (LB) liquid medium** (Oxoid)

1% (w/v) peptone 140  
0.5% (w/v) yeast extract  
85.6 mM sodium chloride

##### **LB agar plate**

LB medium (see above)  
1.2 % agar

##### **SOC medium** (Oxoid)

2% (w/v) tryptone  
0.5% yeast extract  
85.6 mM NaCl  
2.5 mM KCl  
10 mM MgCl<sub>2</sub>  
20 mM glucose

##### **Terrific broth** (Sigma)

1.2% (w/v) tryptone  
2.4% (w/v) yeast extract  
0.94% (w/v) K<sub>2</sub>HPO<sub>4</sub>  
0.22% (w/v) KH<sub>2</sub>PO<sub>4</sub>

### 6.1.8. Buffers

#### **2X Gel Loading buffer II (RNA):**

95%                      formamide  
18 mM                      EDTA, pH 8.0  
0.025% (w/v)              each of SDS, Xylene Cyanol, and Bromophenol Blue

**Caco-buffer (pH 7.3 at 20°C):**

20 mM	sodium cacodylate
20 mM	NaCl
140 mM	KCl
1 mM	MgCl <sub>2</sub>

**Buffer A:**

10 mM	Tris-HCl, pH 7.5
60 mM	potassium chloride
10 mM	magnesium chloride

**Grad-buffer:**

10 mM	Tris-HCl, pH 7.5
50 mM	ammonium chloride
10 mM	magnesium chloride
1 mM	DTT

**Lysis buffer:**

20 mM	sodium acetate, pH 4.2
0.5% (w/v)	SDS
1 mM	EDTA

**Modified Church buffer:**

1 mM	EDTA
0.5 mM	sodium phosphate, pH 7.2
7 % (w/v)	SDS

**Phosphate buffered Saline (10X):**

80 g	sodium chloride
2 g	potassium chloride
17.7 g	disodiumhydrogen phosphate
800 ml	H <sub>2</sub> O
adjust to pH 7.4	
add H <sub>2</sub> O to a final volume of 1 l	

**RNA elution buffer:**

0.1 M	sodium acetate, pH 5.2
0.1% (w/v)	SDS
1 mM	EDTA

**RNA elution buffer-NH<sub>4</sub>Ac:**

0.3 M	ammonium acetate
0.1% (w/v)	SDS
1 mM	EDTA

**RNase J-exo reaction buffer (5X)**

100 mM	Tris-HCl pH 7.0
40 mM	MgCl <sub>2</sub>
500 mM	NH <sub>4</sub> Cl
0.5 mM	DTT

**RNA structure buffer (1X):**

10 mM	Tris-HCl, pH 7.0
-------	------------------

10 mM	magnesium chloride
100 mM	potassium chloride

**RNase J-endo reaction buffer (5X)**

100 mM	Tris-HCl pH 8.0
40 mM	MgCl <sub>2</sub>
500 mM	NaCl

**Sequencing buffer (1X):**

20 mM	sodium citrate, pH 5.0
7 M	Urea
1 mM	EDTA

**SSC (saline-sodium citrate) buffer (20X stock solution):**

173.5 g	sodium chloride
88.2 g	sodium citrate
800 ml	H <sub>2</sub> O

adjust to pH 7.0 (using HCl)  
add H<sub>2</sub>O to a final volume of 1 l

**Stains-All:**

30 ml	Stains-All stock
90 ml	formamide

add H<sub>2</sub>O to a final volume of 200 ml

**Stains-All stock:**

0.03 g dissolved in 30 ml formamide

**Stop Solution:**

95 % (v/v)	ethanol
5 % (v/v)	phenol

**TBE buffer (10X stock):**

108 g	Tris base
55 g	boric acid
40 ml	0.5 M EDTA, pH 8.0

add H<sub>2</sub>O to a final volume of 1 l

**TE buffer (1X):**

100 mM	Tris-HCl, pH 8.0
10 mM	EDTA, pH 8.0

**Tris-HCl/SDS buffer (3M Tris-Cl, 0.3% SDS, pH 8.45):**

18.2 g	Tris-Base (TRIZMA base)
--------	-------------------------

ad 300 ml H<sub>2</sub>O  
adjust to pH 8.45  
add 7.5 ml 20 % SDS  
add H<sub>2</sub>O to a final volume of 500 ml



**Washing-buffer (1X):**

25 mM                      Tris-HCl, pH 8.0  
300 mM                     NaCl

**1X cathode buffer (0.1 M Tris, 0.1M tricine, 0.1% SDS, pH ≈8.3, do not adjust):**

12.11 g                     Tris-Base (TRIZMA base)  
17.92                        Tricine  
5 ml                         20 % SDS  
add H<sub>2</sub>O to a final volume of 1 l

**5X anode buffer (0.2 M Tris, pH 8.9):**

121.1 g                     Tris-Base (TRIZMA base)  
add H<sub>2</sub>O to a final volume of 500 ml  
adjust to pH 8.9 with HCl 10 N

**6.1.9. Sterilization**

Media solutions were sterilized before use by autoclaving at 120°C and 1 bar for 20 min. Thermolabile solutions were filtered-sterilized.

## 6.2. General Methods

This section describes general methods that were routinely used throughout the study. For a detailed protocol of the FASTBAC-Seq method see [Chapter 4 section 4.1 \(paper II, \(Masachis et al., 2018\)\)](#).

### 6.2.1. *E. coli* growth conditions and strain generation

*E. coli* strains were grown at 37°C and normal aeration conditions. Cultures were inoculated either from overnight grown single-colonies or freshly grown single-colony 5 ml LB cultures diluted 1/100 and incubated at 37°C and 200 rpm.

For inducible P<sub>Lac</sub> or P<sub>BAD</sub> promoters, cultures were supplemented with 0.25 mM or 13.3 mM IPTG or L-(+)-arabinose, respectively.

When needed, antibiotics were used as follows:

Kanamycin: 50 µg/ml

Ampicillin: 100 µg/ml

Chloramphenicol: 25 or 30 µg/ml for low- or high-copy plasmids, respectively

#### 6.2.1.1. Transformation of *E. coli* chemically competent cells

50 µl of homemade chemically competent TOP10 cells (Invitrogen) were mixed with 1 µl ligation product or 0.5-2 µl plasmid (concentration 10-150 ng/µl). After 5 min pre-incubation on ice, cells were subjected to a heat-shock for 45 seconds at 42°C. Next, cells were chilled on ice for 1 min and resuspended in 500 µl SOC medium. Recovery was carried out for 30 or 60 min for cells carrying ampicillin or chloramphenicol resistance markers, respectively. Cells were pelleted by centrifugation at 3000 rpm for 5 min and resuspended in 50 µl SOC medium before being plate on LB-agar plates supplemented with the appropriate antibiotics.

### 6.2.2. *H. pylori* growth conditions and strain generation

The *H. pylori* strains used in this study were 26695 ([Tomb et al., 1997](#)), B128 ([Farnbacher et al., 2010](#); [McClain et al., 2009](#)) and X47-2AL ([Ermak et al., 1998](#)). Strains were grown on Columbia agar plates supplemented with 7% horse blood and Dent selective supplement for 24 h to 48 h depending on the strain. Liquid cultures were performed in Brain-Heart Infusion (BHI) medium supplemented with 10% fetal bovine serum (FBS) and Dent. *H. pylori* plates and liquid cultures were incubated at 37°C under microaerobic conditions (10% CO<sub>2</sub>, 6% O<sub>2</sub>; 84% N<sub>2</sub>) using an Anoxomat atmosphere generator. For liquid cultures, bacteria harvested from plates were inoculated at an optical density at 600 nm of 0.02 (OD<sub>600</sub> = 0.02) into tubes containing 10 mL BHI medium supplemented with 10% FBS and Dent supplement and incubated at 37°C and 200 rpm shaking. After 8 h, pre-cultures were diluted to the desired OD<sub>600</sub> into 25 mL flasks and incubated at 37°C and 120 rpm shaking.

For *H. pylori* mutant selection and culture, antibiotics were used at the following final concentrations:

Kanamycin: 20 µg/ml

Chloramphenicol: 8 µg/ml

Streptomycin: 10 µg/ml

Erythromycin: 10 µg/ml

### 6.2.2.1. *H. pylori* chromosomal manipulation techniques

All mutant *H. pylori* strains listed in the [Tables 8.1.1](#) and [8.2.3](#) in the Appendix were generated by chromosomal homologous recombination of PCR-generated constructs, introduced by natural transformation, as previously described ([Bury-Moné et al., 2001](#)). In all cases, constructs contained  $\approx$  400 nt of the up- and downstream regions of the target gene of the chromosome, flanking the DNA fragment to be introduced (*i.e.*, antibiotic resistance marker to generate deletions or a wild type copy of the target gene for complementation). DNA fragments of interest previously cloned in *E. coli* vectors to avoid *H. pylori* wild type genomic DNA (gDNA) contamination. Constructs were generated by PCR assembly of PCR products amplified from the plasmids shown in [Tables 8.1.2](#) and [8.2.5](#) in the Appendix with the oligonucleotides shown in [Tables 8.1.4](#) and [8.2.6](#) in the Appendix. Prior to transformation, *H. pylori* strains were grown from  $-80^{\circ}\text{C}$  stocks for two passages, then resuspended in 1 ml BHI and non-selective CAB plates were inoculated (by making a small circle) with the volume equivalent to 1 OD<sub>600</sub>. After 4 hours incubation at  $37^{\circ}\text{C}$  under microaerobic conditions, 1  $\mu\text{g}$  of PCR assembly product was added to the cells and plates were incubated for another 16 hours. Transformed cells were then passed into plates supplemented with the appropriate antibiotics for the selection of recombinant cells and incubated for 4-6 days until isolated colonies appeared. Genomic DNA from potential clones was purified using the Quick Bacteria Genomic DNA extraction Kit and subjected to PCR and Sanger sequencing for mutant validation.

### 6.2.2.2. Generation of an *H. pylori* 26695 streptomycin resistant strain

The use of traditional drug resistance selection markers for the generation of mutant strains in *H. pylori* is possible, however, it can rapidly become limiting (*e.g.*, if several genes need to be deleted, if resistance determinant affects cellular metabolism, or if alleles with small differences need to be compared). The use of a recipient strain containing a counterselectable marker at the locus of interest can overcome such limitations. In this study, the *rpsL<sub>CJ</sub>-erm* counterselectable marker conferring streptomycin sensitivity (*rpsL<sub>CJ</sub>*) and erythromycin resistance (*erm*) was used as described by Dailidiene and colleagues ([Dailidiene et al., 2006](#)) and ([Pernitzsch et al., 2014](#)). This allows the selection of unmarked transformants by the loss of this *rpsL<sub>CJ</sub>-erm* marker, avoiding the use of traditional resistance determinants in the donor DNA. However, the use of this counterselectable marker requires the prior generation of a streptomycin resistant strain. To that aim, a PCR fragment containing the *H. pylori* 26695 *rpsL* gene carrying a mutation in the Lysine at position 43 to Arginine (*rpsL*K43R), was used to transform a wild type *H. pylori* 26695 strain as described in the General Methods section in [subheading 8.2.2.1](#) generating the strain *rpsL*K43R ([Table 8.1.1](#) in Appendix). This mutation in the *rpsL* gene, coding for the protein S12, has been previously demonstrated to confer streptomycin resistance by prompting S12 mutated binding to the same 16S rRNA loops that are specifically targeted by streptomycin ([Gregory et al., 2001](#); [Torii et al., 2003](#)), thereby avoiding its action.

### 6.2.2.3. Deletion of the *H. pylori* 26695 *aapA3/IsoA3* locus using the *rpsL<sub>CJ</sub>-erm* cassette

The method here developed for the analysis of toxicity determinants on toxin-encoding genes belonging to type I TA systems (FASTBAC-Seq, [Chapter 4 section 4.1](#)) required the previous generation of a strain on which the TA locus of interest have been replaced with the counterselection cassette *rpsL<sub>CJ</sub>-erm* on a streptomycin resistant background (*rpsL* K43R), as described by ([Dailidiene et al., 2006](#)). To this aim, 1) a fragment upstream of the locus was amplified with the primer pair FA406/FA407 ([Table 8.1.4](#) in Appendix), representing a 415 nt homology region, 2) a second fragment downstream of the locus was amplified with the primer pair FA408/FA409 ([Table 8.1.4](#) in Appendix), representing a 418 nt homology region. Note, the internal primers (FA407 and FA408)

were used to introduce a 3'- and 5'- *rpsL<sub>CJ</sub>-erm* cassette homology tails, respectively, that will allow subsequent PCR assembly. The *rpsL<sub>CJ</sub>-erm* cassette was amplified from the pSP60-2 plasmid (Table 8.1.3 in the Appendix) using the primer pair FA110/FA111 (Table 8.1.4 in Appendix). The upstream and downstream fragments were then assembled with the *rpsL<sub>CJ</sub>-erm* cassette by PCR assembly using the external primers (FA406/FA409). This construct was used to transform the *H. pylori* 26695 strain as described in the General Methods section subheading 6.2.2.1. This process generated the strain that was posteriorly used as recipient strain in all our transformation experiments, the  $\Delta aapA3/IsoA3::rpsL_{CJ}\text{-erm}/rpsL_{K43R}$  strain (Table 8.1.1 in the Appendix).

#### 6.2.2.4. *AapA3/IsoA3* locus sub-cloning in *E. coli*

Because *H. pylori* has a highly active homologous recombination machinery, a cloning step of the *aapA3/IsoA3* locus in an *E. coli* vector was essential to avoid contamination with WT *H. pylori* gDNA of the PCR products used in the transformation assays. To this end, the *aapA3/IsoA3* locus was split into two fragments amplified with the Phusion High-Fidelity Hot Start DNA Polymerase and the primer pairs FA406/FA386 (fragment “Up”, of 638 nt containing 415 nt of homology region, the *aapA3* promoter and the first 10 amino acids residues of AapA3 peptide, see Figure 8.1.1A in the Appendix) and FA409/FA387 (fragment “Down”, of 680 nt containing IsoA3 promoter, the rest of *aapA3* mRNA and 418 nt of homology region, Figure 8.1.1A1A in the Appendix). Note that the FA386 and FA387 primers have 25 nucleotides overlap to allow PCR assembly. Each fragment was cloned in a separate pGEM®-T plasmid (Table 8.1.3 in the Appendix) and transformed into One Shot TOP10 chemically competent *E. coli* cells.

#### 6.2.2.5. Mutant generation by Site-Directed mutagenesis PCR

Plasmids generated in the General Methods subheading 8.2.2.4. were isolated. Custom designed overlapping oligonucleotides containing the desired mutations were used to perform Site-Directed mutagenesis PCR. To inactivate the IsoA3 –10 box two synonymous point mutations (at position +87 and +90 from the AapA3 TSS) were introduced using the primer pair FA283/FA284 (Table 8.1.4 in the Appendix, see Figure 4.2.1 for details). This strategy preserves the toxin coding sequence while completely abolishing the transcription of the antitoxin. To inactivate the toxin start codon, a single point mutation in the third codon position was introduced using the primer pair FA281/FA282 (Table 8.1.4 in the Appendix, see Figure 4.2.1 for details). Wild-type or mutated fragments were amplified from the previously-generated and isolated plasmids using 35 amplification cycles, the Phusion High-Fidelity Hot Start DNA Polymerase, and the same primer pairs as those used for insert amplification prior to cloning. PCR assembly with 35 amplification cycles, the Phusion High-Fidelity Hot Start DNA Polymerase and the external primers FA406/FA409 was performed to construct the *aapA3/IsoA3* locus variants (1294 nt amplicon) posteriorly used as DNA substrates for *H. pylori* transformation. For the *in vivo* ‘clean’ validation of all the mutants here studied, this same protocol was used adapting the DNA oligonucleotides used for Site-Directed mutagenesis PCR.

#### 6.2.2.6. Determining *H. pylori* transformation efficiency

*H. pylori* transformation was performed as described in General Methods subheading 6.2.2.1. The transformation assays aimed at the determination of the transformation efficiency as an indirect prove of the toxicity produced by the expression of a PCR construct. For this purpose, transformation patches (after 16 hours growth upon DNA addition) were recovered and resuspended in 1 ml BHI. Ten-fold serial dilutions adapted to each transformation case ( $10^{-7}$ ,  $10^{-6}$  and  $10^{-5}$  for non-selective media; and  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$  for selective media upon transformation with water or a toxic construct; and  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-3}$  for selective media upon transformation with non-toxic constructs) were performed.

Allelic replacement events were selected by the use of streptomycin-containing plates (selection of loss of the *rpsl<sub>Cj</sub>-erm* cassette, Str<sup>R</sup>). The number of Str<sup>R</sup> CFU/total CFU was calculated, plotted on base-10 logarithmic scale and statistically analyzed by unpaired t-test (GraphPad Prism software version 7).

#### 6.2.2.7. *H. pylori* transformation assay to identify toxicity suppressors by Illumina sequencing

Transformation assays aiming at the identification to toxicity suppressors by Illumina sequencing were performed in three biological replicates, using the wild-type (WT) or antitoxin promoter inactivated PRC-generated constructs (pIsoA3\*). Upon transformation, all bacteria were recovered and serially diluted. Colonies were isolated on streptomycin-containing plates by using optimized dilutions (9 plates/replicate of 10<sup>-1</sup> dilution for pIsoA3\* and 3 plates/replicate of 10<sup>-3</sup> dilution for WT). Three days after transformation, isolated colonies were pooled (approximately 60,000 colonies per transformation) and genomic DNA was extracted. Next, a 426 nt-long amplicon corresponding to the *aapA3/IsoA3* locus was amplified with the primer pair FA395/FA396 (introducing the DNA adapters for Illumina paired-end sequencing, [Table 8.1.4 in the Appendix](#)). Importantly, to avoid amplification from *rpsl<sub>Cj</sub>* revertant strains (mutated *rpsl<sub>Cj</sub>* gene), the FA395 and FA396 primers are nested to the ones used for locus deletion (FA407 and FA408), thus, binding to deleted regions that are re-introduced only upon recombination. For this PCR, the Phusion High-Fidelity Hot Start DNA polymerase and 35 amplification cycles were used. Finally, the samples were sent for sequencing at the Plateforme GeT-PlaGe-, Genotoul Centre INRA, Toulouse. Sequencing was done on an Illumina MiSeq instrument in paired-end mode 2 x 250 nt (overlapping reads).

### 6.2.3. Nucleic acid techniques

Molecular biology experiments were performed according to standard procedures and the supplier recommendations. High Purity Plasmid Miniprep Kit and Quick Bacteria Genomic DNA extraction Kit were used for plasmid preparations and *H. pylori* genomic DNA extractions, respectively. PCR were performed either with Dream Taq DNA polymerase, or with Phusion High-Fidelity Hot Start DNA polymerase when the product required high fidelity polymerase. Site directed mutagenesis PCR was performed with the PfuUltra High-Fidelity DNA Polymerase.

#### 6.2.3.1. Agarose gel electrophoresis

DNA fragments were separated in agarose gels. For preparation 0.8 to 2% (w/v) agarose was dissolved in 1X TBE buffer. At a gel solution temperature of ≈55°C, ethidium bromide was added to a final concentration of ≈50 µg/100 ml. Samples were mixed with five volumes of 6X Sample Loading Buffer, loaded and run in 1X TBE buffer at 120 V for 20-40 min (depending on the fragment size).

#### 6.2.3.2. Polyacrylamide gel electrophoresis

RNA fragments were separated in polyacrylamide (PAA) gels.

**Denaturing PAGE:** denaturing gels (native structure of RNA molecules is destroyed) were supplemented with urea to a final concentration of 7 M.

Prior to use, all equipment was carefully cleaned with 70% ethanol. Polymerization was initiated by adding 1/100 volume of ammonium persulfate (APS) and 1/1000 volume of N,N,N',N'-tetramethylethane-1,2-diamine (TEMED). RNA samples were mixed with one volume of 2X Gel Loading Buffer II, denatured for 5 min at 90°C and loaded. Gel runs were performed in 1 x TBE at



18W for 1 to 4 hours (according to the size of the RNA species to be analyzed). When needed (*i.e.*, sequencing gels), gels were dried for 45 min at 80°C, and revealed using a Pharos FX phosphorimager.

#### **6.2.3.3. Total RNA preparation with hot phenol**

Bacterial growth was stopped at the desired OD<sub>600</sub> by addition of 650 µl Stop Solution (pre-chilled at -20°C) to 5 mL of culture and rapidly placed on ice. Cells were then centrifuged for 10 min at 3,500 rpm and 4°C and the pellets were stored at -80°C. Cell pellets were resuspended in 600 µl Lysis Solution and added to 600 µl hot phenol pH 5.2. After incubation for 10 min at 65°C, the mixture was centrifuged for 10 min at 13,000 rpm and room temperature. The aqueous phase was transferred to a phase-locked gel tube with an equal volume of chloroform and centrifuged for 10 min at 13,000 rpm and room temperature. Total RNA was precipitated from the aqueous phase by adding 2.5 volumes of EtOH 100% and 1/10 volumes 3 M NaOAc pH 5.2. After centrifugation for 30 min at 13,000 rpm and 4°C, the supernatant was discarded and the pellet was washed with 75% EtOH. Finally, the supernatant was discarded and the RNA pellet air-dried and resuspended in H<sub>2</sub>O. For RNA half-life determinations, rifampicin (prepared at 34 mg/ml in methanol) was added to the culture at a final concentration of 80 µg/ml and cells were harvested at the desired time points. A culture where rifampicin was replaced by the same volume of methanol served as a non-treated control.

#### **6.2.3.4. DNase I digestion**

RNA samples resuspended in H<sub>2</sub>O denatured at 65°C for 5 min and subsequently cooled down on ice for other 5 min. Then, 1 µl of DNase I per µg of RNA was added and samples were incubated for 30 min at 37°C. After DNase I treatment RNA was isolated by phenol:chloroform:isoamylalcohol (P:C:I) extraction. RNA was mixed with one volume of P:C:I (25:24:1 v/v) by vortexing for 10 s. Samples were centrifuged for 10 min at 13,000 rpm and 4°C. The aqueous phase (upper) was mixed with 2.5 volumes of 30:1 (v/v) EtOH:3M NaOAc pH 5.2 and RNA was precipitated overnight at -20°C.

#### **6.2.3.5. Radioactive labeling of DNA oligonucleotides for RNA detection**

For DNA oligonucleotide labeling 10 pmol of the oligonucleotide was incubated in a 10 µl reaction volume with 25 µCi of <sup>32</sup>P-γ-ATP in the presence of 1 u T4 polynucleotide kinase (PNK) for 30 min at 37°C. MicroSpin<sup>TM</sup> G-25 columns were used to get rid of unincorporated nucleotides.

#### **6.2.3.6. Northern Blot**

For Northern blot analysis, 1 to 10 µg RNA were separated on an 8% polyacrylamide (PAA), 7M urea, 1x Tris-Borate-EDTA (TBE) gel. RNA was transferred to a Hybond<sup>TM</sup>-N nylon membrane by electroblotting in TBE 1X at 8V and 4°C overnight. Then, RNA was cross-linked to the membrane by UV irradiation (302 nm) for 2 min in a UV-crosslinker. After pre-hybridization for 1 h in 15 ml modified Church Buffer at 42°C or 65°C for DNA or RNA probes, respectively, the radioactive labeled probe (2-5 pmol) was added. After a hybridization period of 3-18 h at 42°C or 65°C for DNA or RNA probes, respectively, the Membranes were washed two times 5 minutes in 2X SSC, 0.1% SDS, and revealed using a Pharos FX phosphorimager. For the detection of *aapA3* mRNA species the labeled primer FD38 was used (Table 8.1.4 in the Appendix). IsoA3 small RNA was detected with a riboprobe corresponding to the *aapA3*-Tr RNA species transcribed from a PCR fragment containing the T7 promoter amplified with the primer pair FA170/FA173 (Table 8.1.4 in the Appendix).

#### **6.2.3.7. Polysome fractionation in sucrose gradients**

*H. pylori* strains were grown as previously described. At an early exponential phase (OD<sub>600</sub><0.9), chloramphenicol (100 µg/ml) was added to the culture to stabilize translating ribosomes. After 5 min

incubation at 37°C, cultures were quickly cooled by transferring them into pre-chilled flasks immersed in a dry ice/ethanol bath. Cultures were then centrifuged for 10 min at 3,500 rpm and 4°C and pellets were washed with Buffer A and frozen at -80°C. Then, pellets were resuspended in 500 µl of Buffer A containing RNasin® Ribonuclease Inhibitor and cells were lysed with glass beads in a Precellys homogenizer. Lysates were recovered and immediately frozen in liquid nitrogen. About 10 OD<sub>260</sub> units of lysate were layered onto 10% - 40% sucrose gradients in Grad-Buffer and centrifuged at 35,000 rpm for 3.75 h at 4°C in a SW41 Ti rotor. Gradients were analyzed with an ISCO UA-6 detector with continuous monitoring at 254 nm. Fractions of 500 µl were collected and RNA was precipitated by adding 1 volume of ethanol containing 150 mM of sodium acetate pH 5.2 and stored at -20°C. RNA was extracted as described above and subjected to Northern Blot analysis following the above-described protocol.

#### 6.2.3.8. Cold and hot *in vitro* transcription

For cold *in vitro* synthesis of the *aapA3* and IsoA3 RNAs, DNA templates were amplified from *H. pylori* 26695 genomic DNA using primer pairs: FA170/FA175 (*aapA3*-FL), FA170/FA173 (*aapA3*-Tr), FD11/FD17 (IsoA3), each forward primer carrying a T7 promoter sequence (see primer list in Table 8.1.4 in Appendix). *In vitro* transcription was carried out using the MEGAscript® T7 Transcription Kit according to the manufacturer's protocol. After P:C:I (25:24:1 v/v) extraction followed by overnight isopropanol precipitation at -20°C and 30 min centrifugation at 13,000 rpm and 4°C, the RNA pellets were resuspended in H<sub>2</sub>O. Samples were desalted and unincorporated nucleotides were removed by gel filtration using a Sephadex G-25 1 ml column. RNA integrity was checked on PAA denaturing gel and visualized by Stains-All. Hot *in vitro* transcription, was performed as described above but in the presence of 50 µCi of <sup>32</sup>P-α-UTP and 1 mM cold UTP (instead of 10 mM used for all dNTPs during cold *in vitro* transcription). To generate 5'-monophosphate internally labelled RNAs, 30 mM GMP (8-fold excess of guanosine monophosphate over guanosine triphosphate to obtain uniformly labeled RNA that has a 5'-monophosphorylated end) was added to the hot *in vitro* transcription reaction. Unincorporated nucleotides were removed using a MicroSpin™ G-25 column and labeled RNA was purified on an 8% PAA containing 7 M urea 1X TBE gel. Upon visualization of the labeled RNA on an autoradiography film, the band corresponding to the RNA species of interest was cut from the gel and eluted overnight at 4°C under shaking in 750 µl RNA elution buffer. RNA was extracted by P:C:I (25:24:1 v/v), desalted and concentrated by ethanol precipitation, pellets were resuspended in 100 µl H<sub>2</sub>O and stored at -20°C.

Sequences of the T7 RNAs used in Chapter 4 sections 4.2 and 4.3 are shown in Table 6.3.2 and Table 6.4.2, respectively, in this Chapter.

#### 6.2.3.9. 5' end labeling of RNA

20 pmol RNA was dephosphorylated with 10 u of calf alkaline phosphatase (CIP) in a 20 µl reaction at 37°C for 1 h. RNA was isolated by phenol extraction and precipitated overnight at -20°C in the presence of 30:1 EtOH: NaOAc pH 5.2 and 20 µg GlycoBlue™. The dephosphorylated RNA was then 5' end-labeled with 10 pmol <sup>32</sup>P-γ -ATP using the T4 polynucleotide kinase (PNK) for 30 min at 37°C in a 20 µl reaction. Labeled RNA was gel purified as described in the subheading 6.2.3.8. of the General Methods section, resuspended in 50 µl H<sub>2</sub>O and stored at -20°C.

## 6.2.4. Protein techniques

### 6.2.4.1. One-dimensional SDS PAGE

Denaturing separation of protein samples was performed in 15% SDS-PAGE gels. Preparation was performed as follows:

#### PAA gel for 15 % resolving gel ( $V_r=30$ ml)

40% PAA solution (29:1 acrylamide/bisacrylamide)	11.25 ml
Tris-HCl 3M pH 8.45, 0.3% SDS	10 ml
100% Glycerol	3.2 ml
H <sub>2</sub> O	5.5 ml
10% APS	100 $\mu$ l
TEMED	30 $\mu$ l

#### PAA gel for 6 % stacking gel ( $V_r=15$ ml)

40% PAA solution (29:1 acrylamide/bisacrylamide)	2.25 ml
Tris-HCl 3M pH 8.45, 0.3% SDS	3.75 ml
100% Glycerol	3.2 ml
H <sub>2</sub> O	9 ml
10% APS	60 $\mu$ l
TEMED	37.5 $\mu$ l

Gels were run for 1h at 80 V (stacking gel) and for 2 h at 120 V on 1X cathode (upper tank) and 1X anode (lower tank) buffers.

### 6.2.4.2. *In vitro* translation assays

*In vitro* translation assays were performed as described in (Sharma et al., 2010) with minor modifications. For *in vitro* translation of the *aapA3*-FL and *aapA3*-Tr mRNAs, 0.5  $\mu$ g ( $\approx 7$  pmol *aapA3*-Tr, 8 pmol *aapA3*-FL) of cold *in vitro* transcribed RNA was added to the *E. coli* S30 Extract System for Linear Templates Kit in 25  $\mu$ l reactions containing 0.2 mM [<sup>35</sup>S]-methionine (1175 Ci/nmol; Perkin-Elmer), 7.5  $\mu$ l S30 extract, 10  $\mu$ l S30 premix without amino acids, and 0.1 mM of each amino acid minus methionine. Reactions were incubated for 1 h at 73°C, stopped with 100  $\mu$ l NaOH and precipitated with 1 ml of TCA (25%). After 30 min centrifugation at 13,000 rpm, pellets were washed twice with acetone, then resuspended in 20  $\mu$ l of 1X Tricine Sample Buffer, heat-denatured for 5 min at 90°C and separated in 16.5% Tris-tricine gels using Low-Range Rainbow molecular weight marker as size standard. Gels were run at 120V for 1h at 80 V (stacking gel) and for 2 h at 120 V, dried at 80°C for 2 h and revealed using a Pharos FX phosphorimager.

## 6.3. Methods Paper III

### 6.3.1. Bacterial strains and oligonucleotides

*H. pylori* strains used in Chapter 4 section 4.2 (Paper III) are listed in Table 8.1.1 in the Appendix. *H. pylori* chromosome manipulations and mutant strain generation were performed as described in the General Methods section. Selection of toxicity suppressor mutations and analysis by Illumina sequencing was performed as described in the General Methods section and in more detail in Chapter 4 section 4.1 (Paper II).

For the *in vivo* validation of suppressor mutants, single nucleotide mutations of interest were introduced by Site-Directed mutagenesis PCR using the pA3-Up and pA3-Down plasmids (Table 8.1.2 in the Appendix, see Table 6.3.1 for sequences details) as DNA templates and the DNA oligonucleotides listed in the Table 8.1.4 in the Appendix.

### 6.3.2. *In vitro* structure probing

Structure probing analyses were performed as described previously (Darfeuille et al., 2007; Sharma et al., 2007, 2010), using 0.1 pmol of 5' end-labeled RNA. Before use, each *in vitro* transcribed RNA was denatured by incubation at 90°C for 2 min in the absence of magnesium and salt, then chilled on ice for 1 min, followed by a renaturation step at room temperature for 15 min in 1X Structure Buffer. To determine the secondary structure of RNA, 0.01 u RNase T1 was added to the labeled RNA and incubated in 1X Sequencing Buffer for 5 min at 37°C. Lead acetate (5 mM final concentration) digestions of both *aapA3-Tr* and *aapA3-FL* were done in the absence or in the presence of 2-10 fold excess IsoA3 cold RNA. All the reactions were stopped by adding 10 µl of 2X Gel Loading Buffer II and stored at -20°C. Cleaved fragments were then analyzed on an 8% denaturing PAA gel containing 7M urea and 1X TBE. Gels were dried 45 min at 80°C, and revealed using a Pharos FX phosphorimager.

### 6.3.3. RNase H1/oligonucleotide accessibility assays

Before use, each *in vitro* transcribed RNA and DNA oligonucleotides were denatured by incubation at 90°C for 2 min in the absence of magnesium and salt, then chilled on ice for 1 min. RNA followed a renaturation step at 37°C for 5 min in 1X Structure Buffer. Next, DNA oligonucleotides complementary to the region around the SD sequence (FA633 for WT and T78C (aSD2) mRNA; FA644 for A40T mRNA (aSD1); FA651 for the double mutant 33A>T/A40T mRNA (aSD1); and FA652 for A28C mRNA; sequences listed in Table 8.1.4 in the Appendix) were added to a final concentration of 0 to 10 µM. Reactions were adjusted to a final volume of 10 µl with RNase free H<sub>2</sub>O and incubated for 30 min at 30°C in the presence or absence (control) of 0.25 u *E. coli* RNase H1. Reactions were then stopped by addition of 10 µl of 2X Gel Loading Buffer and stored at -20°C. Cleaved fragments were analyzed on an 8% denaturing PAA gel containing 7M urea and 1X TBE. Gels were dried 45 min at 80°C, and revealed using a Pharos FX phosphorimager.

**Table 6.3.1. Inserts of *rpsL<sub>C</sub>-erm* and *aapA3/IsoA3* plasmids used in Paper III.**

Restriction enzymes sites are highlighted in blue; promoters are highlighted in yellow; Transcriptional Start Site (TSS) is highlighted in yellow and in bold; *aapA3*-FL mRNA 3' end is highlighted in yellow and bold; Shine-Dalgarno sequences are highlighted in green; ORFs are shown in red; translation start and stop sites are highlighted in bold; mutations are underlined. Cloning primers determine the borders of the insert, and add when need it the restriction sites to allow cloning. SDS primers are the ones used for Site-Directed mutagenesis PCR. All nucleotide positions (mutations) are relative to Aapa3 TSS. A3 refers to the *aapA3/IsoA3* chromosomal locus of the *H. pylori* 26695 strain.

Name	Plasmid/ Origin/ Marker	Insert (nt)	Cloning primers	SDM primers
pSP60-2 ( <i>rpsL<sub>C</sub>-erm</i> )	pSP60/ pSC101* / Amp <sup>R</sup>	GGATCCTGCTTTATAACTATGGATTAAACACTTTTTAGTAAT TTCTTAAATCAATTTTGAATTTTTCATTTTGTCTAAAATGA GTAAGATTATTTTATAAGTAAACAATAATTAATAAATTAGT TTATGTTTATATAATATTAAATTATGTAGGTTTATTTT ATAAAAAATTTTAAAGCATTTAAATTATTCTAGTTTATATT TTTGTAAGTAAATTTTAGATATCATCCAAAGTTTATTACATTT TTAAAGAAAGGAATTATTGTGCCTACCATAAATCAATTGGTT AGAAAAGAGCGCAAAAAGTTTGTAGAAAAATCTAAATCTCC AGCGCTTAAAAATTGTCCACAAAGAAGGGGAGTTTGCCTA GGGTTTATACTACAACACCTAAAAAACCAAACTCAGCGTTA AGAAAAGTTGCCAAAGTAAGACTTACTAGTGGCTTTGAAGT GATCAGCTATATCGGCGGTGAAGGTCATAACTTGCAAGAAC ACAGCATTGTTTGTGCGTGGTGGTAGGGTAAAAGACTTAC CAGGGGTTAAATATCACATCGTTCGTGGTGCTTGTGATACAG CAGGTGTGCAAAAAGAACAGTTTCTCGTTCTAAATATGGTG CTAAACGTCCTGAAGCAGGCGCTGCAAAATAATCATTACATAC AGACAAATCCGTTAGACCAATAATCGCATCAGATTGCAGTAT AAATTAACGATCACTCATCATGTATATTATCAGAGCTCGTG CTATAATTATACTAATTTTATAAGGAGGAAAAATAAAGAG GTTTATAATGAACGAGAAAAATATAAACACAGTCAAAACT TTATTACTTCAAAACATAATATAGATAAAATAATGACAAATA TAAGATTAAATGAACATGATAATATCTTTGAAATCGGCTCAG GAAAAGGGCATTTTACCCTTGAATTAGTACAGAGGTGTAATT TCGTAACGCAATTGAAATAGACCATAAATTATGCAAAACTA CAGAAAAATAAATCTGTTGATCAGGATAATTTCCAAGTTTAA ACAAGGATATATTGCAGTTTAAATTTCTAAAAACCAATCCT ATAAAATATTGGTAATATACCTTATAACATAAGTACGGATA TAATGCGCAAAATTGTTTTGATAGTATAGCTGATGAGATTT ATTTAATCGTGGAATACGGGTTTGCTAAAAGATTATTAATA TAAAACGCTCATCGGCATTATTTTAAATGGCAGAAAGTTGATA TTTCTATTAAGTATGGTCCCAAGAGAATATTTTCATCCTA AACCTAAAGTGAATAGCTCACTTATCAGATTAATAGAAAA AAATCAAGAATATCACACAAAGATAAACAGAAAGTAAATTA TTTCGTTATGAAATGGGTTAACAAAGAATACAAGAAAATATT TACAAAAAATCAATTTAACAATTCCTTAAACATGCAGGAAT TGACGATTTAAACAATATTAGCTTTGAACAATTCCTATCTCT TTCAATAGCTATAAATTATTAAATAAAGTAAGAATTC (1545)	(Pernitzsch et al., 2014)	-
pA3- Up WT	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCAACCACTTACTACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAAGCACACCAACAACAACA CAAAAAGCGCAACAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTAATCTAGTAAATTTATCCATTTAACTAACTTTTTTA ACTCCGTTAGATTTTAACTTTTAAATCCCTCTTTTGTATG CTTGTCAAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTAACAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTTAAGGGGGTTTGT GTAGGATTTTCATCACGCCCATAGTTGCAATATGGGGCAAA CCTATGTCAAAAGGAGCATGCCATGAAACACAAAAGTGGCA AACGCTCTTG (638)	FA406/ FA386	-
pA3- Down WT	pGEM-T/ CoE11/ Amp <sup>R</sup>	CACAAAAGTGGCAACGCTCTTGGAACAAATTTACTTTGA GTTTGCTTTTGGGGCTTAAAGTGATCGTTCTGTGAAACGC TGATTTTCTAAGAGCGTTCCCTTAAGCTTTTAAGCTTAGGG GTTTCCTTATGAAGCGGGTTGTCTTGACTTTTGGGGGCTTG GTCATTTTGTGTAACCTTATTTTCTTTAGGGGATAGAAAT AATTCTCCCTCTATAAACTAAATCCCTCATAGAGTGCTT	FA387/ FA409	-



		TTAAAAAATACCGCTTGCTCTAGCTTTTGGATTATTTGTCTTA AAAACTCACTCCCTTAAAAACCCGTATTCATAGACTTTAGGCT GTATTTCTTTGGCGAGATCTTTTAGGGCGTCTTGCAAATTAG CGTAAAGCTGTTTGTAATATTCATCTTTAGCTTTAGCGGGTA AATTTAATTTGCCTTGCTTGTTACGGCCTTGAAAAAATCTTT GATGTCATAAAGGCTTGCGTTAGCGTTATAGGGGCGGTTTGT GAAATCTTGTGTGGTAATAGCGATAAACCTCTCTGCTAGC GTCAAACACCTTTGAAGCGCTCGGGCTGAATTTAGGGGCTT GTTTTCTTTTGGCACTTAAAAAGAGGCTATCGTTCTCTTA AGTTCTTTGATTTGCGCTTTTAAAAACTCTAATAGAGCGTGG CTAG (680)		
pA3- Up start	pGEM-T/ CoEl1/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACA CAAAAAGCGCAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTAATCTAGTAAATTTATCCATTTAACTAACTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGATG CTTGTCAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTAATAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTTAAGGGGGTTGTT GTAGGATTTTCATCAACGCCCATAGTTGCAATATGGGGCAAAAT CCTATGTCAAAAGGAGCATGCCATTAAACACAAAAGTGGCA AACGCTCTTGG (638)	FA406/ FA386	FA281/ FA282
pA3- Down pIsoA3 *	pGEM-T/ CoEl1/ Amp <sup>R</sup>	CACAAAAGTGGCAAACGCTCTTGGAAAAACGTTCTACTTTGAG TTTGTCTTTTGGGGCTTAAAGTGATCGTTTCTGTGAAACGCT GATTTTTCTAAGAGCGTTCCTTAAGCTTTAAGCTTAGGGG TTTCCTTATGAAGCGGGTTTGCTCTGACTTTTGGGGCGTTGG TCATTTTGTGTAAAACTATTTTCTTTAGGGGATAGAAATA ATTCTCCCTCTATAAACTAAATCCCTCATAGAGTGCTTTT AAAAAATACCGCTTGCTCTAGCTTTTGTATTATTGTCTTAAA AATCACTCCCTTAAAAACCCGTATTCATAGACTTTAGGCTGT ATTTCTTTGGCGAGATCTTTAGGGCGTCTTGCAAAATTAGCG TAAAGCTGTTGTAAATATTCATCTTTAGCTTTAGCGGGTAAA TTTAATTGCTTGCTTGTTACGGCCTTGAAAAAATCTTTGA TGTCATAAAGGCTTGCGTTAGCGTTATAGGGGCGGTTTGTA AATCTTGTGTGTGGTAATAGCGATAAACCTCTCTGCTAGCGT CAAACACCTTTGAAGCGCTCGGGCTGAATTTAGGGGCTTGT TTCTTTTGGCACTTAAAAAGAGGCTATCGTTCTCTTTAAG TTCTTTGATTTGCGCTTTTAAAAACTCTAATAGAGCGTGGCTA G (680)	FA387/ FA409	FA283/ FA284
pA3- Up A28C	pGEM-T/ CoEl1/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACA CAAAAAGCGCAACAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTAATCTAGTAAATTTATCCATTTAACTAACTTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGATG CTTGTCAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTAATAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTTAAGGGGGTTGTT GTAGGATTTTCATCAACGCCCATAGTTGCAATATGGGGCAAAAT CCTATGTCAAAAGGAGCATGCCATTGAAACACAAAAGTGGCA AACGCTCTTGG (638)	FA406/ FA386	FA465/ FA466
pA3- Up A33T	pGEM-T/ CoEl1/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACA CAAAAAGCGCAACAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTAATCTAGTAAATTTATCCATTTAACTAACTTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGATG CTTGTCAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTAATAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTTAAGGGGGTTGTT GTAGGATTTTCATCAACGCCCATAGTTGCAATATGGGGCAAAAT CCTTTGTCAAAAGGAGCATGCCATTGAAACACAAAAGTGGCA AACGCTCTTGG (638)	FA406/ FA386	FA546/ FA547

pA3-Up A40T	pGEM-T/ CoE11/ Amp <sup>R</sup>	AACGCTCTTGG (638) GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACAA CAAAAAGCGCAACAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTTAATCTAGTAAATTTATCCATTTAAACTAACTTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGTATG CTTGTCAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTA AAAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTAAAGGGGGTTTGT G TAGGATTTTCATCAACGCCCATAGTTGCAATATGGGGCAAAAT CCTATGTCAATAGGAGCATGCCATGAAACACAAAAGTGGCA AACGCTCTTGG (638)	FA406/ FA386	FA467/ FA468
pA3-Up A33T A40T	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACAA CAAAAAGCGCAACAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTTAATCTAGTAAATTTATCCATTTAAACTAACTTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGTATG CTTGTCAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTA AAAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTAAAGGGGGTTTGT G TAGGATTTTCATCAACGCCCATAGTTGCAATATGGGGCAAAAT CCTTTGTCAATAGGAGCATGCCATGAAACACAAAAGTGGCA AACGCTCTTGG (638)	FA406/ FA386	FA548/ FA549
pA3-Up T78C	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGGTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACAA CAAAAAGCGCAACAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGCGGCCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTTAATCTAGTAAATTTATCCATTTAAACTAACTTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGTATG CTTGTCAAATGGCTCTGTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTTCGTAATGGGGTGTA AAAACTTTGTCTT GCGTATTTGATGCGGTCTAGGGGTGTTTAAAGGGGGTTTGT G TAGGATTTTCATCAACGCCCATAGTTGCAATATGGGGCAAAAT CCTATGTCAAAAGGAGCATGCCATGAAACACAAAAGTGGCA AACGCTCCTGG (638)	FA406/ FA386	FA511/ FA512
pA3-Down T78C	pGEM-T/ CoE11/ Amp <sup>R</sup>	CACAAAAGTGGCAACGCTCCTGGAAAACATTATACTTTGA GTTTGCTTTTTTGGGGCTTAAAGTGATCGTTTCTGTGAAACGC TGA TTTTCTAAGAGCGTTCCCTTAAGCTTTTAAGCTTAGGG GTTTCCTTATGAAGCGGGTTTGTCTTGACTTTTGGGGCGTTG GTCATTTTGTGTAAAACCTATTTTCTTTAGGGGATAGAAAAT AATTCTCCCTCTATAAAAACCTAAATCCCTCATAGAGTGCTT TTAAAAAATACCGCTTGCTCTAGCTTTTGTATTATTTGTCTTA AAAATCACTCCCTTAAAAACCGTATTCATAGACTTTAGGCT GTATTTCTTTGGCGAGATCTTTTAGGGCGCTTTGCAAATTAG CGTAAAGCTGTTGTAAATATTCATCTTTAGCTTTAGCGGGTA AATTTAATTTGCCTTGCTTGTACGGCCTTGAAAAAATCTTT GATGTCATAAAGGCTTGCGTTAGCGTTATAGGGGCGGTTTGT GAAATCTTGTGTGTGTAATAGCGATAAACCTCTCTGCTAGC GTCAAACACCTTTGAAGCGCTCGGGCTGAATTTACGGGGCTT GTTTTCTTTTTTGGCACTTAAAAAGAGGCTATCGTTCTTTTA AGTTCTTTGATTTTCGCTTTTAAAAACTCTAATAGAGCGTGG CTAG (680)	FA387/ FA409	FA511/ FA512
pA3-Down T78C/p IsoA3*	pGEM-T/ CoE11/ Amp <sup>R</sup>	CACAAAAGTGGCAACGCTCCTGGAAAACGTTCTACTTTGA GTTTGCTTTTTTGGGGCTTAAAGTGATCGTTTCTGTGAAACGC TGA TTTTCTAAGAGCGTTCCCTTAAGCTTTTAAGCTTAGGG GTTTCCTTATGAAGCGGGTTTGTCTTGACTTTTGGGGCGTTG GTCATTTTGTGTAAAACCTATTTTCTTTAGGGGATAGAAAAT AATTCTCCCTCTATAAAAACCTAAATCCCTCATAGAGTGCTT TTAAAAAATACCGCTTGCTCTAGCTTTTGTATTATTTGTCTTA AAAATCACTCCCTTAAAAACCGTATTCATAGACTTTAGGCT GTATTTCTTTGGCGAGATCTTTTAGGGCGCTTTGCAAATTAG CGTAAAGCTGTTGTAAATATTCATCTTTAGCTTTAGCGGGTA AATTTAATTTGCCTTGCTTGTACGGCCTTGAAAAAATCTTT GATGTCATAAAGGCTTGCGTTAGCGTTATAGGGGCGGTTTGT GAAATCTTGTGTGTGTAATAGCGATAAACCTCTCTGCTAGC GTCAAACACCTTTGAAGCGCTCGGGCTGAATTTACGGGGCTT GTTTTCTTTTTTGGCACTTAAAAAGAGGCTATCGTTCTTTTA AGTTCTTTGATTTTCGCTTTTAAAAACTCTAATAGAGCGTGG CTAG (680)	FA387/ FA409	FA511/ FA512

		CGTAAAGCTGTTTGTAAATATTCATCTTTAGCTTTAGCGGGTA AATTTAATTTGCCTTGCTTGTACGGCCTTGAAAAAATCTTT GATGTCATAAAAGGCTTGCGTTAGCGTTATAGGGGCGGTTTGT GAAATCTTGTGTGTGGTAATAGCGATAAACCTCTCTGCTAGC GTCAAACACCTTTGAAGCGCTCGGGCTGAATTTACAGGGGCTT GTTTTCTTTTTTGGCACTTAAAAAGAGGCTATCGTTCTCTTTA AGTTCTTTGATTTTCGCCTTTTAAAACTCTAATAGAGCGTGG CTAG (680)		
pA3- Up G43A	pGEM-T/ CoEI1/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTAAA TTTCAATTAAAGGAGTCATCATGGCACACCATGAACAACAA CAGCAACAACAGGCTAACAGCCAACACCACCACCATCACCA TGCGCACCACCACCATTAACGCTGGCGAACACCATCACCA TAATGCGCAACAACACGCCGAACAACAAGCAGAGCAACAAG CTCAGCAACAGCAACAACAACAAGCACACCAACAACAACAA CAAAAAGCGCAACAACAACAACAATATTGATTGGGGCG TTTGTGGGGGCGGCTTAGGGCTACTCCTAGCTTTTAACTCTT TCTTTTAATCTAGTAAATTTATCCATTTAACTAACTTTTTTA ACTCCGTTAGATTTTTAACTTTTAAAAATCCCTCTTTTGTATG CTTGTCAAATGGCTCTGTTTTCTATCCAATAAAGATAAGCCT TATTAATTGAGTGGTCGTAATGGGGTGTA AAAA ACTTTGTCTT GCGTATTGATGCGGTCTAGGGGTGTTTTAAGGGGGTTTGT GTAGGATTTTCATCAACGCCCCATAGTTGCAATATGGGGCAAAT CCTATGTCAAAAGAAGCATGCCATGAAACACAAAAGTGGCA AACGCTCTTGG (638)	FA406/ FA386	FA535/ FA536

**Table 6.3.2. Sequence of T7 RNAs used for in vitro studies in Paper III.**

Lower-case letter indicated the T7 promoter used. Red letters indicate ORFs. Translation start and stop codons are shown in bold. Mutated positions are underlined. nt\* indicates the length in nucleotides of the RNA (excluding the T7 promoter sequence).

T7 RNA	Sequence (5'→3') (nt*)	Primer pair	Template
<i>aapA3</i> -FL WT	gaaattaatagcactcactataggACGCCCCATAGTTGCAA TATGGGGCAAATCCTATGTCAAAAGGAGCAT GCC <b>ATGAAACACAAAAGTGGCAAACGCTCTT</b> <b>GGAAAACATTATACTTTGAGTTTGCTTTTTTG</b> <b>GGGCTTAAAGTGATCGTTTCTGTGAAACGCT</b> <b>GA</b> TTTTTCTAAGAGCGTTCCCTTAAGCTTTTA AGCTTAGGGGTTTCCTTATGAAGCGGGTTTGT CCTTGACTTTTGGGGCGTTGGTC (229)	FA170/FA175	gDNA <i>H. pylori</i> 26695 WT
<i>aapA3</i> -Tr WT	gaaattaatagcactcactataggACGCCCCATAGTTGCAA TATGGGGCAAATCCTATGTCAAAAGGAGCAT GCC <b>ATGAAACACAAAAGTGGCAAACGCTCTT</b> <b>GGAAAACATTATACTTTGAGTTTGCTTTTTTG</b> <b>GGGCTTAAAGTGATCGTTTCTGTGAAACGCT</b> <b>GA</b> TTTTTCTAAGAGCGTTCCCTTAAGCTTTTA AGCTTAGGGGTTTCCT (190)	FA170/FA173	gDNA <i>H. pylori</i> 26695 WT
<i>aapA3</i> -Tr A28C	gaaattaatagcactcactataggACGCCCCATAGTTGCAA TATGGGGCAAATCCTATGTCAAAAGGAGCAT GCC <b>ATGAAACACAAAAGTGGCAAACGCTCTT</b> <b>GGAAAACATTATACTTTGAGTTTGCTTTTTTG</b> <b>GGGCTTAAAGTGATCGTTTCTGTGAAACGCT</b> <b>GA</b> TTTTTCTAAGAGCGTTCCCTTAAGCTTTTA AGCTTAGGGGTTTCCT (190)	FA170/FA173	gDNA <i>H. pylori</i> 26695 <i>aapA3</i> A28C/ pIsoA3*
<i>aapA3</i> -Tr A40T	gaaattaatagcactcactataggACGCCCCATAGTTGCAA TATGGGGCAAATCCTATGTCAATAGGAGCAT GCC <b>ATGAAACACAAAAGTGGCAAACGCTCTT</b> <b>GGAAAACATTATACTTTGAGTTTGCTTTTTTG</b> <b>GGGCTTAAAGTGATCGTTTCTGTGAAACGCT</b> <b>GA</b> TTTTTCTAAGAGCGTTCCCTTAAGCTTTTA AGCTTAGGGGTTTCCT (190)	FA170/FA173	gDNA <i>H. pylori</i> 26695 <i>aapA3</i> A40T/ pIsoA3*
<i>aapA3</i> -Tr A33T A40T	gaaattaatagcactcactataggACGCCCCATAGTTGCAA TATGGGGCAAATCCTTTGTCAATAGGAGCAT GCC <b>ATGAAACACAAAAGTGGCAAACGCTCTT</b> <b>GGAAAACATTATACTTTGAGTTTGCTTTTTTG</b> <b>GGGCTTAAAGTGATCGTTTCTGTGAAACGCT</b> <b>GA</b> TTTTTCTAAGAGCGTTCCCTTAAGCTTTTA AGCTTAGGGGTTTCCT (190)	FA170/FA173	gDNA <i>H. pylori</i> 26695 <i>aapA3</i> A33T A40T/ pIsoA3*
<i>aapA3</i> -Tr T78C	gaaattaatagcactcactataggACGCCCCATAGTTGCAA TATGGGGCAAATCCTATGTCAAAAGGAGCAT GCC <b>ATGAAACACAAAAGTGGCAAACGCTCTT</b> <b>GGAAAACATTATACTTTGAGTTTGCTTTTTTG</b> <b>GGGCTTAAAGTGATCGTTTCTGTGAAACGCT</b> <b>GA</b> TTTTTCTAAGAGCGTTCCCTTAAGCTTTTA AGCTTAGGGGTTTCCT (190)	FA170/FA173	gDNA <i>H. pylori</i> 26695 <i>aapA3</i> T78C/pIsoA3*
IsoA3 WT	CAAGAGCGTTTGCCACTTTTGTGTTTCATGGC ATGCTCCTTTTGACATAGGATTTGCCCATAT TGCAACTATGGGGCGT (80)	FD110/FD17	gDNA <i>H. pylori</i> 26695 WT

## 6.6. Methods Paper IV

### 6.4.1. Bacterial strains and oligonucleotides

Bacterial strains and oligonucleotides used in the Chapter 4 section 4.3 (paper IV) are listed in the Table 8.2.3 and 8.2.6 in the Appendix, respectively. Sequences of the pGEM-T cloned inserts of the *aapA3*/IsoA3 locus mutated variants are shown in Table 6.4.1 in this Chapter.

#### **Generation of the *H. pylori* PNPase and PNPase/*aapA3*/IsoA3 deleted strains**

To generate the *pnp* deletion in the *H. pylori* 26995/*rpsL*K43R strain, the PNPase residues 1-658 were replaced by the counterselection marker *rpsL<sub>CJ</sub>-erm* in a *rpsL*K43R genetic background as described by (Dailidiene et al., 2006). PCR-construction was performed as described in the General Methods. In detail, the primer pairs FD699/FA109 and FA112/FD702 (Table 8.2.6 in the Appendix) were used to amplify the upstream and downstream to *pnp* deletion regions and introduce the *rpsL<sub>CJ</sub>-erm* homology tails to allow the subsequent PCR assembly. The *rpsL<sub>CJ</sub>-erm* cassette was amplified from the pSP60-2 plasmid (Table 8.2.5 in the Appendix) with the primer pair FA110/FA111. Fragments were assembled by PCR using the primer pair FD699/FD702. This PCR-construct was then used to perform natural transformation of the *H. pylori* 26995/*rpsL*K43R strain and generate the  $\Delta pnp::rpsL_{CJ}\text{-erm}/rpsL_{K43R}$  strain (Table 8.2.3 in the Appendix). The *rpsL<sub>CJ</sub>-erm* cassette was excised from the *pnp* locus by transforming the *H. pylori* 26995  $\Delta pnp::rpsL_{CJ}\text{-erm}/rpsL_{K43R}$  strain with a PCR fragment generated with the primer pair DD699/FA113. This lead to the generation of a ‘clean’ *pnp* deleted strain (lacking the residues 1-658), the  $\Delta pnp/rpsL_{K43R}$  strain (Table 8.2.3 in the Appendix). The  $\Delta pnp/rpsL_{K43R}$  strain was used as background to generate the double deletion strain by deleting the *aapA3*/IsoA3 strain using the *rpsL<sub>CJ</sub>-erm* cassette as described in the General Methods section subheading 6.2.2.3.

#### **Generation of the *H. pylori* RNase Y, RhpA and RppH deletion strains**

To generate the *H. pylori* 26695 RNase Y deleted strain, the primer pairs FD631/FD632 and FD633/FD634 (Table 8.2.6 in the Appendix) were used to amplify the up- and downstream to *rny* deletion regions and introduce the *aphA3* homology tails to allow the subsequent PCR assembly. The *aphA3* cassette was amplified from the pUCK18K2 plasmid (Table 8.2.5 in the Appendix) with the primer pair FD70/FD255 (Table 8.2.6 in the Appendix). Fragments were assembled by PCR using the primer pair FD631/FD634 and the product was used to transform the *H. pylori* 26695 strain generating the  $\Delta rny::aphA3$  strain (Table 8.2.3 in the Appendix). To obtain the double RNase Y/PNPase deletion mutant ( $\Delta rny::aphA3/(\Delta pnp::catCG$ , Table 8.2.3 in Appendix), the *pnp* gene was replaced by the *catCG* gene (amplified from pILL2150, Table 8.2.5 in Appendix) in the  $\Delta rny::aphA3$  genetic background. The *H. pylori* 26695 RhpA and RppH deletion strains were generated by replacing the *rhpA* or the *rppH* genes by the *aphA3* gene or the *rpsL-erm* cassette, respectively, using standard PCR assembly (primers listed in Table 8.2.6 in Appendix) and natural transformation protocols.

#### **Expression of the *aapA3*/IsoA3 locus in *E. coli***

To express the *H. pylori* 26995 *aapA3*/IsoA3 locus in *E. coli* cells, a PCR fragment containing the full locus sequence (1294 nt-long amplicon) was amplified with the primer pair FA406/FA409 (Table 8.2.6 in the Appendix) using the Dream Taq DNA polymerase, which adds 3’ adenine overhangs and cloned into pGEM-T vector allowing T-A cloning strategy (*aapA3*/IsoA3 locus expression occurs from *H. pylori* endogenous promoters). The ligation product was then used to transform TOP10 chemocompetent cells and the strains were stored in glycerol -80°C stocks upon Sanger sequencing verification of the inserts.



***In vivo validation of mRNA decay-related suppressors***

The  $\Delta pnp/\Delta aapA3/IsoA3::rpslC_7-erm/rpslK43R$  strain was used as genetic background to assess the effect of the selected mutations in the 3'-terminal hairpin of the *aapA3* mRNA. In detail, the *aapA3/IsoA3* locus (1294 nt-long amplicon) containing the G183A mutation was amplified from gDNA of an isolated G183A suppressor strain using the primer pair FA406/FA409 (Table 8.2.6 in Appendix). This PCR-product was used substrate for natural *H. pylori* transformation in both *+pnp* and *-pnp* genetic backgrounds, achieving the allelic replacement of the *rpslC\_7-erm* cassette present in the *aapA3/IsoA3* locus. As the G183A mutation was expected to be toxic in the absence of PNPase, a mutation in the AapA3 start codon (G54T) was introduced. To that aim, the G54T mutation was introduced by Site-Directed mutagenesis PCR in a previously isolated pA3-Up plasmid (Table 8.2.5 in Appendix). PCR assembly using the primer pair FA406/FA409 (Table 8.2.6 in Appendix) was used to construct the *aapA3*G183A G54A/pIsoA3\* fragment used for *H. pylori* *+pnp* and *-pnp* genetic background transformation. This same protocol was used to perform chromosomal targeted mutagenesis to introduce the different point mutations here studied. Note that the T128C –Val26Ala-mutant was an isolated suppressor therefore it was directly split-cloned in pGEM-T as described above (with no need of Site-Directed mutagenesis PCR).

**6.4.2. Rapid amplification of cDNA 3' ends (3'RACE)**

Mapping of mRNA 3'ends by 3'RACE was performed following the protocol for Term-seq library preparation described in Dar et al., (2016) with minor modifications. In bacteria, contrary to the 5'-triphosphate end of primary transcripts, the 3' end carries a 3'hydroxyl group that can be directly ligated to the 5'monophosphate of a DNA oligonucleotide (3' adapter). In detail, 5 µg DNase I treated total RNA was subjected to 3'end specific ligation by mixing with 1.5 µl of 100 µM adapter solution (FA609, see Table 8.2.6 in Appendix), 2.5 µl 10X T4 RNA ligase buffer (containing 10 mM ATP), 2 µl DMSO, 9.5 µl 50% PEG4000 and 2.5 µl T4 RNA ligase enzyme in a final volume reaction of 25 µl. The reaction was incubated at 23°C for 2.5 h. Then, the DNA-RNA hybrid was extracted by P:C:I (25:24:1 v/v) followed by overnight precipitation at -20°C in the presence of 2.5 volumes of 30:1 (v/v) EtOH:3M NaOAc pH 5.2 and resuspended in 10 µl H<sub>2</sub>O. For cDNA synthesis, 10 µl ligation product was mixed with 0.5 µl of 100 µM FA120 deoxyoligonucleotide (50 pmol) (Table 8.2.6 in Appendix), 2 µl of 10 mM dNTPs solution and incubated at 65°C for 5 min then chilled on ice for 2 min. Next, 4 µl of 5X First Strand buffer, 1 µl of 100 mM DTT, 1 µl RNaseOUT™ (10 u) and 2 µl of SuperScriptII enzyme (20 u) were added to the mix. Reaction was incubated for 30 min at 37°C. Subsequently, 3.25 µl of the cDNA samples were used as DNA template in standard DreamTaq DNA polymerase reaction using the FA120 reverse primer that anneals with the adapter sequence and the FD49 gene specific forward primer. Following visualization in 2% agarose gel, PCR products were cloned into pGEM®-T *E. coli* vectors and Sanger-sequenced.

**6.4.3. RNA chemical synthesis and purification**

The AapA3 5'- wild type and G24A mutant, and the 3'-wild type and G183A mutant terminal stem-loops were synthesized in RNA chemistry using 1 µmol synthesis scale by Brune Violet at the University of Bordeaux, France. Samples were mixed with one volume of Gel Loading Buffer II, denatured at 90°C for 1 min and loaded on a 10 % PAA gel containing 8 M urea and 1X TBE. Gel run for 2 h at 17W in 0.5X TBE. After identifying the RNA species of interest by UV shadowing (254 nm), the RNA was cut from the gel and passively eluted overnight at 4°C in 1 ml RNA elution-buffer-

NH<sub>4</sub>Ac. Gel pieces were removed by using Spin-X®-centrifuge tubes filter. Then RNA was precipitated overnight in the presence of 2.5 volumes of EtOH 100 %. After centrifugation at 13,000 rpm and 4°C for 30 min, RNA pellets were washed with 75 % EtOH and resuspended in 50 µl of H<sub>2</sub>O. RNA integrity was checked on PAA denaturing gel, visualized by Stains-All and 20 µM were used for mass-spectrometry confirmation. Sequences of the chemically synthesized RNAs are shown in the Table 6.4.3.

#### 6.4.4. UV melting experiments

Thermal denaturation of *aapA3* 5'- and 3'-terminal hairpins was performed as described in (Ducongé et al., 2000) with slight modifications. 1 µM final concentration of chemically synthesized, gel purified RNA was denatured for 1 min at 90°C and chilled on ice for 10 min. Then, the RNA was mixed with Caco-buffer pH 7.3 (20 mM sodium cacodylate, 140 mM KCl, 20 mM NaCl, 1 mM MgCl<sub>2</sub>) in a final volume reaction of 250 µl. Denaturation of the samples was achieved by increasing the temperature at 0.4°C/min from 5°C to 90°C. Melting temperatures were determined as the maximum of the first derivate of the UV melting curves.

#### 6.4.5. RNase J overexpression and purification

For the *rnj* gene cloning, a fragment containing the full-length RNase J ORF (residues 1-693) was amplified by PCR using the primer pair FA369/FA370 (Table 8.2.6 in Appendix) and genomic DNA from the *H. pylori* B128 strain as template. Ligation of sequentially NdeI and BlnI digested PCR products, purified with the PCR and DNA fragment purification Kit, with sequentially NdeI and BlnI digested pET-15b plasmid was performed using the T4 DNA ligase following the manufacturer instructions. This plasmid allows the expression of a fusion protein containing a hexa-histidine tag and thrombine cleavage site at the N-terminus of RNase J. The ligation product (named pRNaseJ, Table 8.2.5 in Appendix) was subcloned into *E. coli* TOP10 cells. For overexpression of the recombinant RNase J, *E. coli* BL21 (DE3) Rosetta2 cells carrying the pET-15b-RNaseJ plasmid were grown in Terrific Broth supplemented 100 µg/ml ampicillin and 25 µg/ml chloramphenicol at 37°C and 200 rpm shaking. Induction of protein expression was initiated by addition of 1 mM IPTG at an OD<sub>600</sub> ≈ 1.5–2 and expression continued for 3 h at 37°C and 200 rpm shaking. Cells were harvested by centrifugation at 4,000 g for 15 min, and the cell pellet was frozen at –20°C overnight. Cell pellets were then lysed by sonication and centrifuged at 20,000 g for 45 min at 4°C. The supernatant was applied to a nickel-affinity gel for 1 h at 4°C. After extensive washes with Washing-buffer (25 mM Tris, pH 8.0 and 300 mM NaCl), the protein was eluted with a gradient of imidazole (0–500 mM). Elution fractions corresponding to the protein were then pooled. A second chromatography step was performed using a heparin HiTrap Heparin-HP affinity column with a linear gradient of 300 mM–1 M NaCl.

#### 6.4.6. RNase J *in vitro* cleavage assays

For RNase J-6XHist-Nter exonucleolytic activity assays, 1 µg PCR product (amplified with the primer pair FA170/FA624, Table 8.2.6 in Appendix) containing the *aapA3* mRNA 5'-terminal stem-loop (5'SL, 47 nt) with wild type or G24A mutated sequence and carrying the T7 promoter was used as template for hot *in vitro* transcription in the presence of 50 µCi of <sup>32</sup>P-α-UTP and 30 mM UMP as described in the 8.2.3.6 subheading of the General Methods section. Kinetics reactions were performed in 10 µl reaction volume using uniformly <sup>32</sup>-P labeled 5'monophosphate RNA and 28 µM *H. pylori* RNase J, purified as described in this section subheading 6.4.6. Reactions were incubated at 37°C during 0-30 min and time points were taken every 5 min. Labeled RNA was denatured at 90°C

for 1 min in the absence of salt and magnesium, then chilled on ice for 1 min, followed by the addition of 2.2 µl 5X RNase J reaction buffer and an excess of cold 5'SL RNA (47 nt). Reactions were stopped by the addition of 10 µl 2X Gel Loading Buffer II, run on a 20% denaturing PAA gel and analyzed using a Pharos FX phosphorimager.

#### 6.4.7. Mapping of RNase J and RNase III endonucleolytic cleavages

20 pmol IsoA3 or *aapA3*-Tr *in vitro* transcribed RNA was dephosphorylated with 10 u of calf alkaline phosphatase (CIP) in a 20 µl reaction at 37°C for 1 h. RNA was isolated by phenol extraction and precipitated overnight at -20°C in the presence of 30:1 EtOH: NaOAc pH 5.2 and 20 µg GlycoBlue™. The dephosphorylated RNA was then 5' end-labeled with 10 pmol <sup>32</sup>P-γ -ATP using the T4 polynucleotide kinase (PNK) for 30 min at 37°C in a 20 µl reaction. Labeled RNA was gel purified, resuspended in 50 µl H<sub>2</sub>O and stored at -20°C. Before use, 0.1 pmol of 5' end-labeled RNA was denatured by incubation at 90°C for 2 min in the absence of magnesium and salt, then chilled on ice for 1 min, followed by a renaturation step at 37°C for 5 min. For RNase J cleavage assays, RNA was incubated at 37°C in the presence of 1µg Yeast RNA and 28 µM *H. pylori* RNase J in 1X RNase J-endo reaction buffer. For RNase III assays, reactions were incubated at 37°C for 5 min in the presence of 0.1, 0.01, or 0.001 u RNase III in 1X Structure Buffer. Reactions were stopped by the addition of 10 µl 2X Gel Loading Buffer II at the indicated time points. Cleavage products were run on a 6% denaturing PAA gel and analysed using a Pharos FX phosphorimager.

**Table 6.4.1. Inserts of *aapA3*/IsoA3 mutant plasmids used in Paper IV.**

Promoters are highlighted in yellow; Transcriptional Start Site (TSS) is highlighted in yellow and bold; *aapA3*-FL mRNA 3' end is highlighted in yellow and bold; Shine-Dalgarno sequences are highlighted in green; ORFs are shown in red; translation start and stop sites are highlighted in bold; mutations are underlined. Cloning primers determine the borders of the insert, and add when need it the restriction sites to allow cloning. SDS primers are the ones used for Site-Directed mutagenesis PCR. All nucleotide positions (mutations) in the *aapA3*/IsoA3 constructs are relative to Aapa3 TSS. A3 refers to the *aapA3*/IsoA3 chromosomal locus of the *H. pylori* 26695 strain.

Name	Plasmid/ Origin/ Marker	Insert (nt)	Cloning primers	SDM primers
pA3-Up G3C	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATT AATTTCAATTAAAGGAGTCATCATGGCACACCATGAACAA CAACAGCAACAACAGGCTAACAGCCAACACCACCACCATC ACCATGCGCACCACCACCATTAACGTTGGCGAACACCA TCACCATAATGCGCAACAACACGCCGAACAACAAGCAGAG CAACAAGCTCAGCAACAGCAACAACAAGCACACCAA CAACAACAACAAAAAGCGCAACAACAACCAACAATAT TGATTGGGGCGTTTGTGGGGCGGCCTTAGGGCTACTCCTA GCTTTAACTCTTTCTTTAATCTAGTAAATTTATCCATTTA AACTAACTTTTAACTCCGTTAGATTTTAACTTTTAA ATCCCTCTTTTGATGCTGTCAAATGGCTCTGTTTCTATC CAATAAAGATAAGCCTTATTAATTGAGTGGTCGTAATGGG GTGTAAAAAAGTTTGTCTTGGCTATTGATGCGGTCTAGGGG TGTTTAAAGGGGGTTTGTG <b>TAGGAT</b> TTCATC <b>ACCC</b> CCAT AGTTGCAATATGGGGCAAATCCTATGTCA <b>AAAGGAG</b> CATG CC <b>ATGAAACACAAAAGTGGCAACGCTCTTGG</b> (638)	FA406/ FA386	FA540/ FA541
pA3-Up C4T	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATT AATTTCAATTAAAGGAGTCATCATGGCACACCATGAACAA CAACAGCAACAACAGGCTAACAGCCAACACCACCACCATC ACCATGCGCACCACCACCATTAACGTTGGCGAACACCA TCACCATAATGCGCAACAACACGCCGAACAACAAGCAGAG CAACAAGCTCAGCAACAGCAACAACAAGCACACCAA CAACAACAACAAAAAGCGCAACAACAACCAACAATAT TGATTGGGGCGTTTGTGGGGCGGCCTTAGGGCTACTCCTA GCTTTAACTCTTTCTTTAATCTAGTAAATTTATCCATTTA AACTAACTTTTAACTCCGTTAGATTTTAACTTTTAA	FA406/ FA386	FA515/ FA516

		ATCCCTCTTTTGGATGCTTGTCAAATGGCTCTGTTTCTATC CAATAAAGATAAGCCTTATTAATTGAGTGGTCGTAATGGG GTGTA AAAA ACTTTGTCTTGGCTATTTGATGCGGTCTAGGGG TGTTTAAAGGGGGTTTGTGTTAGGATTTTCATCAGTCCCAT AGTTGCAATATGGGGCAAATCCTATGTCAAAAGGAGCATG CCATGAAACACAAAAGTGGCAAACGCTCTTGG (638)		
pA3-Up G24A	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTA AATTTCAATTAAAGGAGTCATCATGGCACACCATGAACAA CAACAGCAACAACAGGCTAACAGCCAACACCACCACATC ACCATGCGCACCACCACCATTACTACGGTGGCGAACACCA TCACCATAATGCGCAACAACACGCCGAACAACAAGCAGAG CAACAAGCTCAGCAACAGCAACAACAACAAGCACACCAA CAACAACAACAAAAGCGCAACAACAACAACAACAATAT TGATTGGGGCGTTTGTGGGGGCGGCCTTAGGGCTACTCCTA GCTTTTAACTCTTTCTTTAATCTAGTAAATTTATCCATTTA AACTAACTTTTTTAACTCCGTTAGATTTTTTAACTTTTAAA ATCCCTCTTTTGGATGCTTGTCAAATGGCTCTGTTTCTATC CAATAAAGATAAGCCTTATTAATTGAGTGGTCGTAATGGG GTGTA AAAA ACTTTGTCTTGGCTATTTGATGCGGTCTAGGGG TGTTTAAAGGGGGTTTGTGTTAGGATTTTCATCAGCCTCAT AGTTGCAATATGGGACAAATCCTATGTCAAAAGGAGCATG CCATGAAACACAAAAGTGGCAAACGCTCTTGG (638)	FA406/ FA386	FA463/ FA464
pA3-Up C4T G24A	pGEM-T/ CoE11/ Amp <sup>R</sup>	GCATTATAAAATGAAATCCTAGCTAATGAGCTAGAATTTA AATTTCAATTAAAGGAGTCATCATGGCACACCATGAACAA CAACAGCAACAACAGGCTAACAGCCAACACCACCACATC ACCATGCGCACCACCACCATTACTACGGTGGCGAACACCA TCACCATAATGCGCAACAACACGCCGAACAACAAGCAGAG CAACAAGCTCAGCAACAGCAACAACAACAAGCACACCAA CAACAACAACAAAAGCGCAACAACAACAACAACAATAT TGATTGGGGCGTTTGTGGGGGCGGCCTTAGGGCTACTCCTA GCTTTTAACTCTTTCTTTAATCTAGTAAATTTATCCATTTA AACTAACTTTTTTAACTCCGTTAGATTTTTTAACTTTTAAA ATCCCTCTTTTGGATGCTTGTCAAATGGCTCTGTTTCTATC CAATAAAGATAAGCCTTATTAATTGAGTGGTCGTAATGGG GTGTA AAAA ACTTTGTCTTGGCTATTTGATGCGGTCTAGGGG TGTTTAAAGGGGGTTTGTGTTAGGATTTTCATCAGTCCCAT AGTTGCAATATGGGACAAATCCTATGTCAAAAGGAGCATG CCATGAAACACAAAAGTGGCAAACGCTCTTGG (638)	FA406/ FA386	FA461/ FA462
pA3- Down G168A pIsoA3*	pGEM-T/ CoE11/ Amp <sup>R</sup>	CACAAAAGTGGCAAACGCTCTTGGAAAACGTTCTACTTTG AGTTTGCTTTTTTGGGGCTTAAAGTGATCGTTTCTGTGAAA CGCTGATTTTTCTAAGAGCGTTCCTTAACTTTAAGCTT AGGGGTTTCCTTATGAAGCGGTTTGTCTTGACTTTTGGG GCCTTGGTCAATTTGTGTAAAACTTATTTCTTTAGGGGA TAGAAATAATTCTCCCTCTATAAACTAAATCCCTCATA GAGTGCTTTTAAAAAATACCGCTTGCTCTAGCTTTTGTATT ATTTGTCTTAAAAATCACTCCCTTAAAAACCGTATTTCATA GACTTTAGGCTGTATTTCTTTGGCGAGATCTTTAGGGCGT CTTGCAAATTAGCGTAAAGCTGTTGTAAATATTCATCTTTA GCTTTAGCGGGTAAATTTAATTGCTTGCTTGTTACGGCC TTGAAAAAATCTTTGATGTCATAAAGGCTTGCGTTAGCGT TATAGGGGCGGTTTGTGAAATCTTGTGTGGTAAATAGCG ATAAACCTCTCTGCTAGCGTCAAACACCTTTGAAGCGCTCG GGCTGAATTTACAGGGCTTGTTTCTTTTGGCACTTAAA AAGAGGCTATCGTTCTCTTTAAGTTCTTTGATTTTCGCCTTT AAAACTCTAATAGAGCGTGGCTAG (680)	FA387/ FA409	FA559/ FA560
pA3- Down G176A pIsoA3*	pGEM-T/ CoE11/ Amp <sup>R</sup>	CACAAAAGTGGCAAACGCTCTTGGAAAACGTTCTACTTTG AGTTTGCTTTTTTGGGGCTTAAAGTGATCGTTTCTGTGAAA CGCTGATTTTTCTAAGAGCGTTCCTTAACTTTAAGCTT AGGGGTTTCCTTATGAAGCGGTTTGTCTTGACTTTTGGG GCCTTGGTCAATTTGTGTAAAACTTATTTCTTTAGGGGA TAGAAATAATTCTCCCTCTATAAACTAAATCCCTCATA GAGTGCTTTTAAAAAATACCGCTTGCTCTAGCTTTTGTATT ATTTGTCTTAAAAATCACTCCCTTAAAAACCGTATTTCATA GACTTTAGGCTGTATTTCTTTGGCGAGATCTTTAGGGCGT CTTGCAAATTAGCGTAAAGCTGTTGTAAATATTCATCTTTA GCTTTAGCGGGTAAATTTAATTGCTTGCTTGTTACGGCC TTGAAAAAATCTTTGATGTCATAAAGGCTTGCGTTAGCGT TATAGGGGCGGTTTGTGAAATCTTGTGTGGTAAATAGCG ATAAACCTCTCTGCTAGCGTCAAACACCTTTGAAGCGCTCG GGCTGAATTTACAGGGCTTGTTTCTTTTGGCACTTAAA AAGAGGCTATCGTTCTCTTTAAGTTCTTTGATTTTCGCCTTT AAAACTCTAATAGAGCGTGGCTAG (680)	FA387/ FA409	FA471/ FA472
pA3- Down	pGEM-T/ CoE11/	CACAAAAGTGGCAAACGCTCTTGGAAAACGTTCTACTTTG AGTTTGCTTTTTTGGGGCTTAAAGTGATCGTTTCTGTGAAA CGCTGATTTTTCTAAGAGCGTTCCTTAACTTTAAGCTT	FA387/ FA409	- (isolated suppressor

G183A pIsoA3*	Amp <sup>R</sup>	AGGAGTTTCCTTATGAAGCGGGTTTGTCTTGACTTTTGGG GC <u>G</u> TTGGTCATTTTGTGTAAAACTTATTTCTTTAGGGGA TAGAAATAATTCTCCCTCTATAAACTAAATCCCTCATA GAGTGCTTTTAAAAAATACCGCTTGCTCTAGCTTTTGTATT ATTTGTCTTAAAAATCACTCCCTTAAAAACCGTATTCATA GACTTTAGGCTGTATTTCTTTGGCGAGATCTTTAGGGCGT CTTGCAAATTAGCGTAAAGCTGTTTGTAAATATTCATCTTTA GCTTTAGCGGGTAAATTTAATTGCCTTGCTTGTTACGGCC TTGAAAAAATTCTTTGATGTCATAAAGGCTTGCGTTAGCGT TATAGGGGCGGTTTGTGAAATCTTGTGTGTGGTAATAGCG ATAAACCTCTCTGCTAGCGTCAAACACCTTTGAAGCGCTCG GGCTGAATTTCAAGGGCTTGTTCCTTTTGGCACTTAAA AAGAGGCTATCGTTCCTTTAAGTCTTTGATTTTCGCCTTTT AAAACTCTAATAGAGCGTGGCTAG (680)	strain)
------------------	------------------	--	---------

Note: the pSP60-2, pA3-Up, pA3-UP start and pA3-Down pIsoA3\* plasmids shown in Table 8.1.2 in the Appendix.

**Table 6.4.2. Sequence of T7 RNAs used for *in vitro* studies in Paper IV.**

Lower-case letter indicated the T7 promoter used. Mutated positions are underlined. Nt\* indicates the length in nucleotides of the RNA (excluding the T7 promoter sequence). The wild type *aapA3* full-length and truncated isoforms as well as the IsoA3 RNA used in this study were already described in the Table 6.3.2.

T7 RNA	Sequence (5'→3') (nt*)	Primer pair	Template
<i>aapA3</i> -5'SL WT	gaaattaatacgactcactataggACGCCCCATAGTTGC AATATGGGGCAAATCCTATGTCAAAAGGA GCA (47)	FA170/FA624	gDNA <i>H. pylori</i> 26695 WT
<i>aapA3</i> -5'SL G24A	gaaattaatacgactcactataggACGCCCCATAGTTGC AATATGGG <u>A</u> CAAATCCTATGTCAAAAGGA GCA (47)	FA170/FA624	gDNA <i>H. pylori</i> 26695 <i>aapA3</i> G24A/pIsoA3*

**Table 6.4.3. Sequence of chemically synthesized RNAs used in Paper IV.**

Mutated positions are underlined.

RNA	Sequence (5'→3') (nt*)
<i>aapA3</i> -5'SL <sub>25</sub> WT	ACGCCCCATAGTTGCAATATGGGGC (25)
<i>aapA3</i> -5'SL <sub>25</sub> G24A	ACGCCCCATAGTTGCAATATGGG <u>A</u> C (25)
<i>aapA3</i> -3'SL <sub>24</sub> WT	CCCTTAAGCTTTTAAGCTTAGGGG (24)
<i>aapA3</i> -3'SL <sub>24</sub> G183A	CCCTTAAGCTTTTAAGCTTAGG <u>A</u> G (24)



## CHAPTER 7. References

- Aiba, H. (2007). Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.* *10*, 134–139.
- Al-Hashimi, H.M., and Walter, N.G. (2008). RNA dynamics: it is about time. *Curr. Opin. Struct. Biol.* *18*, 321–329.
- Alm, R.A., Ling, L.-S.L., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., et al. (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* *397*, 176–180.
- Altuvia, S., Storz, G., and Papenfort, K. (2018). Cross-Regulation between Bacteria and Phages at a Posttranscriptional Level. *Microbiol. Spectr.* *6*.
- Alvarez-Martinez, C.E., and Christie, P.J. (2009). Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiol. Mol. Biol. Rev.* *73*, 775–808.
- Andrade, J.M., Pobre, V., Matos, A.M., and Arraiano, C.M. (2012). The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* *18*, 844–855.
- Andreeva, I., Belardinelli, R., and Rodnina, M. V. (2018). Translation initiation in bacterial polysomes through ribosome loading on a standby site on a highly translated mRNA. *Proc. Natl. Acad. Sci.* *115*, 4411–4416.
- Argaman, L., Hershsberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H., and Altuvia, S. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* *11*, 941–950.
- Arnion, H., Korkut, D.N., Masachis Gelo, S., Chabas, S., Reignier, J., Iost, I., and Darfeuille, F. (2017). Mechanistic insights into type I toxin antitoxin systems in *Helicobacter pylori*: the importance of mRNA folding in controlling toxin expression. *Nucleic Acids Res.* gkw1343.
- Atmakuri, K., Cascales, E., and Christie, P.J. (2004). Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. *Mol. Microbiol.* *54*, 1199–1211.
- Awano, N., Rajagopal, V., Arbing, M., Patel, S., Hunt, J., Inouye, M., and Phadtare, S. (2010). *Escherichia coli* RNase R Has Dual Activities, Helicase and RNase. *J. Bacteriol.* *192*, 1344–1352.
- Babu, M.M., Lang, B., and Aravind, L. (2009). Methods to Reconstruct and Compare Transcriptional Regulatory Networks. In *Methods in Molecular Biology* (Clifton, N.J.), pp. 163–180.
- Bakshi, S., Siryaporn, A., Goulian, M., and Weisshaar, J.C. (2012). Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Mol. Microbiol.* *85*, 21–38.
- Balaban, N., and Novick, R.P. (1995). Translation of RNAIII, the *Staphylococcus aureus* agr regulatory RNA molecule, can be activated by a 3'-end deletion. *FEMS Microbiol. Lett.* *133*, 155–161.
- Bandyra, K.J., Sinha, D., Syrjanen, J., Luisi, B.F., and De Lay, N.R. (2016). The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes. *RNA* *22*, 360–372.
- Bartel, P.L., Roecklein, J.A., SenGupta, D., and Fields, S. (1996). A protein linkage map of *Escherichia coli* bacteriophage T7. *Nat. Genet.* *12*, 72–77.
- Bauer, S., Endres, M., Lange, M., Schmidt, T., Schumbrutzki, C., Sickmann, A., and Beier, D. (2013). Novel function assignment to a member of the essential HP1043 response regulator family of epsilon-proteobacteria. *Microbiology* *159*, 880–889.
- Beier, D., and Frank, R. (2000). Molecular characterization of two-component systems of *Helicobacter pylori*. *J. Bacteriol.* *182*, 2068–2076.
- Bentele, K., Saffert, P., Rauscher, R., Ignatova, Z., and Bluthgen, N. (2014). Efficient translation initiation dictates codon usage at gene start. *Mol. Syst. Biol.* *9*, 675–675.
- Berghoff, B.A., Hoekzema, M., Aulbach, L., and Wagner, E.G.H. (2017). Two regulatory RNA elements affect TisB-dependent depolarization and persister formation. *Mol. Microbiol.* *103*, 1020–1033.
- Bhattacharyya, S., Jacobs, W.M., Adkar, B. V., Yan, J., Zhang, W., and Shakhnovich, E.I. (2018). Accessibility of the Shine-Dalgarno Sequence Dictates N-Terminal Codon Bias in *E. coli*. *Mol. Cell* *70*, 894–905.e5.
- Bik, E.M., Eckburg, P.B., Gill, S.R., Nelson, K.E., Purdom, E.A., Francois, F., Perez-Perez, G., Blaser, M.J., and

- Relman, D.A. (2006). Molecular analysis of the bacterial microbiota in the human stomach. *Proc. Natl. Acad. Sci.* *103*, 732–737.
- Binns, N., and Masters, M. (2002). Expression of the *Escherichia coli* *pcnB* gene is translationally limited using an inefficient start codon: a second chromosomal example of translation initiated at AUU. *Mol. Microbiol.* *44*, 1287–1298.
- Bischler, T., Hsieh, P., Resch, M., Liu, Q., Tan, H.S., Foley, P.L., Hartleib, A., Sharma, C.M., and Belasco, J.G. (2017). Identification of the RNA Pyrophosphohydrolase RppH of *Helicobacter pylori* and Global Analysis of Its RNA Targets. *J. Biol. Chem.* *292*, 1934–1950.
- Blaszczyk, J., Gan, J., Tropea, J.E., Court, D.L., Waugh, D.S., and Ji, X. (2004). Nucleolytic Assembly of Ribonuclease III with Double-Stranded RNA. *Structure* *12*, 457–466.
- Boneca, I.G., Ecobichon, C., Chaput, C., Mathieu, A., Guadagnini, S., Prévost, M.-C., Colland, F., Labigne, A., and Reuse, H. de (2008). Development of Inducible Systems To Engineer Conditional Mutants of Essential Genes of *Helicobacter pylori*. *Appl. Environ. Microbiol.* *74*, 2095–2102.
- Bonis, M., Ecobichon, C., Guadagnini, S., Prévost, M.-C., and Boneca, I.G. (2010). A M23B family metalloprotease of *Helicobacter pylori* required for cell shape, pole formation and virulence. *Mol. Microbiol.* *78*, 809–819.
- Boschirolì, M.L., Ouahrani-Bettache, S., Foulongne, V., Michaux-Charachon, S., Bourg, G., Allardet-Servent, A., Cazevielle, C., Lavigne, J.-P., Liautard, J.P., Ramuz, M., et al. (2002). Type IV secretion and *Brucella* virulence. *Vet. Microbiol.* *90*, 341–348.
- Bouvet, P., and Belasco, J.G. (1992). Control of RNase E-mediated RNA degradation by 5'-terminal base pairing in *E. coli*. *Nature* *360*, 488–491.
- Brantl, S., and Jahn, N. (2015). SRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.* *39*, 413–427.
- Brielle, R., Pinel-Marie, M.-L., and Felden, B. (2016). Linking bacterial type I toxins with their actions. *Curr. Opin. Microbiol.* *30*, 114–121.
- Brombach, M., and Pon, C.L. (1987). The unusual translational initiation codon AUU limits the expression of the *infC* (initiation factor IF3) gene of *Escherichia coli*. *Mol. Gen. Genet.* *208*, 94–100.
- Bumann, D., Habibi, H., Kan, B., Schmid, M., Goosmann, C., Brinkmann, V., Meyer, T.F., and Jungblut, P.R. (2004). Lack of Stage-Specific Proteins in Cocoid *Helicobacter pylori* Cells. *Infect. Immun.* *72*, 6738–6742.
- Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rösch, P. (2010). A NusE:NusG complex links transcription and translation. *Science* *328*, 501–504.
- Bury-Moné, S., Skouloubris, S., Labigne, A., and De Reuse, H. (2001). The *Helicobacter pylori* UreI protein: role in adaptation to acidity and identification of residues essential for its activity and for acid activation. *Mol. Microbiol.* *42*, 1021–1034.
- Butler, J.S., Springer, M., and Grunberg-Manago, M. (1987). AUU-to-AUG mutation in the initiator codon of the translation initiation factor IF3 abolishes translational autocontrol of its own gene (*infC*) in vivo. *Proc. Natl. Acad. Sci. U. S. A.* *84*, 4022–4025.
- Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G., and Luisi, B.F. (2005). Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* *437*, 1187–1191.
- Cárdenas-Mondragón, M.G., Ares, M.A., Panunzi, L.G., Pacheco, S., Camorlinga-Ponce, M., Girón, J.A., Torres, J., and De la Cruz, M.A. (2016). Transcriptional Profiling of Type II Toxin-Antitoxin Genes of *Helicobacter pylori* under Different Environmental Conditions: Identification of HP0967-HP0968 System. *Front. Microbiol.* *7*, 1872.
- Carpousis, A.J. (2007). The RNA Degradosome of *Escherichia coli*: An mRNA-Degrading Machine Assembled on RNase E. *Annu. Rev. Microbiol.* *61*, 71–87.
- Carpousis, A.J., Van Houwe, G., Ehretsmann, C., and Krisch, H.M. (1994). Copurification of *E. coli* RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* *76*, 889–900.
- CARRON, M., TRAN, V., SUGAWA, C., and COTICCHIA, J. (2006). Identification of *Helicobacter pylori* Biofilms in Human Gastric Mucosa. *J. Gastrointest. Surg.* *10*, 712–717.
- Carter-Muenchau, P., and Wolf, R.E. (1989). Growth-rate-dependent regulation of 6-phosphogluconate dehydrogenase level mediated by an anti-Shine-Dalgarno sequence located within the *Escherichia coli* *gnd*

- structural gene. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 1138–1142.
- Castillo, A.R., Arevalo, S.S., Woodruff, A.J., and Ottemann, K.M. (2007). Experimental analysis of *Helicobacter pylori* transcriptional terminators suggests this microbe uses both intrinsic and factor-dependent termination. *Mol. Microbiol.* **67**, 155–170.
- Causton, H., Py, B., McLaren, R.S., and Higgins, C.F. (1994). mRNA degradation in *Escherichia coli*: a novel factor which impedes the exoribonucleolytic activity of PNPase at stem-loop structures. *Mol. Microbiol.* **14**, 731–741.
- Celli, J.P., Turner, B.S., Afdhal, N.H., Keates, S., Ghiran, I., Kelly, C.P., Ewoldt, R.H., McKinley, G.H., So, P., Erramilli, S., et al. (2009). *Helicobacter pylori* moves through mucus by reducing mucin viscoelasticity. *Proc. Natl. Acad. Sci.* **106**, 14321–14326.
- Cellini, L. (2014). *Helicobacter pylori*: A chameleon-like approach to life. *World J. Gastroenterol.* **20**, 5575.
- Chalker, A.F., Minehart, H.W., Hughes, N.J., Koretke, K.K., Lonetto, M.A., Brinkman, K.K., Warren, P. V., Lupas, A., Stanhope, M.J., Brown, J.R., et al. (2001). Systematic identification of selective essential genes in *Helicobacter pylori* by genome prioritization and allelic replacement mutagenesis. *J. Bacteriol.* **183**, 1259–1268.
- Chang, Y.-W., Shaffer, C.L., Rettberg, L.A., Ghosal, D., and Jensen, G.J. (2018). In Vivo Structures of the *Helicobacter pylori* cag Type IV Secretion System. *Cell Rep.* **23**, 673–681.
- Chao, Y., and Vogel, J. (2016). A 3' UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. *Mol. Cell* 1–12.
- Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., and Vogel, J. (2012). An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J.* **31**, 4005–4019.
- Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., Śmiga, M., Papenfort, K., Reinhardt, R., Wieden, H.-J., et al. (2017). In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol. Cell* **65**, 39–51.
- Chen, H., Bjerknes, M., Kumar, R., and Jay, E. (1994). Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res.* **22**, 4953–4957.
- Chen, L.H., Emory, S.A., Bricker, A.L., Bouvet, P., and Belasco, J.G. (1991). Structure and function of a bacterial mRNA stabilizer: analysis of the 5' untranslated region of *ompA* mRNA. *J. Bacteriol.* **173**, 4578–4586.
- Chen, M., Andersen, L.P., Zhai, L., and Kharazmi, A. (1999). Characterization of the respiratory chain of *Helicobacter pylori*. *FEMS Immunol. Med. Microbiol.* **24**, 169–174.
- Chen, Z., Itzek, A., Malke, H., Ferretti, J.J., and Kreth, J. (2013). Multiple roles of RNase Y in *Streptococcus pyogenes* mRNA processing and degradation. *J. Bacteriol.* **195**, 2585–2594.
- Cheng, Z.-F., and Deutscher, M.P. (2002). Purification and Characterization of the *Escherichia coli* Exoribonuclease RNase R. *J. Biol. Chem.* **277**, 21624–21629.
- Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005). BIOGENESIS, ARCHITECTURE, AND FUNCTION OF BACTERIAL TYPE IV SECRETION SYSTEMS. *Annu. Rev. Microbiol.* **59**, 451–485.
- Contreras, M., Thiberge, J.-M., Mandrand-Berthelot, M.-A., and Labigne, A. (2003). Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of *Helicobacter pylori*. *Mol. Microbiol.* **49**, 947–963.
- Corbin, R.W., Paliy, O., Yang, F., Shabanowitz, J., Platt, M., Lyons, C.E., Root, K., McAuliffe, J., Jordan, M.I., Kustu, S., et al. (2003). Toward a protein profile of *Escherichia coli*: Comparison to its transcription profile. *Proc. Natl. Acad. Sci.* **100**, 9232–9237.
- Covacci, A., Censini, S., Bugnoli, M., Petracca, R., Burrone, D., Macchia, G., Massone, A., Papini, E., Xiang, Z., and Figura, N. (1993). Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5791–5795.
- Croxen, M.A., Sisson, G., Melano, R., and Hoffman, P.S. (2006). The *Helicobacter pylori* chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa. *J. Bacteriol.* **188**, 2656–2665.
- Dailidienė, D., Dailidienė, G., Kersulyte, D., and Berg, D.E. (2006). Contraselectable Streptomycin Susceptibility Determinant for Genetic Manipulation and Analysis of *Helicobacter pylori*. **72**, 5908–5914.
- Dam Mikkelsen, N., and Gerdes, K. (1997). *Sok* antisense RNA from plasmid R1 is functionally inactivated by

- RNase E and polyadenylated by poly(A) polymerase I. *Mol. Microbiol.* 26, 311–320.
- Danielli, A., Amore, G., and Scarlato, V. (2010). Built Shallow to Maintain Homeostasis and Persistent Infection: Insight into the Transcriptional Regulatory Network of the Gastric Human Pathogen *Helicobacter pylori*. *PLoS Pathog.* 6, e1000938.
- Daou-Chabo, R., and Condon, C. (2009). RNase J1 endonuclease activity as a probe of RNA secondary structure. *RNA* 15, 1417–1425.
- Dar, D., and Sorek, R. (2018). Extensive reshaping of bacterial operons by programmed mRNA decay. *PLOS Genet.* 14, e1007354.
- Dar, D., Shamir, M., Mellin, J.R., Koutero, M., Stern-Ginossar, N., Cossart, P., and Sorek, R. (2016). Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* (80-. ). 352.
- Darfeuille, F., Unoson, C., Vogel, J., and Wagner, E.G.H. (2007). An Antisense RNA Inhibits Translation by Competing with Standby Ribosomes. *Mol. Cell* 26, 381–392.
- Darty, K., Denise, A., and Ponty, Y. (2009). VARNAs: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25, 1974–1975.
- Das, H.K., Goldstein, A., and Lowney, L.I. (1967). Attachment of ribosomes to nascent messenger RNA in *Escherichia coli*. *J. Mol. Biol.* 24, 231–245.
- Deana, A., and Belasco, J.G. (2005). Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.* 19, 2526–2533.
- Debowski, A.W., Walton, S.M., Chua, E.-G., Tay, A.C.-Y., Liao, T., Lamichhane, B., Himbeck, R., Stubbs, K.A., Marshall, B.J., Fulurija, A., et al. (2017). *Helicobacter pylori* gene silencing in vivo demonstrates urease is essential for chronic infection. *PLoS Pathog.* 13, e1006464.
- Delany, I., Spohn, G., Rappuoli, R., and Scarlato, V. (2002). Growth phase-dependent regulation of target gene promoters for binding of the essential orphan response regulator HP1043 of *Helicobacter pylori*. *J. Bacteriol.* 184, 4800–4810.
- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607.
- Deutscher, M.P., Marshall, G.T., and Cudny, H. (1988). RNase PH: an *Escherichia coli* phosphate-dependent nuclease distinct from polynucleotide phosphorylase. *Proc. Natl. Acad. Sci. U. S. A.* 85, 4710–4714.
- Dombroski, A.J., and Platt, T. (1988). Structure of rho factor: an RNA-binding domain and a separate region with strong similarity to proven ATP-binding domains. *Proc. Natl. Acad. Sci. U. S. A.* 85, 2538–2542.
- Dorer, M.S., Sessler, T.H., and Salama, N.R. (2011). Recombination and DNA Repair in *Helicobacter pylori*. *Annu. Rev. Microbiol.* 65, 329–348.
- Dorléans, A., Li de la Sierra-Gallay, I., Piton, J., Zig, L., Gilet, L., Putzer, H., and Condon, C. (2011). Molecular Basis for the Recognition and Cleavage of RNA by the Bifunctional 5′–3′ Exo/Endoribonuclease RNase J. *Structure* 19, 1252–1261.
- Dörr, T., Vulić, M., and Lewis, K. (2010). Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biol.* 8, e1000317.
- Dressaire, C., Laurent, B., Loubière, P., Besse, P., and Coccagn-Bousquet, M. (2010). Linear covariance models to examine the determinants of protein levels in *Lactococcus lactis*. *Mol. Biosyst.* 6, 1255.
- Ducongé, F., Di Primo, C., and Toulme, J.J. (2000). Is a Closing “GA Pair” a Rule for Stable Loop-Loop RNA Complexes? *J. Biol. Chem.* 275, 21287–21294.
- Duhring, U., Axmann, I.M., Hess, W.R., and Wilde, A. (2006). An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc. Natl. Acad. Sci.* 103, 7054–7058.
- Dunn, J.J., and Studier, F.W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* 166, 477–535.
- Dunn, J.J., Buzash-Pollert, E., and Studier, F.W. (1978). Mutations of bacteriophage T7 that affect initiation of synthesis of the gene 0.3 protein. *Proc. Natl. Acad. Sci. U. S. A.* 75, 2741–2745.
- Durand, S., Gilet, L., and Condon, C. (2012a). The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet.* 8, e1003181.

- Durand, S., Jahn, N., Condon, C., and Brantl, S. (2012b). Type I toxin-antitoxin systems in *Bacillus subtilis*. *RNA Biol.* 9, 1491–1497.
- Durand, S., Jahn, N., Condon, C., and Brantl, S. (2012c). Type I toxin-antitoxin systems in *Bacillus subtilis*. *RNA Biol.* 9, 1491–1497.
- Duss, O., Michel, E., Diarra dit Konté, N., Schubert, M., and Allain, F.H.-T. (2014). Molecular basis for the wide range of affinity found in Csr/Rsm protein–RNA recognition. *Nucleic Acids Res.* 42, 5332–5346.
- Dutta, T., and Deutscher, M.P. (2009). Catalytic properties of RNase BN/RNase Z from *Escherichia coli*: RNase BN is both an exo- and endoribonuclease. *J. Biol. Chem.* 284, 15425–15431.
- Duval, M., Korepanov, A., Fuchsbaauer, O., Fechter, P., Haller, A., Fabbretti, A., Choulier, L., Micura, R., Klaholz, B.P., Romby, P., et al. (2013). *Escherichia coli* Ribosomal Protein S1 Unfolds Structured mRNAs Onto the Ribosome for Active Translation Initiation. *PLoS Biol.* 11, 12–14.
- Duval, M., Simonetti, A., Caldelari, I., and Marzi, S. (2015). Multiple ways to regulate translation initiation in bacteria: Mechanisms, regulatory circuits, dynamics. *Biochimie* 114, 18–29.
- Dwivedi, G.R., Sharma, E., and Rao, D.N. (2013). *Helicobacter pylori* DprA alleviates restriction barrier for incoming DNA. *Nucleic Acids Res.* 41, 3274–3288.
- Ehretsmann, C.P., Carpousis, A.J., and Krisch, H.M. (1992). Specificity of *Escherichia coli* endoribonuclease RNase E: in vivo and in vitro analysis of mutants in a bacteriophage T4 mRNA processing site. *Genes Dev.* 6, 149–159.
- Emory, S.A., Bouvet, P., and Belasco, J.G. (1992). A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev.* 6, 135–148.
- Ermak, T.H., Giannasca, P.J., Nichols, R., Myers, G.A., Nedrud, J., Weltzin, R., Lee, C.K., Kleanthous, H., and Monath, T.P. (1998). Immunization of mice with urease vaccine affords protection against *Helicobacter pylori* infection in the absence of antibodies and is mediated by MHC class II-restricted responses. *J. Exp. Med.* 188, 2277–2288.
- Espah Borujeni, A., Channarasappa, A.S., and Salis, H.M. (2014). Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* 42, 2646–2659.
- Even, S., Pellegrini, O., Zig, L., Labas, V., Vinh, J., Bréchemmier-Baey, D., and Putzer, H. (2005). Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E. *Nucleic Acids Res.* 33, 2141–2152.
- Falush, D., Wirth, T., Linz, B., Pritchard, J.K., Stephens, M., Kidd, M., Blaser, M.J., Graham, D.Y., Vacher, S., Perez-Perez, G.I., et al. (2003). Traces of Human Migrations in *Helicobacter pylori* Populations. *Science* (80-. ). 299, 1582–1585.
- Farnbacher, M., Jahns, T., Willrodt, D., Daniel, R., Haas, R., Goesmann, A., Kurtz, S., and Rieder, G. (2010). Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. *BMC Genomics* 11, 335.
- Feng, L., Rutherford, S.T., Papenfort, K., Bagert, J.D., van Kessel, J.C., Tirrell, D.A., Wingreen, N.S., and Bassler, B.L. (2015). A Qrr Noncoding RNA Deploys Four Different Regulatory Mechanisms to Optimize Quorum-Sensing Dynamics. *Cell* 160, 228–240.
- Fields, J.A., Li, J., Gulbranson, C.J., Hendrixson, D.R., and Thompson, S.A. (2016). *Campylobacter jejuni* CsrA Regulates Metabolic and Virulence Associated Proteins and Is Necessary for Mouse Colonization. *PLoS One* 11, e0156932.
- Fischer, W., Hofreuter, D., and Haas, R. (2001). *Natural Transformation, Recombination, and Repair* (ASM Press).
- Fox, J.G. (2002). The non-H pylori helicobacters: their expanding role in gastrointestinal and systemic diseases. *Gut* 50, 273–283.
- Foyes, S., Dorrell, N., Ward, S.J., Stabler, R.A., McColm, A.A., Rycroft, A.N., and Wren, B.W. (2000). *Helicobacter pylori* possesses two CheY response regulators and a histidine kinase sensor, CheA, which are essential for chemotaxis and colonization of the gastric mucosa. *Infect. Immun.* 68, 2016–2023.
- Fozo, E.M., Hemm, M.R., and Storz, G. (2008a). Small Toxic Proteins and the Antisense RNAs That Repress Them. 72, 579–589.
- Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Kathy, S., Jones, K.L., Ocampo, A., Rudd, K.E., and Storz, G.



- (2008b). Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *70*, 1076–1093.
- Franch, T., Gulyaev, A.P., and Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: Processing at the hok mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding. *J. Mol. Biol.* *273*, 38–51.
- Franch, T., Petersen, M., Wagner, E.G.H., Jacobsen, J.P., and Gerdes, K. (1999). Antisense RNA regulation in prokaryotes: Rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *J. Mol. Biol.* *294*, 1115–1125.
- Fröhlich, K.S., Papenfort, K., Fekete, A., and Vogel, J. (2013). A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J.* *32*, 2963–2979.
- Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2006). Structural Insight into the Mechanism of Double-Stranded RNA Processing by Ribonuclease III. *Cell* *124*, 355–366.
- García, A., Salas-Jara, M.J., Herrera, C., and González, C. (2014). Biofilm and *Helicobacter pylori*: from environment to human host. *World J. Gastroenterol.* *20*, 5632–5638.
- Ge, Z. (2002). Potential of fumarate reductase as a novel therapeutic target in *Helicobacter pylori* infection. *Expert Opin. Ther. Targets* *6*, 135–146.
- Ge, Z., Feng, Y., Dangler, C.A., Xu, S., Taylor, N.S., and Fox, J.G. (2000). Fumarate reductase is essential for *Helicobacter pylori* colonization of the mouse stomach. *Microb. Pathog.* *29*, 279–287.
- Gerdes, K. (2016). Hypothesis: type I toxin-antitoxin genes enter the persistence field—a feedback mechanism explaining membrane homeostasis. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *371*, 20160189.
- Gerdes, K., and Wagner, E.G.H. (2007). RNA antitoxins. *Curr. Opin. Microbiol.* *10*, 117–124.
- Gerdes, K., Bech, F.W., Jorgensen, S.T., Lobner-olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlstrom, O., Molin, S., and Von, K. (1986a). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. *5*, 2023–2029.
- Gerdes, K., Rasmussen, P.B., and Molin, S. (1986b). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. U. S. A.* *83*, 3116–3120.
- Giangrossi, M., Prosseda, G., Tran, C.N., Brandi, A., Colonna, B., and Falconi, M. (2010). A novel antisense RNA regulates at transcriptional level the virulence gene icsA of *Shigella flexneri*. *Nucleic Acids Res.* *38*, 3362–3375.
- Giangrossi, M., Giuliadori, A.M., Tran, C.N., Amici, A., Marchini, C., and Falconi, M. (2017). VirF Relieves the Transcriptional Attenuation of the Virulence Gene icsA of *Shigella flexneri* Affecting the icsA mRNA–RnaG Complex Formation. *Front. Microbiol.* *8*, 650.
- Gimpel, M., and Brantl, S. (2017). Dual-function small regulatory RNAs in bacteria. *Mol. Microbiol.* *103*, 387–397.
- Gimpel, M., Heidrich, N., Mäder, U., Krügel, H., and Brantl, S. (2010). A dual-function sRNA from *B. subtilis*: SR1 acts as a peptide encoding mRNA on the gapA operon. *Mol. Microbiol.* *76*, 990–1009.
- Göbl, C., Kosol, S., Stockner, T., Rückert, H.M., and Zangger, K. (2010). Solution Structure and Membrane Binding of the Toxin Fst of the *par* Addiction Module. *Biochemistry* *49*, 6567–6575.
- Gold, L. (1988). Posttranscriptional Regulatory Mechanisms in *Escherichia coli*. *Annu. Rev. Biochem.* *57*, 199–233.
- Goormaghtigh, F., Fraikin, N., Putrinš, M., Hallaert, T., Hauryliuk, V., Garcia-Pino, A., Sjödin, A., Kasvandik, S., Udekwu, K., Tenson, T., et al. (2018). Reassessing the Role of Type II Toxin-Antitoxin Systems in Formation of *Escherichia coli* Type II Persister Cells. *MBio* *9*.
- Gottesman, S. (2005). Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet.* *21*, 399–404.
- Grage, S.L., Afonin, S., Kara, S., Butth, G., and Ulrich, A.S. (2016). Membrane Thinning and Thickening Induced by Membrane-Active Amphipathic Peptides. *Front. Cell Dev. Biol.* *4*, 65.
- Graham, J.E., Peek, R.M., Krishna, U., and Cover, T.L. (2002). Global analysis of *Helicobacter pylori* gene expression in human gastric mucosa. *Gastroenterology* *123*, 1637–1648.
- Greenfield, T.J., and Weaver, K.E. (2000). Antisense RNA regulation of the pAD1 par post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Mol. Microbiol.* *37*, 661–670.

- Greenfield, T.J., Ehli, E., Kirshenmann, T., Franch, T., Gerdes, K., and Weaver, K.E. (2000). The antisense RNA of the *par* locus of *pAD1* regulates the expression of a 33-amino-acid toxic peptide by an unusual mechanism. *Mol. Microbiol.* 37, 652–660.
- Greenfield, T.J., Franch, T., Gerdes, K., and Weaver, K.E. (2001). Antisense RNA regulation of the *par* post-segregational killing system: structural analysis and mechanism of binding of the antisense RNA, *RNAI* and its target, *RNAI*. *Mol. Microbiol.* 42, 527–537.
- Gregory, S.T., Cate, J.H., and Dahlberg, A.E. (2001). Streptomycin-resistant and streptomycin-dependent mutants of the extreme thermophile *Thermus thermophilus*. *J. Mol. Biol.* 309, 333–338.
- Groeneveld, H., Thimon, K., and van Duin, J. (1995). Translational control of maturation-protein synthesis in phage MS2: a role for the kinetics of RNA folding? *RNA* 1, 79–88.
- Grohmann, E., Muth, G., and Espinosa, M. (2003). Conjugative plasmid transfer in gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* 67, 277–301, table of contents.
- Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., and Hofacker, I.L. (2008). The Vienna RNA websuite. *Nucleic Acids Res.* 36, W70–4.
- Guo, M.S., Updegrove, T.B., Gogol, E.B., Shabalina, S.A., Gross, C.A., and Storz, G. (2014). MicL, a new  $\sigma^E$ -dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev.* 28, 1620–1634.
- Gurnev, P.A., Ortenberg, R., Dörr, T., Lewis, K., and Bezrukov, S.M. (2012). Persister-promoting bacterial toxin TisB produces anion-selective pores in planar lipid bilayers. *FEBS Lett.* 586, 2529–2534.
- Han, K.-D., Matsuura, A., Ahn, H.-C., Kwon, A.-R., Min, Y.-H., Park, H.-J., Won, H.-S., Park, S.-J., Kim, D.-Y., and Lee, B.-J. (2011). Functional Identification of Toxin-Antitoxin Molecules from *Helicobacter pylori* 26695 and Structural Elucidation of the Molecular Interactions. *J. Biol. Chem.* 286, 4842–4853.
- Han, K.-D., Ahn, D.-H., Lee, S.-A., Min, Y.-H., Kwon, A.-R., Ahn, H.-C., and Lee, B.-J. (2013). Identification of Chromosomal HP0892-HP0893 Toxin-Antitoxin Proteins in *Helicobacter pylori* and Structural Elucidation of Their Protein-Protein Interaction. *J. Biol. Chem.* 288, 6004–6013.
- Han, K., Tjaden, B., and Lory, S. (2016). GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by in vivo proximity ligation. *Nat. Microbiol.* 2, 16239.
- Harms, A., Fino, C., Sørensen, M.A., Semsey, S., and Gerdes, K. (2017). Prophages and Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant Persister Cells. *MBio* 8.
- den Hartigh, A.B., Rolan, H.G., de Jong, M.F., and Tsolis, R.M. (2008). VirB3 to VirB6 and VirB8 to VirB11, but Not VirB7, Are Essential for Mediating Persistence of *Brucella* in the Reticuloendothelial System. *J. Bacteriol.* 190, 4427–4436.
- Hartz, D., McPheeters, D.S., and Gold, L. (1989). Selection of the initiator tRNA by *Escherichia coli* initiation factors. *Genes Dev.* 3, 1899–1912.
- Hathroubi, S., Servetas, S.L., Windham, I., Merrell, D.S., and Ottemann, K.M. (2018). *Helicobacter pylori* Biofilm Formation and Its Potential Role in Pathogenesis. *Microbiol. Mol. Biol. Rev.* 82, e00001-18.
- Hecht, A., Glasgow, J., Jaschke, P.R., Bawazer, L.A., Munson, M.S., Cochran, J.R., Endy, D., and Salit, M. (2017). Measurements of translation initiation from all 64 codons in *E. coli*. *Nucleic Acids Res.* 45, 3615–3626.
- Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., and Block, S.M. (2006). Sequence-Resolved Detection of Pausing by Single RNA Polymerase Molecules. *Cell* 125, 1083–1094.
- Hofreuter, D., Odenbreit, S., and Haas, R. (2001). Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. *Mol. Microbiol.* 41, 379–391.
- Hossain, S.T., Malhotra, A., and Deutscher, M.P. (2016). How RNase R Degrades Structured RNA. *J. Biol. Chem.* 291, 7877–7887.
- Hua, J., and Ho, B. (1996). Is the coccoid form of *Helicobacter pylori* viable? *Microbios* 87, 103–112.
- Humbert, O., Dorer, M.S., and Salama, N.R. (2011). Characterization of *Helicobacter pylori* factors that control transformation frequency and integration length during inter-strain DNA recombination. *Mol. Microbiol.* 79, 387–401.
- Im, H., Jang, S.-B., Pathak, C., Yang, Y.-J., Yoon, H.-J., Yu, T.-K., Suh, J.-Y., and Lee, B.-J. (2014). Crystal structure of toxin HP0892 from *Helicobacter pylori* with two Zn(II) at 1.8 Å resolution. *Protein Sci.* 23, 819–832.

- Incarnato, D., Morandi, E., Anselmi, F., Simon, L.M., Basile, G., and Oliviero, S. (2017). In vivo probing of nascent RNA structures reveals principles of cotranscriptional folding. *Nucleic Acids Res.* *45*, 9716–9725.
- Inouye, M., Arnheim, N., and Sternglanz, R. (1973). Bacteriophage T7 lysozyme is an N-acetylmuramyl-L-alanine amidase. *J. Biol. Chem.* *248*, 7247–7252.
- Iost, I., and Dreyfus, M. (1995). The stability of *Escherichia coli* lacZ mRNA depends upon the simultaneity of its synthesis and translation. *EMBO J.* *14*, 3252–3261.
- Jahn, N., and Brantl, S. (2013a). One antitoxin-two functions: SR4 controls toxin mRNA decay and translation. *Nucleic Acids Res.* *41*, 9870–9880.
- Jahn, N., and Brantl, S. (2013b). One antitoxin — two functions: SR4 controls toxin mRNA decay and translation. *41*, 9870–9880.
- Jahn, N., Preis, H., Wiedemann, C., and Brantl, S. (2012a). BsrG/SR4 from *Bacillus subtilis*- the first temperature-dependent type I toxin-antitoxin system. *Mol. Microbiol.* *83*, 579–598.
- Jahn, N., Preis, H., Wiedemann, C., Brantl, S., Jena, F., and Fakultät, B. (2012b). BsrG / SR4 from *Bacillus subtilis* – the first temperature-dependent type I toxin – antitoxin system. *83*, 579–598.
- Jimenez-Pearson, M.-A., Delany, I., Scarlato, V., and Beier, D. (2005). Phosphate flow in the chemotactic response system of *Helicobacter pylori*. *Microbiology* *151*, 3299–3311.
- Jonathan P. Celli, †, Bradley S. Turner, †,‡, Nezam H. Afdhal, ‡, Randy H. Ewoldt, §, Gareth H. McKinley, §, Rama Bansil, \*,† and, and Shyamsunder Erramilli\*, † (2007). Rheology of Gastric Mucin Exhibits a pH-Dependent Sol–Gel Transition.
- Kaito, C., Saito, Y., Nagano, G., Ikuo, M., Omae, Y., Hanada, Y., Han, X., Kuwahara-Arai, K., Hishinuma, T., Baba, T., et al. (2011). Transcription and Translation Products of the Cytolysin Gene psm-mec on the Mobile Genetic Element SCCmec Regulate *Staphylococcus aureus* Virulence. *PLoS Pathog.* *7*, e1001267.
- Kang, S.-Y., Heo, K.T., and Hong, Y.-S. (2018). Optimization of Artificial Curcumin Biosynthesis in *E. coli* by Randomized 5'-UTR Sequences To Control the Multienzyme Pathway. *ACS Synth. Biol.* *accsynbio.8b00198*.
- Kawamoto, H., Koide, Y., Morita, T., and Aiba, H. (2006). Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol. Microbiol.* *61*, 1013–1022.
- Kawano, M. (2012). Divergently overlapping *cis* -encoded antisense RNA regulating toxin-antitoxin systems from *E. coli*. *RNA Biol.* *9*, 1520–1527.
- Kawano, M., Oshima, T., Kasai, H., and Mori, H. (2002). Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a *cis*-encoded small antisense RNA in *Escherichia coli*. *Mol. Microbiol.* *45*, 333–349.
- Kawano, M., Reynolds, A.A., Miranda-Rios, J., and Storz, G. (2005a). Detection of 5'- and 3'-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* *33*, 1040–1050.
- Kawano, M., Reynolds, A.A., Miranda-Rios, J., and Storz, G. (2005b). Detection of 5'- and 3'-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* *33*, 1040–1050.
- Kawano, M., Aravind, L., and Storz, G. (2007a). An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol. Microbiol.* *64*, 738–754.
- Kawano, M., Aravind, L., and Storz, G. (2007b). An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *64*, 738–754.
- Kennell, D., and Riezman, H. (1977). Transcription and translation initiation frequencies of the *Escherichia coli* lac operon. *J. Mol. Biol.* *114*, 1–21.
- Keshavarz, T., Walker, M.M., Karim, Q.N., Sidebotham, R.L., and Worku, M.L. (1999). The relationship between *Helicobacter pylori* motility, morphology and phase of growth: implications for gastric colonization and pathology. *Microbiology* *145*, 2803–2811.
- Khemici, V., and Carpousis, A.J. (2004). The RNA degradosome and poly(A) polymerase of *Escherichia coli* are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Mol. Microbiol.* *51*, 777–790.
- Khemici, V., Poljak, L., Luisi, B.F., and Carpousis, A.J. (2008). The RNase E of *Escherichia coli* is a membrane-binding protein. *Mol. Microbiol.* *70*, 799–813.
- Kim, J.M., Kim, J.S., Lee, J.Y., Kim, Y.-J., Youn, H.-J., Kim, I.Y., Chee, Y.J., Oh, Y.-K., Kim, N., Jung, H.C., et al. (2007). Vacuolating Cytotoxin in *Helicobacter pylori* Water-Soluble Proteins Upregulates Chemokine

- Expression in Human Eosinophils via  $\text{Ca}^{2+}$  Influx, Mitochondrial Reactive Oxygen Intermediates, and NF- $\kappa$ B Activation. *Infect. Immun.* **75**, 3373–3381.
- Komarova, A. V., Tchufistova, L.S., Dreyfus, M., and Boni, I. V. (2005). AU-Rich Sequences within 5' Untranslated Leaders Enhance Translation and Stabilize mRNA in *Escherichia coli*. *J. Bacteriol.* **187**, 1344–1349.
- Kortmann, J., and Narberhaus, F. (2012). Bacterial RNA thermometers: molecular zippers and switches. *Nat. Rev. Microbiol.* **10**, 255–265.
- Kozak, M. (2005). Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**, 13–37.
- Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258.
- Kushner, S.R. (2002). mRNA decay in *Escherichia coli* comes of age. *J. Bacteriol.* **184**, 4658–65; discussion 4657.
- Kusters, J.G., van Vliet, A.H.M., and Kuipers, E.J. (2006). Pathogenesis of *Helicobacter pylori* Infection. *Clin. Microbiol. Rev.* **19**, 449–490.
- Kwok, T., Zabler, D., Urman, S., Rohde, M., Hartig, R., Wessler, S., Misselwitz, R., Berger, J., Sewald, N., König, W., et al. (2007). *Helicobacter* exploits integrin for type IV secretion and kinase activation. *Nature* **449**, 862–866.
- Kwon, A.-R., Kim, J.-H., Park, S.J., Lee, K.-Y., Min, Y.-H., Im, H., Lee, I., Lee, K.-Y., and Lee, B.-J. (2012). Structural and biochemical characterization of HP0315 from *Helicobacter pylori* as a VapD protein with an endoribonuclease activity. *Nucleic Acids Res.* **40**, 4216–4228.
- Laalami, S., Bessi eres, P., Rocca, A., Zig, L., Nicolas, P., and Putzer, H. (2013). *Bacillus subtilis* RNase Y Activity In Vivo Analysed by Tiling Microarrays. *PLoS One* **8**, e54062.
- Lai, D., Proctor, J.R., and Meyer, I.M. (2013). On the importance of cotranscriptional RNA structure formation. *RNA* **19**, 1461–1473.
- Lalaouna, D., and Mass e, E. (2015). Identification of sRNA interacting with a transcript of interest using MS2-affinity purification coupled with RNA sequencing (MAPS) technology. *Genomics Data* **5**, 136–138.
- De Lay, N., and Gottesman, S. (2011). Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA* **17**, 1172–1189.
- Lehnik-Habrink, M., Newman, J., Rothe, F.M., Solovyova, A.S., Rodrigues, C., Herzberg, C., Commichau, F.M., Lewis, R.J., and Stulke, J. (2011). RNase Y in *Bacillus subtilis*: a Natively Disordered Protein That Is the Functional Equivalent of RNase E from *Escherichia coli*. *J. Bacteriol.* **193**, 5431–5441.
- Li, Y., and Altman, S. (2003). A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. *Proc. Natl. Acad. Sci.* **100**, 13213–13218.
- Lin, E.A., Zhang, X.-S., Levine, S.M., Gill, S.R., Falush, D., and Blaser, M.J. (2009). Natural Transformation of *Helicobacter pylori* Involves the Integration of Short DNA Fragments Interrupted by Gaps of Variable Size. *PLoS Pathog.* **5**, e1000337.
- Linnstaedt, S., Kasprzak, W., Shapiro, B., and Casey, J. (2006). The role of a metastable RNA secondary structure in hepatitis delta virus genotype III RNA editing. *RNA* **12**, 1521–1533.
- Liu, Z., Trevi o, J., Ramirez-Pe a, E., and Sumby, P. (2012). The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen group A *Streptococcus*. *Mol. Microbiol.* **86**, 140–154.
- Livny, J., and Waldor, M.K. (2007). Identification of small RNAs in diverse bacterial species. *Curr. Opin. Microbiol.* **10**, 96–101.
- Lopez, P.J., Marchand, I., Joyce, S.A., and Dreyfus, M. (1999). The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Mol. Microbiol.* **33**, 188–199.
- Loughlin, M.F., Barnard, F.M., Jenkins, D., Sharples, G.J., and Jenks, P.J. (2003). *Helicobacter pylori* mutants defective in RuvC Holliday junction resolvase display reduced macrophage survival and spontaneous clearance from the murine gastric mucosa. *Infect. Immun.* **71**, 2022–2031.
- Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E.M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* **25**, 117–124.



- Mackie, G.A. (1998). Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* 395, 720–724.
- Mackie, G.A. (2000). Stabilization of circular rpsT mRNA demonstrates the 5'-end dependence of RNase E action in vivo. *J. Biol. Chem.* 275, 25069–25072.
- Maier, T., Schmidt, A., Guell, M., Kuhner, S., Gavin, A.-C., Aebersold, R., and Serrano, L. (2014). Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* 7, 511–511.
- Maixner, F., Krause-Kyora, B., Turaev, D., Herbig, A., Hoopmann, M.R., Hallows, J.L., Kusebauch, U., Vigl, E.E., Malferttheiner, P., Megraud, F., et al. (2016). The 5300-year-old *Helicobacter pylori* genome of the Iceman. *Science* 351, 162–165.
- Makarova, K.S., Grishin, N. V., and Koonin, E. V. (2006). The HicAB cassette, a putative novel, RNA-targeting toxin-antitoxin system in archaea and bacteria. *Bioinformatics* 22, 2581–2584.
- Mangold, M., Siller, M., Roppenser, B., Vlamincx, B.J.M., Penfound, T.A., Klein, R., Novak, R., Novick, R.P., and Charpentier, E. (2004). Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. *Mol. Microbiol.* 53, 1515–1527.
- Marincola, G., Schäfer, T., Behler, J., Bernhardt, J., Ohlsen, K., Goerke, C., and Wolz, C. (2012). RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. *Mol. Microbiol.* 85, 817–832.
- Marsich, E., Zuccato, P., Rizzi, S., Vetere, A., Tonin, E., and Paoletti, S. (2002). *Helicobacter pylori* expresses an autolytic enzyme: gene identification, cloning, and theoretical protein structure. *J. Bacteriol.* 184, 6270–6279.
- Marsin, S., Mathieu, A., Kortulewski, T., Guérois, R., and Radicella, J.P. (2008). Unveiling novel RecO distant orthologues involved in homologous recombination. *PLoS Genet.* 4.
- Marsin, S., Lopes, A., Mathieu, A., Dizet, E., Orillard, E., Guérois, R., and Radicella, J.P. (2010). Genetic dissection of *Helicobacter pylori* AddAB role in homologous recombination. *FEMS Microbiol. Lett.* 311, 44–50.
- Masachis, S., and Darfeuille, F. (2018). Type I Toxin-Antitoxin Systems: Regulating Toxin Expression via Shine-Dalgarno Sequence Sequestration and Small RNA Binding. *Microbiol. Spectr.* 6.
- Masachis, S., Tourasse, N.J., Chabas, S., Bouchez, O., and Darfeuille, F. (2018). FASTBAC-Seq: Functional Analysis of Toxin–Antitoxin Systems in BACteria by Deep Sequencing. *Methods Enzymol.*
- Masse, E., Escorcia, F.E., and Gottesman, S. (2003). Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.* 17, 2374–2383.
- Masse, E., Vanderpool, C.K., and Gottesman, S. (2005). Effect of RyhB Small RNA on Global Iron Use in *Escherichia coli*. *J. Bacteriol.* 187, 6962–6971.
- Massé, E., Salvail, H., Desnoyers, G., and Arguin, M. (2007). Small RNAs controlling iron metabolism. *Curr. Opin. Microbiol.* 10, 140–145.
- Mathy, N., Bénard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007). 5'-to-3' Exoribonuclease Activity in Bacteria: Role of RNase J1 in rRNA Maturation and 5' Stability of mRNA. *Cell* 129, 681–692.
- Matsunaga, J., Simons, E.L., and Simons, R.W. (1996). RNase III autoregulation: structure and function of rncO, the posttranscriptional “operator”. *RNA* 2, 1228–1240.
- McClain, M.S., Shaffer, C.L., Israel, D.A., Peek, R.M., and Cover, T.L. (2009). Genome sequence analysis of *Helicobacter pylori* strains associated with gastric ulceration and gastric cancer. *BMC Genomics* 10, 3.
- McDowall, K.J., and Cohen, S.N. (1996). The N-terminal Domain of the rne Gene Product has RNase E Activity and is Non-overlapping with the Arginine-rich RNA-binding Site. *J. Mol. Biol.* 255, 349–355.
- McDowall, K.J., Lin-Chao, S., and Cohen, S.N. (1994). A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* 269, 10790–10796.
- McGary, K., and Nudler, E. (2013). RNA polymerase and the ribosome: the close relationship. *Curr. Opin. Microbiol.* 16, 112–117.
- van Meerten, D., Girard, G., and van Duin, J. (2001). Translational control by delayed RNA folding: identification of the kinetic trap. *RNA* 7, 483–494.
- Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y.E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., and Margalit, H. (2016). Global Mapping of Small RNA-Target Interactions in Bacteria. *Mol. Cell* 63, 884–897.
- Mell, J.C., and Redfield, R.J. (2014). Natural competence and the evolution of DNA uptake specificity. *J. Bacteriol.* 196, 1471–1483.



- Meng, W., and Nicholson, A.W. (2008). Heterodimer-based analysis of subunit and domain contributions to double-stranded RNA processing by *Escherichia coli* RNase III *in vitro*. *Biochem. J.* 410, 39–48.
- Meyer, I.M. (2017a). In silico methods for co-transcriptional RNA secondary structure prediction and for investigating alternative RNA structure expression. *Methods* 120, 3–16.
- Meyer, M.M. (2017b). The role of mRNA structure in bacterial translational regulation. *Wiley Interdiscip. Rev. RNA* 8, e1370.
- Miczak, A., Kaberdin, V.R., Wei, C.L., and Lin-Chao, S. (1996). Proteins associated with RNase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. U. S. A.* 93, 3865–3869.
- Milón, P., and Rodnina, M. V. (2012). Kinetic control of translation initiation in bacteria. *Crit. Rev. Biochem. Mol. Biol.* 47, 334–348.
- Miroux, B., and Walker, J.E. (1996). Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol.* 260, 289–298.
- Mizote, T., Yoshiyama, H., and Nakazawa, T. (1997). Urease-independent chemotactic responses of *Helicobacter pylori* to urea, urease inhibitors, and sodium bicarbonate. *Infect. Immun.* 65, 1519–1521.
- Mizuno, T., Chou, M.Y., and Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. U. S. A.* 81, 1966–1970.
- Moffatt, B.A., and Studier, F.W. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* 49, 221–227.
- Mohanty, B.K., and Kushner, S.R. (2010). Processing of the *Escherichia coli* leuX tRNA transcript, encoding tRNA(Leu5), requires either the 3'→5' exoribonuclease polynucleotide phosphorylase or RNase P to remove the Rho-independent transcription terminator. *Nucleic Acids Res.* 38, 597–607.
- Mok, W.W.K., Patel, N.H., and Li, Y. (2010). Decoding Toxicity. *J. Biol. Chem.* 285, 41627–41636.
- Møller-Jensen, J., Franch, T., and Gerdes, K. (2001). Temporal Translational Control by a Metastable RNA Structure. *J. Biol. Chem.* 276, 35707–35713.
- Moodley, Y., and Linz, B. (2009). <i>Helicobacter pylori</i> Sequences Reflect Past Human Migrations. In *Microbial Pathogenomics*, (Basel: KARGER), pp. 62–74.
- Moodley, Y., Linz, B., Yamaoka, Y., Windsor, H.M., Breurec, S., Wu, J.-Y., Maady, A., Bernhoft, S., Thiberge, J.-M., Phuanukoonnon, S., et al. (2009). The Peopling of the Pacific from a Bacterial Perspective. *Science* (80-. ). 323, 527–530.
- Moodley, Y., Linz, B., Bond, R.P., Nieuwoudt, M., Soodyall, H., Schlebusch, C.M., Bernhöft, S., Hale, J., Suerbaum, S., Mugisha, L., et al. (2012). Age of the Association between *Helicobacter pylori* and Man. *PLoS Pathog.* 8, e1002693.
- Morita, T., Maki, K., and Aiba, H. (2005a). RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* 19, 2176–2186.
- Morita, T., Maki, K., and Aiba, H. (2005b). RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* 19, 2176–2186.
- Mörl, M., and Marchfelder, A. (2001). The final cut. *EMBO Rep.* 2, 17–20.
- El Mortaji, L., Aubert, S., Galtier, E., Schmitt, C., Anger, K., Redko, Y., Quentin, Y., and De Reuse, H. (2018). The Sole DEAD-Box RNA Helicase of the Gastric Pathogen *Helicobacter pylori* Is Essential for Colonization. *MBio* 9, e02071-17.
- Mortier-Barrière, I., Velten, M., Dupaigne, P., Mirouze, N., Piétrement, O., McGovern, S., Fichant, G., Martin, B., Noirot, P., Le Cam, E., et al. (2007). A Key Presynaptic Role in Transformation for a Widespread Bacterial Protein: DprA Conveys Incoming ssDNA to RecA. *Cell* 130, 824–836.
- Moss, S.F. (2017). The Clinical Evidence Linking *Helicobacter pylori* to Gastric Cancer. *Cell. Mol. Gastroenterol. Hepatol.* 3, 183–191.
- Mott, J.E., Galloway, J.L., and Platt, T. (1985). Maturation of *Escherichia coli* tryptophan operon mRNA: evidence for 3' exonucleolytic processing after rho-dependent termination. *EMBO J.* 4, 1887–1891.
- Murata-Kamiya, N. (2011). Pathophysiological functions of the CagA oncoprotein during infection by *Helicobacter pylori*. *Microbes Infect.* 13, 799–807.
- Mustoe, A.M., Busan, S., Rice, G.M., Hajdin, C.E., Peterson, B.K., Ruda, V.M., Kubica, N., Nutiu, R., Baryza,

- J.L., and Weeks, K.M. (2018). Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* 173, 181–195.e18.
- Nagase, L., Murata-Kamiya, N., and Hatakeyama, M. (2011). Potentiation of *Helicobacter pylori* CagA Protein Virulence through Homodimerization. *J. Biol. Chem.* 286, 33622–33631.
- Nagel, J.H., Gultyaev, A.P., Gerdes, K., and Pleij, C.W. (1999). Metastable structures and refolding kinetics in hok mRNA of plasmid R1. *RNA* 5, 1408–1418.
- Nakagawa, S., Niimura, Y., and Gojobori, T. (2017). Comparative genomic analysis of translation initiation mechanisms for genes lacking the Shine–Dalgarno sequence in prokaryotes. *Nucleic Acids Res.* 45, 3922–3931.
- Nakamura, H., Yoshiyama, H., Takeuchi, H., Mizote, T., Okita, K., and Nakazawa, T. (1998). Urease plays an important role in the chemotactic motility of *Helicobacter pylori* in a viscous environment. *Infect. Immun.* 66, 4832–4837.
- Narczyk, M., Bertoša, B., Papa, L., Vuković, V., Lešćić Ašler, I., Wielgus-Kutrowska, B., Bzowska, A., Luić, M., and Štefanić, Z. (2018). *Helicobacter pylori* purine nucleoside phosphorylase shows new distribution patterns of open and closed active site conformations and unusual biochemical features. *FEBS J.* 285, 1305–1325.
- Nicolas, P., Mader, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., et al. (2012). Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*. *Science* (80-. ). 335, 1103–1106.
- Nielsen, A.K., Thorsted, P., Thisted, T., Wagner, E.G., and Gerdes, K. (1991). The rifampicin-inducible genes *srnB* from F and *pnd* from R483 are regulated by antisense RNAs and mediate plasmid maintenance by killing of plasmid-free segregants. *Mol. Microbiol.* 5, 1961–1973.
- Olbermann, P., Josenhans, C., Moodley, Y., Uhr, M., Stamer, C., Vauterin, M., Suerbaum, S., Achtman, M., and Linz, B. (2010). A global overview of the genetic and functional diversity in the *helicobacter pylori* cag pathogenicity island. *PLoS Genet.* 6.
- Olekhovich, I.N., Vitko, S., Valliere, M., and Hoffman, P.S. (2014). Response to Metronidazole and Oxidative Stress Is Mediated through Homeostatic Regulator HsrA (HP1043) in *Helicobacter pylori*. *J. Bacteriol.* 196, 729–739.
- Opdyke, J.A., Kang, J.-G., and Storz, G. (2004). GadY, a Small-RNA Regulator of Acid Response Genes in *Escherichia coli*. *J. Bacteriol.* 186, 6698–6705.
- Palframan, S.L., Kwok, T., and Gabriel, K. (2012). Vacuolating cytotoxin A (VacA), a key toxin for *Helicobacter pylori* pathogenesis. *Front. Cell. Infect. Microbiol.* 2, 92.
- Pan, J., and Woodson, S.A. (1998). Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.* 280, 597–609.
- Papenfort, K., and Bassler, B.L. (2016). Quorum sensing signal–response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* 14, 576–588.
- Papenfort, K., and Vogel, J. (2014). Small RNA functions in carbon metabolism and virulence of enteric pathogens. *Front. Cell. Infect. Microbiol.* 4.
- Papini, E., Satin, B., Norais, N., de Bernard, M., Telford, J.L., Rappuoli, R., and Montecucco, C. (1998). Selective increase of the permeability of polarized epithelial cell monolayers by *Helicobacter pylori* vacuolating toxin. *J. Clin. Invest.* 102, 813–820.
- Pathak, C., Im, H., Yang, Y.-J., Yoon, H.-J., Kim, H.-M., Kwon, A.-R., and Lee, B.-J. (2013). Crystal structure of apo and copper bound HP0894 toxin from *Helicobacter pylori* 26695 and insight into mRNAse activity. *Biochim. Biophys. Acta - Proteins Proteomics* 1834, 2579–2590.
- Pedersen, K., and Gerdes, K. (1999). Multiple hok genes on the chromosome of *Escherichia coli*. *Mol. Microbiol.* 32, 1090–1102.
- Pedrolli, D., Langer, S., Hobl, B., Schwarz, J., Hashimoto, M., and Mack, M. (2015). The ribB FMN riboswitch from *Escherichia coli* operates at the transcriptional and translational level and regulates riboflavin biosynthesis. *FEBS J.* 282, 3230–3242.
- Pellicciari, S., Pinatel, E., Vannini, A., Peano, C., Puccio, S., De Bellis, G., Danielli, A., Scarlato, V., and Roncarati, D. (2017). Insight into the essential role of the *Helicobacter pylori* HP1043 orphan response regulator: genome-wide identification and characterization of the DNA-binding sites. *Sci. Rep.* 7, 41063.
- Pernitzsch, S.R., and Sharma, C.M. (2012). Transcriptome Complexity and Riboregulation in the Human

- Pathogen *Helicobacter pylori*. *Front. Cell. Infect. Microbiol.* 2, 14.
- Pernitzsch, S.R., Tirier, S.M., Beier, D., and Sharma, C.M. (2014). A variable homopolymeric G-repeat defines small RNA-mediated posttranscriptional regulation of a chemotaxis receptor in *Helicobacter pylori*. *Proc. Natl. Acad. Sci. U. S. A.* 111, E501-510.
- Perwez, T., and Kushner, S.R. (2006). RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Mol. Microbiol.* 60, 723–737.
- Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J.C.D., and Vogel, J. (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.* 16, 840–846.
- Pflock, M., Dietz, P., Schär, J., and Beier, D. (2004). Genetic evidence for histidine kinase HP165 being an acid sensor of *Helicobacter pylori*. *FEMS Microbiol. Lett.* 234, 51–61.
- Pflock, M., Kennard, S., Delany, I., Scarlato, V., and Beier, D. (2005). Acid-Induced Activation of the Urease Promoters Is Mediated Directly by the ArsRS Two-Component System of *Helicobacter pylori*. *Infect. Immun.* 73, 6437–6445.
- Pflock, M., Finsterer, N., Joseph, B., Mollenkopf, H., Meyer, T.F., and Beier, D. (2006). Characterization of the ArsRS Regulon of *Helicobacter pylori*, Involved in Acid Adaptation. *J. Bacteriol.* 188, 3449–3462.
- Pflock, M., Bathon, M., Schar, J., Muller, S., Mollenkopf, H., Meyer, T.F., and Beier, D. (2007). The Orphan Response Regulator HP1021 of *Helicobacter pylori* Regulates Transcription of a Gene Cluster Presumably Involved in Acetone Metabolism. *J. Bacteriol.* 189, 2339–2349.
- Picard, F., Milhem, H., Loubière, P., Laurent, B., Coccagn-Bousquet, M., and Girbal, L. (2012). Bacterial translational regulations: high diversity between all mRNAs and major role in gene expression. *BMC Genomics* 13, 528.
- Plummer, M., Franceschi, S., Vignat, J., Forman, D., and de Martel, C. (2015). Global burden of gastric cancer attributable to *Helicobacter pylori*. *Int. J. Cancer* 136, 487–490.
- Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation. *Science* (80-. ). 328, 504–508.
- Py, B., Causton, H., Mudd, E.A., and Higgins, C.F. (1994). A protein complex mediating mRNA degradation in *Escherichia coli*. *Mol. Microbiol.* 14, 717–729.
- Py, B., Higgins, C.F., Krisch, H.M., and Carpousis, A.J. (1996). A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* 381, 169–172.
- Raghavan, R., Groisman, E.A., and Ochman, H. (2011). Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.* 21, 1487–1497.
- Rajaure, M., Berry, J., Kongari, R., Cahill, J., and Young, R. (2015). Membrane fusion during phage lysis. *Proc. Natl. Acad. Sci.* 112, 5497–5502.
- Redko, Y., Aubert, S., Stachowicz, A., Lenormand, P., Namane, A., Darfeuille, F., Thibonnier, M., and De Reuse, H. (2013). A minimal bacterial RNase J-based degradosome is associated with translating ribosomes. *Nucleic Acids Res.* 41, 288–301.
- Redko, Y., Galtier, E., Arnion, H., Darfeuille, F., Sismeiro, O., Coppée, J.-Y., Médigue, C., Weiman, M., Cruveiller, S., and De Reuse, H. (2016). RNase J depletion leads to massive changes in mRNA abundance in *Helicobacter pylori*. *RNA Biol.* 13, 243–253.
- Reif, C., Löser, C., and Brantl, S. (2018). *Bacillus subtilis* Type I antitoxin SR6 Promotes Degradation of Toxin yonT mRNA and Is Required to Prevent Toxic yoyJ Overexpression. *Toxins (Basel)*. 10, 74.
- Repsilber, D., Wiese, S., Rachen, M., Schröder, A.W., Riesner, D., and Steger, G. (1999). Formation of metastable RNA structures by sequential folding during transcription: Time-resolved structural analysis of potato spindle tuber viroid (–)-stranded RNA by temperature-gradient gel electrophoresis. 574–584.
- Richards, J., Liu, Q., Pellegrini, O., Celesnik, H., Yao, S., Bechhofer, D.H., Condon, C., and Belasco, J.G. (2011). An RNA pyrophosphohydrolase triggers 5'-exonucleolytic degradation of mRNA in *Bacillus subtilis*. *Mol. Cell* 43, 940–949.
- Richardson, L. V, and Richardson, J.P. (1996). Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (rut) sequences of a terminator. *J. Biol. Chem.* 271, 21597–21603.
- Robertson, H.D. (1982). *Escherichia coli* ribonuclease III cleavage sites. *Cell* 30, 669–672.

- Romero, A., and García, P. (1991). Initiation of translation at AUC, AUA and AUU codons in *Escherichia coli*. *FEMS Microbiol. Lett.* *68*, 325–330.
- Roncarati, D., Danielli, A., Spohn, G., Delany, I., and Scarlato, V. (2007). Transcriptional Regulation of Stress Response and Motility Functions in *Helicobacter pylori* Is Mediated by HspR and HrcA. *J. Bacteriol.* *189*, 7234–7243.
- Roncarati, D., Pellicciari, S., Doniselli, N., Maggi, S., Vannini, A., Valzania, L., Mazzei, L., Zambelli, B., Rivetti, C., and Danielli, A. (2016). Metal-responsive promoter DNA compaction by the ferric uptake regulator. *Nat. Commun.* *7*, 12593.
- Sacerdot, C., Chiaruttini, C., Engst, K., Graffe, M., Milet, M., Mathy, N., Dondon, J., and Springer, M. (1996). The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in *Escherichia coli*. *Mol. Microbiol.* *21*, 331–346.
- Sachs, G., Scott, D., Weeks, D., and Melchers, K. (2002). The compartment buffered by the urease of *Helicobacter pylori*: cytoplasm or periplasm? *Trends Microbiol.* *10*, 217-8; author reply 218-9.
- Salvo, E., Alabi, S., Liu, B., Schlessinger, A., and Bechhofer, D.H. (2016). Interaction of *Bacillus subtilis* Polynucleotide Phosphorylase and RNase Y. *J. Biol. Chem.* *291*, 6655–6663.
- Sanamrad, A., Persson, F., Lundius, E.G., Fange, D., Gynna, A.H., and Elf, J. (2014). Single-particle tracking reveals that free ribosomal subunits are not excluded from the *Escherichia coli* nucleoid. *Proc. Natl. Acad. Sci.* *111*, 11413–11418.
- Sayed, N., Nonin-Lecomte, S., Réty, S., and Felden, B. (2012). Functional and structural insights of a *Staphylococcus aureus* apoptotic-like membrane peptide from a toxin-antitoxin module. *J. Biol. Chem.* *287*, 43454–43463.
- Scarlato, V., Delany, I., Spohn, G., and Beier, D. (2001). Regulation of transcription in *Helicobacter pylori*: simple systems or complex circuits? *Int. J. Med. Microbiol.* *291*, 107–117.
- Schmitt, W., Odenbreit, S., Heuermann, D., and Haas, R. (1995). Cloning of the *Helicobacter pylori* recA gene and functional characterization of its product. *Mol. Gen. Genet.* *248*, 563–572.
- Schreiber, S., Bücker, R., Groll, C., Azevedo-Vethacke, M., Garten, D., Scheid, P., Friedrich, S., Gattermann, S., Josenhans, C., and Suerbaum, S. (2005). Rapid loss of motility of *Helicobacter pylori* in the gastric lumen in vivo. *Infect. Immun.* *73*, 1584–1589.
- Serganov, A., Polonskaia, A., Ehresmann, B., Ehresmann, C., and Patel, D.J. (2003). Ribosomal protein S15 represses its own translation via adaptation of an rRNA-like fold within its mRNA. *EMBO J.* *22*, 1898–1908.
- Shao, Y., Feng, L., Rutherford, S.T., Papenfort, K., and Bassler, B.L. (2013). Functional determinants of the quorum-sensing non-coding RNAs and their roles in target regulation. *EMBO J.* *32*, 2158–2171.
- Sharma, C.M., Darfeuille, F., Plantinga, T.H., and Vogel, J. (2007). A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.* *21*, 2804–2817.
- Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., et al. (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* *464*, 250–255.
- Sharma, I.M., Rappé, M.C., Addepalli, B., Grabow, W.W., Zhuang, Z., Abeysirigunawardena, S.C., Limbach, P.A., Jaeger, L., and Woodson, S.A. (2018). A metastable rRNA junction essential for bacterial 30S biogenesis. *Nucleic Acids Res.* *46*, 5182–5194.
- Shine, J., and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. U. S. A.* *71*, 1342–1346.
- Silvaggi, J.M., Perkins, J.B., and Losick, R. (2005a). Small Untranslated RNA Antitoxin in *Bacillus subtilis*. *J. Bacteriol.* *187*, 6641–6650.
- Silvaggi, J.M., Perkins, J.B., and Losick, R. (2005b). Small untranslated RNA antitoxin in *Bacillus subtilis*. *J. Bacteriol.* *187*, 6641–6650.
- Simonetti, A., Marzi, S., Jenner, L., Myasnikov, A., Romby, P., Yusupova, G., Klaholz, B.P., and Yusupov, M. (2009). A structural view of translation initiation in bacteria. *Cell. Mol. Life Sci.* *66*, 423–436.
- Simonetti, A., Marzi, S., Billas, I.M.L., Tsai, A., Fabbretti, A., Myasnikov, A.G., Roblin, P., Vaiana, A.C., Hazemann, I., Eiler, D., et al. (2013). Involvement of protein IF2 N domain in ribosomal subunit joining revealed from architecture and function of the full-length initiation factor. *Proc. Natl. Acad. Sci.* *110*, 15656–

15661.

Simons, R.W., and Kleckner, N. (1983). Translational control of IS10 transposition. *Cell* 34, 683–691.

Sinha, D., Matz, L., Cameron, T., and De Lay, N.R. (2018). Poly(A) polymerase is required for RyhB sRNA stability and function in *Escherichia coli*. *RNA* rna.067181.118.

Sisto, F., Brenciaglia, M.I., Scaltrito, M.M., and Dubini, F. (2000). *Helicobacter pylori*: ureA, cagA and vacA expression during conversion to the coccoid form. *Int. J. Antimicrob. Agents* 15, 277–282.

Skouloubris, S., Thiberge, J.M., Labigne, A., and De Reuse, H. (1998). The *Helicobacter pylori* Urel protein is not involved in urease activity but is essential for bacterial survival in vivo. *Infect. Immun.* 66, 4517–4521.

Smeets, L.C., Bijlsma, J.J., Kuipers, E.J., Vandenbroucke-Grauls, C.M., and Kusters, J.G. (2000). The dprA gene is required for natural transformation of *Helicobacter pylori*. *FEMS Immunol. Med. Microbiol.* 27, 99–102.

de Smit, M.H., and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. U. S. A.* 87, 7668–7672.

de Smit, M.H., and van Duin, J. (2003). Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J. Mol. Biol.* 331, 737–743.

Solnick, J. V., and Schauer, D.B. (2001). Emergence of Diverse *Helicobacter* Species in the Pathogenesis of Gastric and Enterohepatic Diseases. *Clin. Microbiol. Rev.* 14, 59–97.

Song, T., Mika, F., Lindmark, B., Liu, Z., Schild, S., Bishop, A., Zhu, J., Camilli, A., Johansson, J., Vogel, J., et al. (2008). A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Mol. Microbiol.* 70, 100–111.

Spickler, C., and Mackie, G.A. (2000). Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J. Bacteriol.* 182, 2422–2427.

Štefanić, Z., Mikleušević, G., Luić, M., Bzowska, A., and Leščić Ašler, I. (2017). Structural characterization of purine nucleoside phosphorylase from human pathogen *Helicobacter pylori*. *Int. J. Biol. Macromol.* 101, 518–526.

Steinbrecher, T., Prock, S., Reichert, J., Wadhwani, P., Zimpfer, B., Bürck, J., Berditsch, M., Elstner, M., and Ulrich, A.S. (2012). Peptide-lipid interactions of the stress-response peptide tisb that induces bacterial persistence. *Biophys. J.* 103, 1460–1469.

Steitz, J.A., and Jakes, K. (1975). How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 72, 4734–4738.

Stenström, C.M., Holmgren, E., and Isaksson, L.A. (2001). Cooperative effects by the initiation codon and its flanking regions on translation initiation. *Gene* 273, 259–265.

Sterk, M., Romilly, E., Gerhart, E., and Wagner, H. (2018). Unstructured 5'-tails act through ribosome standby to override inhibitory structure at ribosome binding sites. *Nucleic Acids Res.*

Storz, G., Vogel, J., and Wassarman, K.M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* 43, 880–891.

Stougaard, P., Molin, S., and Nordström, K. (1981). RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc. Natl. Acad. Sci. U. S. A.* 78, 6008–6012.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185, 60–89.

Sundrud, M.S., Torres, V.J., Unutmaz, D., and Cover, T.L. (2004). Inhibition of primary human T cell proliferation by *Helicobacter pylori* vacuolating toxin (VacA) is independent of VacA effects on IL-2 secretion. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7727–7732.

Svenningsen, S. Lo (2018). Small RNA-Based Regulation of Bacterial Quorum Sensing and Biofilm Formation. *Microbiol. Spectr.* 6.

Sycuro, L.K., Wyckoff, T.J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N.R. (2012). Multiple Peptidoglycan Modification Networks Modulate *Helicobacter pylori*'s Cell Shape, Motility, and Colonization Potential. *PLoS Pathog.* 8, e1002603.

Symmons, M.F., Jones, G.H., and Luisi, B.F. (2000). A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* 8, 1215–1226.



- Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., et al. (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* *45*, D362–D368.
- Tan, S., Noto, J.M., Romero-Gallo, J., Peek, R.M., and Amieva, M.R. (2011). *Helicobacter pylori* Perturbs Iron Trafficking in the Epithelium to Grow on the Cell Surface. *PLoS Pathog.* *7*, e1002050.
- La Teana, A., Pon, C.L., and Gualerzi, C.O. (1993). Translation of mRNAs with degenerate initiation triplet AUU displays high initiation factor 2 dependence and is subject to initiation factor 3 repression. *Proc. Natl. Acad. Sci. U. S. A.* *90*, 4161–4165.
- Terry, C.E., McGinnis, L.M., Madigan, K.C., Cao, P., Cover, T.L., Liechti, G.W., Peek, R.M., and Forsyth, M.H. (2005). Genomic Comparison of *cag* Pathogenicity Island (PAI)-Positive and -Negative *Helicobacter pylori* Strains: Identification of Novel Markers for *cag* PAI-Positive Strains. *Infect. Immun.* *73*, 3794–3798.
- Thisted, T., and Gerdes, K. (1992). Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1. *Sok* antisense RNA regulates *hok* gene expression indirectly through the overlapping *mok* gene. *J. Mol. Biol.* *223*, 41–54.
- Thisted, T., Sorensen, N.S., Wagner, E.G.H., and Gerdes, K. (1994). Mechanism of post-segregational killing : *Sok* antisense RNA interacts with *Hok* mRNA via its 5' -end single- stranded leader and competes with the 3' -end of *Hok* mRNA for binding to the *mok* translational initiation region. *13*, 1960–1968.
- Thomason, M.K., Bischler, T., Eisenbart, S.K., Förstner, K.U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C.M., and Storz, G. (2015). Global Transcriptional Start Site Mapping Using Differential RNA Sequencing Reveals Novel Antisense RNAs in *Escherichia coli*. *J. Bacteriol.* *197*, 18–28.
- Thompson, L.J., and de Reuse, H. (2002). Genomics of *Helicobacter pylori*. *Helicobacter* *7 Suppl 1*, 1–7.
- Thompson, S.A., and Blaser, M.J. (1995). Isolation of the *Helicobacter pylori* *recA* gene and involvement of the *recA* region in resistance to low pH. *Infect. Immun.* *63*, 2185–2193.
- Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., et al. (1997a). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* *388*, 539–547.
- Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., et al. (1997b). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* *388*, 539–547.
- Tomizawa, J., Itoh, T., Selzer, G., and Som, T. (1981). Inhibition of *ColE1* RNA primer formation by a plasmid-specified small RNA. *Proc. Natl. Acad. Sci. U. S. A.* *78*, 1421–1425.
- Torii, N., Nozaki, T., Masutani, M., Nakagama, H., Sugiyama, T., Saito, D., Asaka, M., Sugimura, T., and Miki, K. (2003). Spontaneous mutations in the *Helicobacter pylori* *rpsL* gene. *Mutat. Res.* *535*, 141–145.
- Tramonti, A., De Canio, M., and De Biase, D. (2008). *GadX/GadW*-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the *gadY-gadW* divergent promoters and identification of four novel 42 bp *GadX/GadW*-specific binding sites. *Mol. Microbiol.* *70*, 965–982.
- Tsao, M.-Y., Lin, T.-L., Hsieh, P.-F., and Wang, J.-T. (2009). The 3'-to-5' Exoribonuclease (Encoded by HP1248) of *Helicobacter pylori* Regulates Motility and Apoptosis-Inducing Genes. *J. Bacteriol.* *191*, 2691–2702.
- Tsui, H.-C.T., Mukherjee, D., Ray, V.A., Sham, L.-T., Feig, A.L., and Winkler, M.E. (2010). Identification and Characterization of Noncoding Small RNAs in *Streptococcus pneumoniae* Serotype 2 Strain D39. *J. Bacteriol.* *192*, 264–279.
- Tummuru, M.K., Cover, T.L., and Blaser, M.J. (1993). Cloning and expression of a high-molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. *Infect. Immun.* *61*, 1799–1809.
- Uhm, H., Kang, W., Ha, K.S., Kang, C., and Hohng, S. (2018). Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. *Proc. Natl. Acad. Sci.* *115*, 331–336.
- Unoson, C., and Wagner, E.G.H. (2008). A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *70*, 258–270.
- Updegrove, T.B., Kouse, A.B., Bandyra, K.J., and Storz, G. (2018). Stem-loops direct precise processing of 3' UTR-derived small RNA *MicL*. *BioRxiv* 341578.
- Urban, J.H., and Vogel, J. (2007). Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res.* *35*, 1018–1037.

- Valentin-Hansen, P., Johansen, J., and Rasmussen, A.A. (2007). Small RNAs controlling outer membrane porins. *Curr. Opin. Microbiol.* *10*, 152–155.
- Vannini, A., Roncarati, D., and Danielli, A. (2016). The *cag*-pathogenicity island encoded CncR1 sRNA oppositely modulates *Helicobacter pylori* motility and adhesion to host cells. *Cell. Mol. Life Sci.* *73*, 3151–3168.
- Vanzo, N.F., Li, Y.S., Py, B., Blum, E., Higgins, C.F., Raynal, L.C., Krisch, H.M., and Carpousis, A.J. (1998). Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* *12*, 2770–2781.
- Verstraeten, N., Knapen, W.J., Kint, C.I., Liebens, V., Van den Bergh, B., Dewachter, L., Michiels, J.E., Fu, Q., David, C.C., Fierro, A.C., et al. (2015a). O<sub>bg</sub> and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. *Mol. Cell* *59*, 9–21.
- Verstraeten, N., Knapen, J., Kint, I., Jansen, M., Fauvart, M., and Correspondence, J.M. (2015b). O<sub>bg</sub> and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance.
- Viegas, S.C., Apura, P., Martínez-García, E., de Lorenzo, V., and Arraiano, C.M. (2018). Modulating Heterologous Gene Expression with Portable mRNA-Stabilizing 5'-UTR Sequences. *ACS Synth. Biol.* *acssynbio.8b00191*.
- van Vliet, A.H.M., Poppelaars, S.W., Davies, B.J., Stoof, J., Bereswill, S., Kist, M., Penn, C.W., Kuipers, E.J., and Kusters, J.G. (2002). NikR mediates nickel-responsive transcriptional induction of urease expression in *Helicobacter pylori*. *Infect. Immun.* *70*, 2846–2852.
- Vogel, J., Argaman, L., Wagner, E.G.H., and Altuvia, S. (2004). The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr. Biol.* *14*, 2271–2276.
- Wadler, C.S., and Vanderpool, C.K. (2007). A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc. Natl. Acad. Sci.* *104*, 20454–20459.
- Wagner, E.G.H., Altuvia, S., and Romby, P. (2002). Antisense RNAs in bacteria and their genetic elements. *Adv. Genet.* *46*, 361–398.
- Wang, I.-N., Smith, D.L., and Young, R. (2000). Holins: The Protein Clocks of Bacteriophage Infections. *Annu. Rev. Microbiol.* *54*, 799–825.
- Wang, J., Rennie, W., Liu, C., Carmack, C.S., Prévost, K., Caron, M.-P., Massé, E., Ding, Y., and Wade, J.T. (2015). Identification of bacterial sRNA regulatory targets using ribosome profiling. *Nucleic Acids Res.* *43*, 10308–10320.
- Watters, K.E., Strobel, E.J., Yu, A.M., Lis, J.T., and Lucks, J.B. (2016). Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.* *23*, 1124–1131.
- Weaver, K.E. (2012). The *par* toxin-antitoxin system from *Enterococcus faecalis* plasmid pAD1 and its chromosomal homologs. *9*, 1498–1503.
- Weel-Sneve, R., Kristiansen, K.I., Odsbu, I., Dalhus, B., Booth, J., Rognes, T., Skarstad, K., and Bjørås, M. (2013). Single Transmembrane Peptide DinQ Modulates Membrane-Dependent Activities. *PLoS Genet* *9*.
- Weilbacher, T., Suzuki, K., Dubey, A.K., Wang, X., Gudapaty, S., Morozov, I., Baker, C.S., Georgellis, D., Babitzke, P., and Romeo, T. (2003). A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol. Microbiol.* *48*, 657–670.
- Weinberg, Z., Barrick, J.E., Yao, Z., Roth, A., Kim, J.N., Gore, J., Wang, J.X., Lee, E.R., Block, K.F., Sudarsan, N., et al. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* *35*, 4809–4819.
- Wen, J., and Fozo, E.M. (2014). sRNA antitoxins: more than one way to repress a toxin. *Toxins (Basel)*. *6*, 2310–2335.
- Wen, J., Won, D., and Fozo, E.M. (2014). The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin. *Nucleic Acids Res.* *42*, 1930–1946.
- Wen, J., Harp, J.R., and Fozo, E.M. (2017). The 5' UTR of the type I toxin ZorO can both inhibit and enhance translation. *Nucleic Acids Res.* *45*, 4006–4020.
- Wen, Y., Feng, J., and Sachs, G. (2013). *Helicobacter pylori* 5'ureB-sRNA, a cis-Encoded Antisense Small RNA, Negatively Regulates ureAB Expression by Transcription Termination. *J. Bacteriol.* *195*, 444–452.
- Wickiser, J.K., Winkler, W.C., Breaker, R.R., and Crothers, D.M. (2005). The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell* *18*, 49–60.

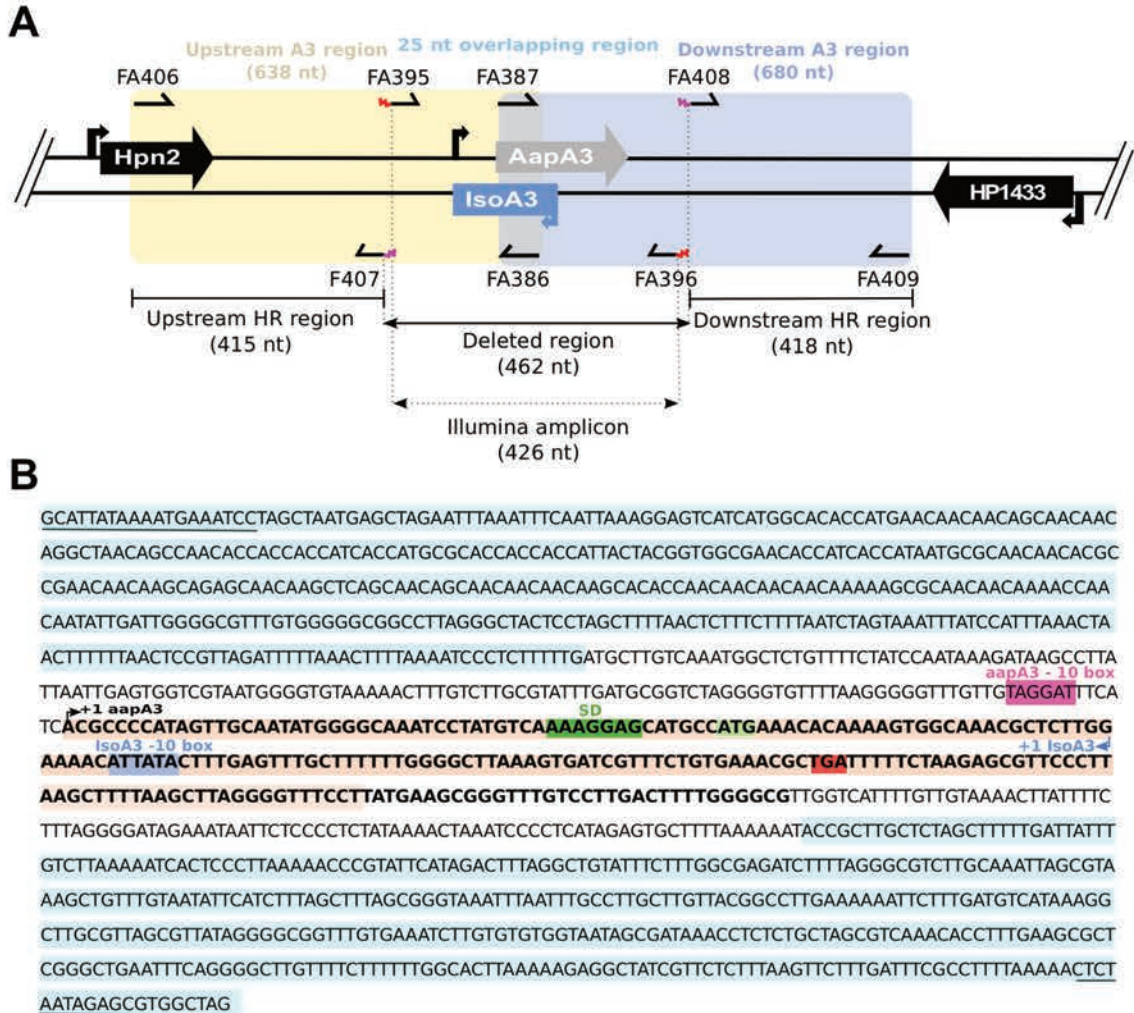
- Williams, J.J., and Hergenrother, P.J. (2012). Artificial activation of toxin-antitoxin systems as an antibacterial strategy. *Trends Microbiol.* *20*, 291–298.
- Wilmaerts, D., Bayoumi, M., Dewachter, L., Knapen, W., Mika, J.T., Hofkens, J., Dedecker, P., Maglia, G., Verstraeten, N., and Michiels, J. (2018). The Persistence-Inducing Toxin HokB Forms Dynamic Pores That Cause ATP Leakage.
- Wirth, T., Wang, X., Linz, B., Novick, R.P., Lum, J.K., Blaser, M., Morelli, G., Falush, D., and Achtman, M. (2004). Distinguishing human ethnic groups by means of sequences from *Helicobacter pylori*: Lessons from Ladakh. *Proc. Natl. Acad. Sci.* *101*, 4746–4751.
- Worku, M.L., Karim, Q.N., Spencer, J., and Sidebotham, R.L. (2004). Chemotactic response of *Helicobacter pylori* to human plasma and bile. *J. Med. Microbiol.* *53*, 807–811.
- Wroblewski, L.E., Peek, R.M., and Wilson, K.T. (2010). *Helicobacter pylori* and Gastric Cancer: Factors That Modulate Disease Risk. *Clin. Microbiol. Rev.* *23*, 713–739.
- Wu, J., Xu, S., and Zhu, Y. (2013). *Helicobacter pylori* CagA: A Critical Destroyer of the Gastric Epithelial Barrier. *Dig. Dis. Sci.* *58*, 1830–1837.
- Xu, F., and Cohen, S.N. (1995). RNA degradation in *Escherichia coli* regulated by 3' adenylation and 5' phosphorylation. *Nature* *374*, 180–183.
- Yang, Q., Figueroa-Bossi, N., and Bossi, L. (2014). Translation Enhancing ACA Motifs and Their Silencing by a Bacterial Small Regulatory RNA. *PLoS Genet.* *10*, e1004026.
- Young, R. (2014). Phage lysis: Three steps, three choices, one outcome. *J. Microbiol.* *52*, 243–258.
- Young, K.A., Allaker, R.P., and Hardie, J.M. (2001). Morphological analysis of *Helicobacter pylori* from gastric biopsies and dental plaque by scanning electron microscopy. *Oral Microbiol. Immunol.* *16*, 178–181.
- Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* *50*, 1111–1124.
- Zhu, A.Q., and von Hippel, P.H. (1998). Rho-dependent Termination within the *trp* *t*<sup>+</sup> Terminator. II. Effects of Kinetic Competition and Rho Processivity<sup>†</sup>. *Biochemistry* *37*, 11215–11222.
- Zhu, J.Y.A., and Meyer, I.M. (2015). Four RNA families with functional transient structures. *RNA Biol.* *12*, 5–20.
- Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* *29*, 1017–1026.
- (1994). Schistosomes, liver flukes and *Helicobacter pylori*. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. IARC Monogr. Eval. Carcinog. Risks to Humans *61*, 1–241.



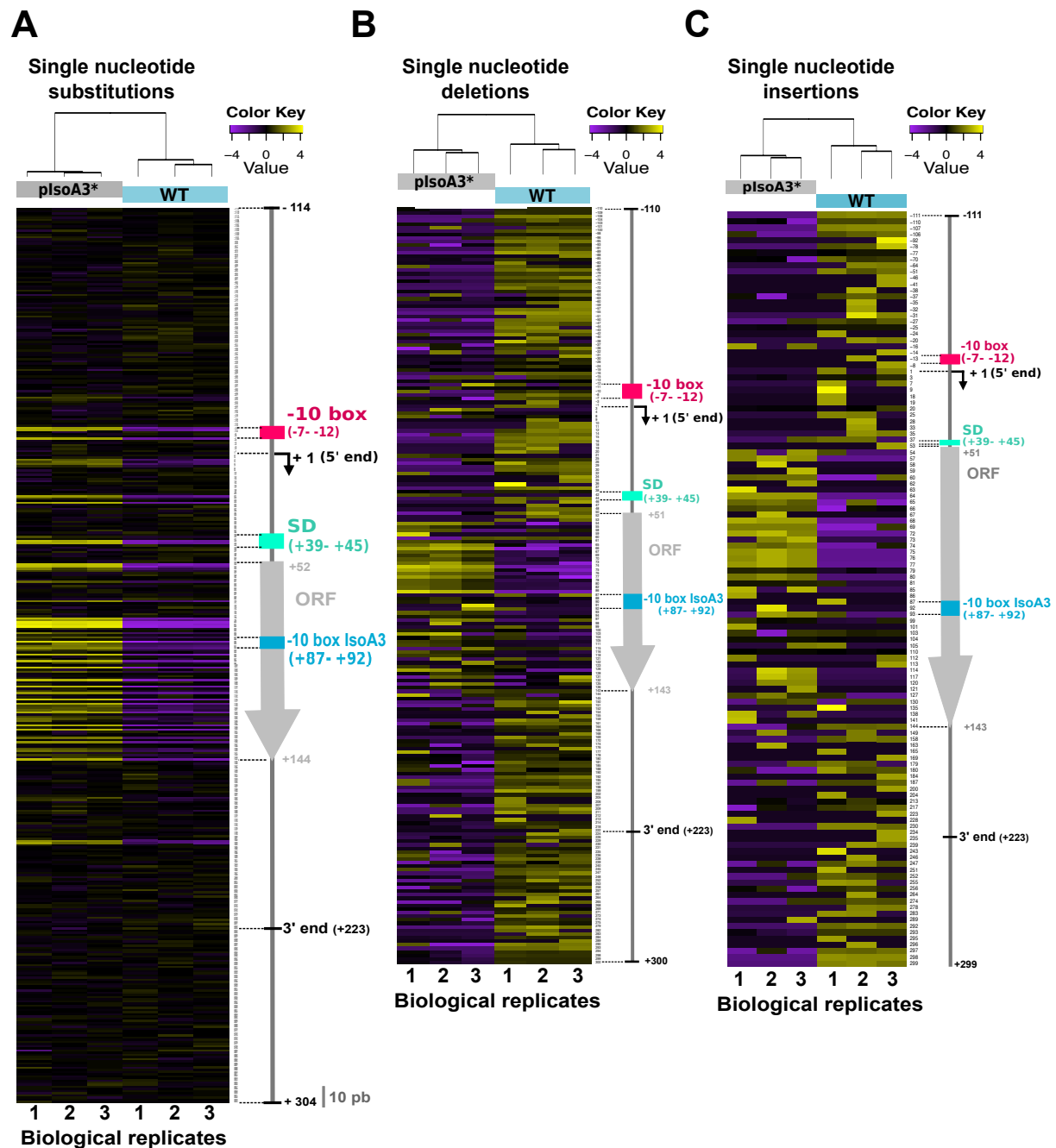
## CHAPTER 8. Appendices

## 8.1. Appendix Paper III

## Supplementary Figures

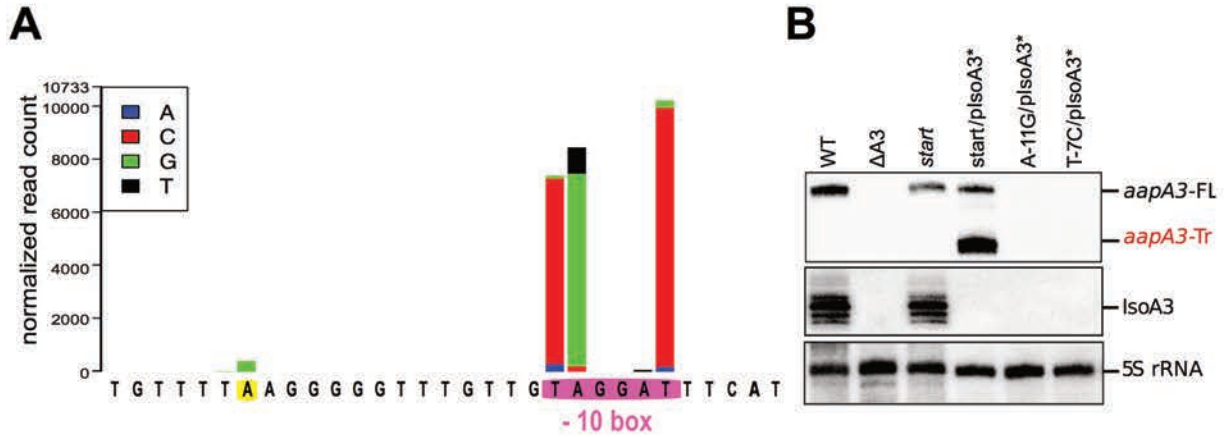


**Figure 8.1.1. Details on *aapA3*/IsoA3 locus deletion and deep-sequencing approaches. (A)** Schematic representation of the oligonucleotides used for the *aapA3*/IsoA3 locus deletion and Illumina paired-end sequencing. Are shown: the upstream and downstream homology regions (HR) used for homologous recombination in all transformation assays; the deleted region and the amplicon used for Illumina paired-end sequencing approach; primers FA407 and FA408 carrying an overhang 5' tail (in pink) containing the 5' and 3' ends, respectively, of the *rpsl-erm* cassette used for locus deletion; primers FA395 and FA396 carrying an overhang 5' tail (in red) containing the adapters for paired-end Illumina sequencing approach (see Table 8.1.4 in the Appendix for oligonucleotides sequences). **(B)** The *aapA3*/IsoA3 module located at the chromosomal locus III of the *H. pylori* 26695 strain. Upstream (415 nt) and downstream (418 nt) regions used for homologous recombination (HR) are highlighted in blue. Deleted region in the *aapA3*/IsoA3 knockout mutant corresponds to the region flanked by the upstream and downstream HR regions. AapA3 -10 box is shown in pink. AapA3 transcription start site (TSS, +1 *aapA3*) determined by RNA-seq analysis is represented with a black arrow. AapA3 full-length transcript is highlighted in bold. 3' end-truncated *aapA3* mRNA is highlighted in orange. AapA3 Shine-Dalgarno sequence (SD) is shown in dark green. AapA3 start (ATG) and stop (TGA) codons are shown in light green and red, respectively. IsoA3 -10 box is highlighted in light blue. IsoA3 transcription start site (TSS, +1 IsoA3) is represented by a light blue arrow.

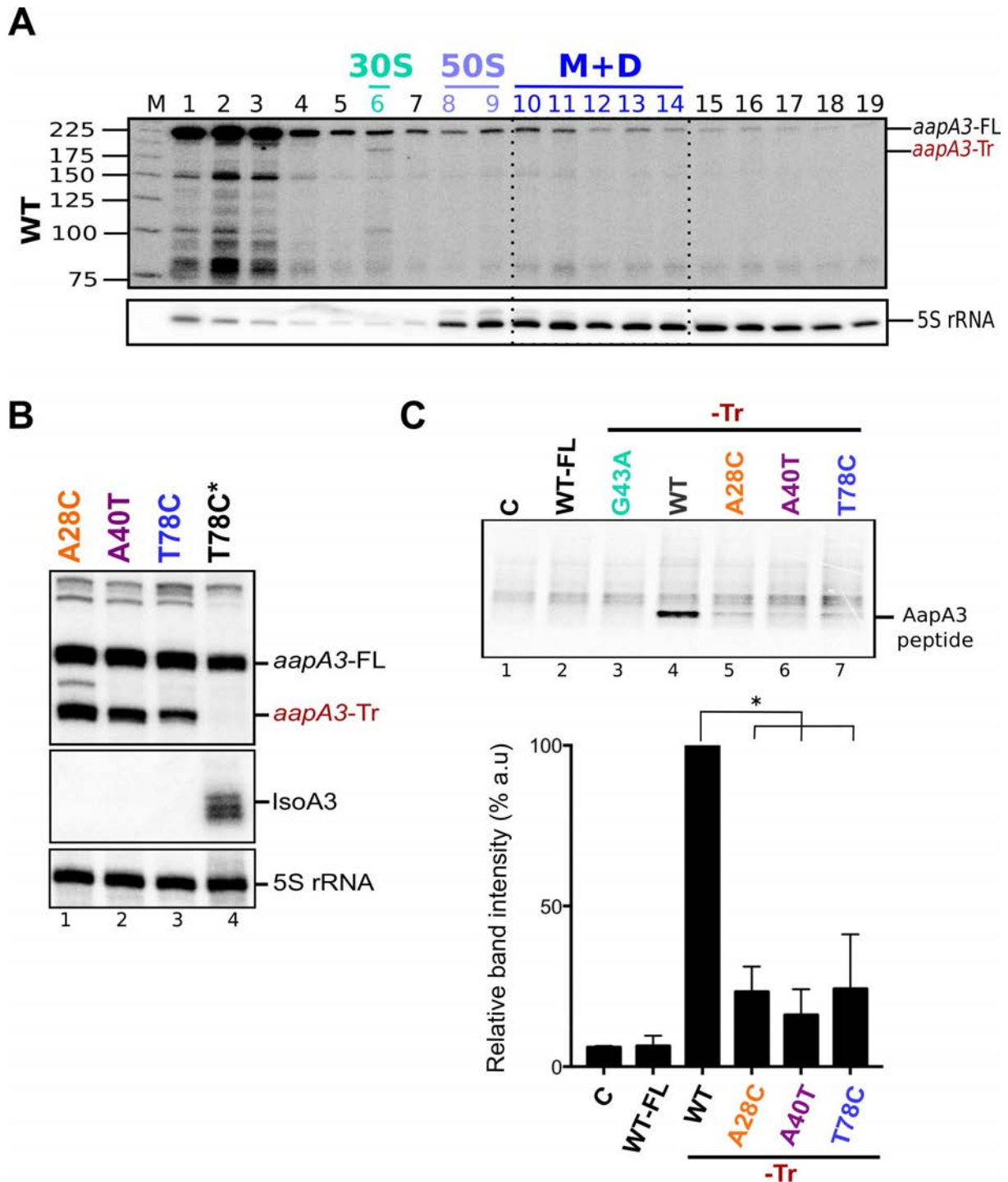


**Figure 8.1.2. Comparison of the distribution and relative frequency of single-nucleotide suppressors in the WT and pIsoA3\* *aapA3*/IsoA3 modules.** Heatmap of the location and relative abundance of single-nucleotide substitutions (A), deletions (B), and insertions (C), in WT (*aapA3*/IsoA3) and pIsoA3\* (*aapA3*/pIsoA3\*) samples coming from Miseq Illumina deep-sequencing. For a given position, all sequences showing a single-nucleotide substitution (A), deletion (B), or insertion (C) at that position were counted irrespective of the nucleotide identity (“positional” analysis). Sequence counts were converted to TMM-normalized median-centered, and log2-transformed FPKM values were hierarchically clustered according to biological replicate sample (see Methods section for details). The color key gives the log2 value scale (negative and positive values represent relative frequencies below and above the median, respectively). Positions are numbered relative to the *aapA3* transcriptional start site (TSS, +1). Colored arrows and boxes on the right-hand side of each heatmap indicate the locations of relevant sequence elements on the *aapA3*/IsoA3 locus: -10 box, *aapA3* promoter -10 box; -10 box IsoA3, IsoA3 promoter -10 box; +1 (5' end), *aapA3* transcription start site; SD, Shine-Dalgarno; ORF, open reading frame; 5' and 3' ends delimit the UTRs (untranslated regions) on the *aapA3* mRNA; the “10 nt” scale bar at the bottom of panel (A) is used to measure intervals of 10 nucleotides alongside the map.



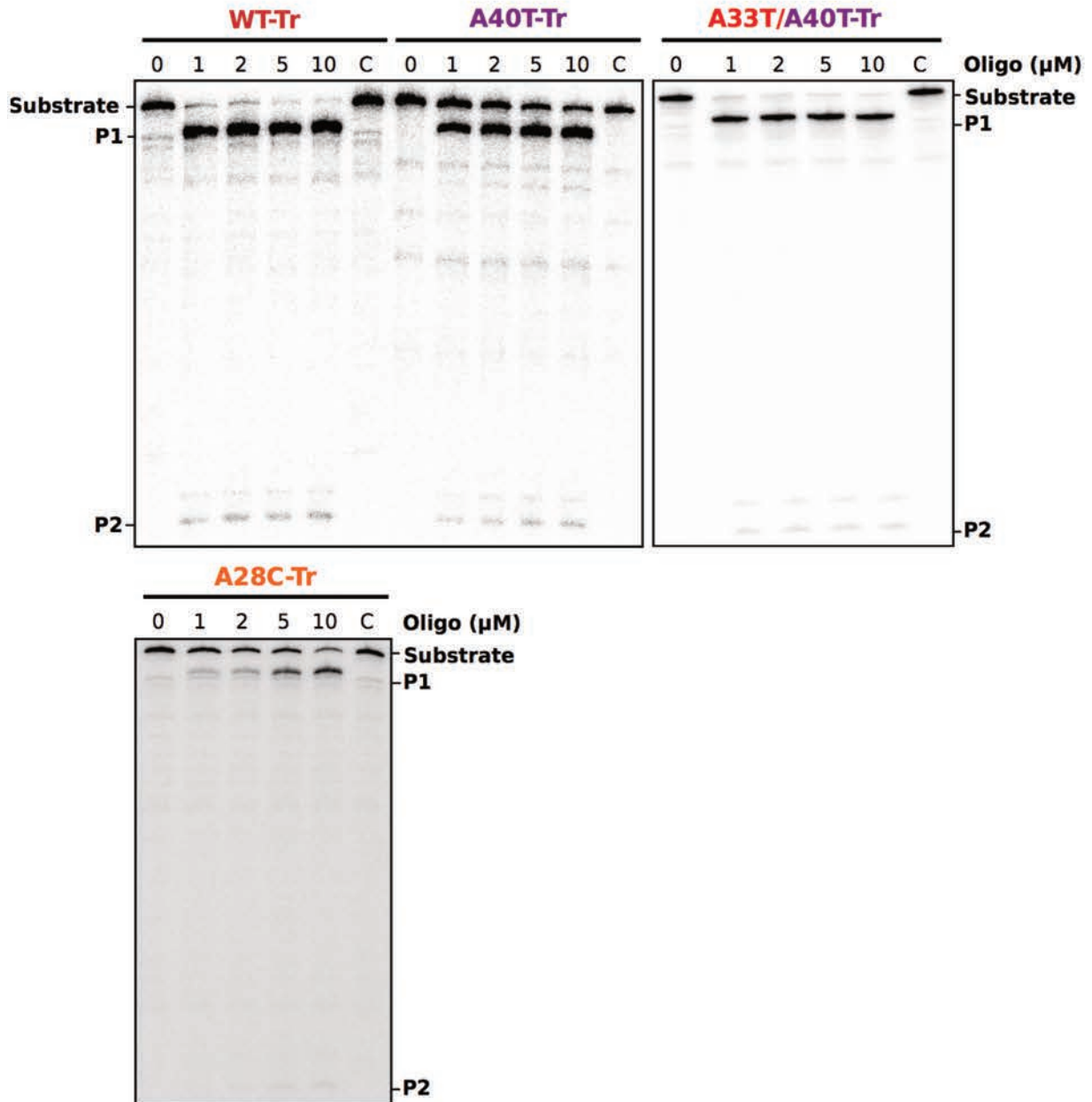


**Figure 8.1.3. Defining and validating AapA3 promoter with nucleotide resolution.** **A)** Statistical analysis of the differential amount of individual single-nucleotide substitution sequences in WT (*aapA3*/IsoA3) and pIsoA3\* (*aapA3*/pIsoA3\*) samples was carried out using DESeq2 (“nucleotide-specific” analysis; see Methods section for details). Nucleotide substitutions that are significantly enriched ( $\text{padj} \leq 0.05$ ) in pIsoA3\* compared to WT sequences in *aapA3* promoter region are shown. The -10 box of the *aapA3* gene promoter is highlighted in pink; an enriched mutation located at position -26 from *aapA3* +1 is highlighted in yellow. **(B)** Total RNA was isolated from the indicated strains and 10  $\mu\text{g}$  were subjected to Northern Blot analysis. The same membrane was successively probed with FD38 labeled oligonucleotide and IsoA3 riboprobe to detect *aapA3* and IsoA3 transcripts, respectively. The different transcripts are annotated as: *aapA3*-FL (full length), *aapA3*-Tr (3’ end truncated), and IsoA3 full-length and processed transcripts. Proper loading was assessed by the level of 5S rRNA using the labeled oligoprobe FD35.



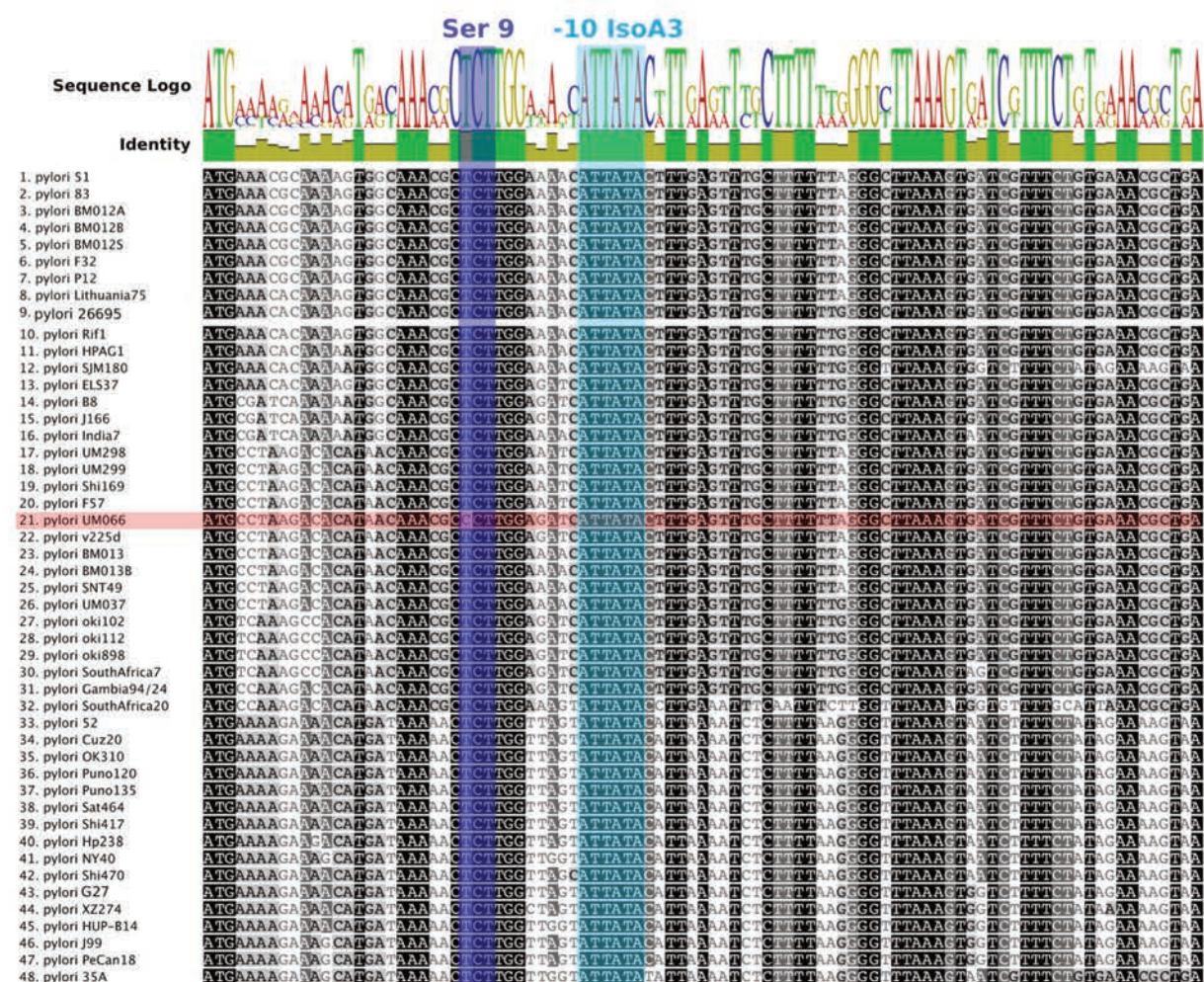
**Figure 8.1.4. The A28C, A40T and T78C suppressors inhibit *aapA3*-Tr mRNA translation.** (A) Only the 3' end truncated *aapA3* mRNA form is translated *in vivo*. Cell lysate of the *H. pylori* 26695 wild-type (WT) strain was subjected to ultracentrifugation through a sucrose gradient under polysome stabilization conditions (+ Chloramphenicol). A profile at OD<sub>254</sub> was recorded. RNA was extracted from each fraction and equal volumes of each extract were subjected to Northern Blot analysis. Fractions corresponding to the free 30S and 50S subunits, 70S ribosomes (free and translating) and polysomes are indicated. The different transcripts are annotated as: *aapA3*-FL (full length, 225 nt), *aapA3*-Tr (3' end truncated, 190 nt), and 5S rRNA as loading control (5S rRNA). M, monosomes; D, disomes. (B) Gene expression analysis of the indicated strains was analyzed by Northern Blot. Transcripts *aapA3*-FL (full length), *aapA3*-Tr (3' end truncated), and IsoA3 full-length and processed transcripts are shown. Proper loading was assessed by the level of 5S rRNA. The T78C\* construct contains a WT IsoA3. (C) Relative peptide production upon *in vitro* translation of *aapA3*-FL, *aapA3*-Tr and the *aapA3*-Tr form of the three independent suppressor

mutants A28C-Tr, A40T-Tr and T78C-Tr (upper panel). A construct with inactivated SD sequence (G43A-Tr) is also included for comparison. Control lane (0) shows the translation background obtained without exogenous mRNA. A representative experiment is shown. Relative peptide production was quantified (lower panel). Error bars represent the s.d; n=3 independent experiments, \* $P<0.0001$  according to unpaired t-test.

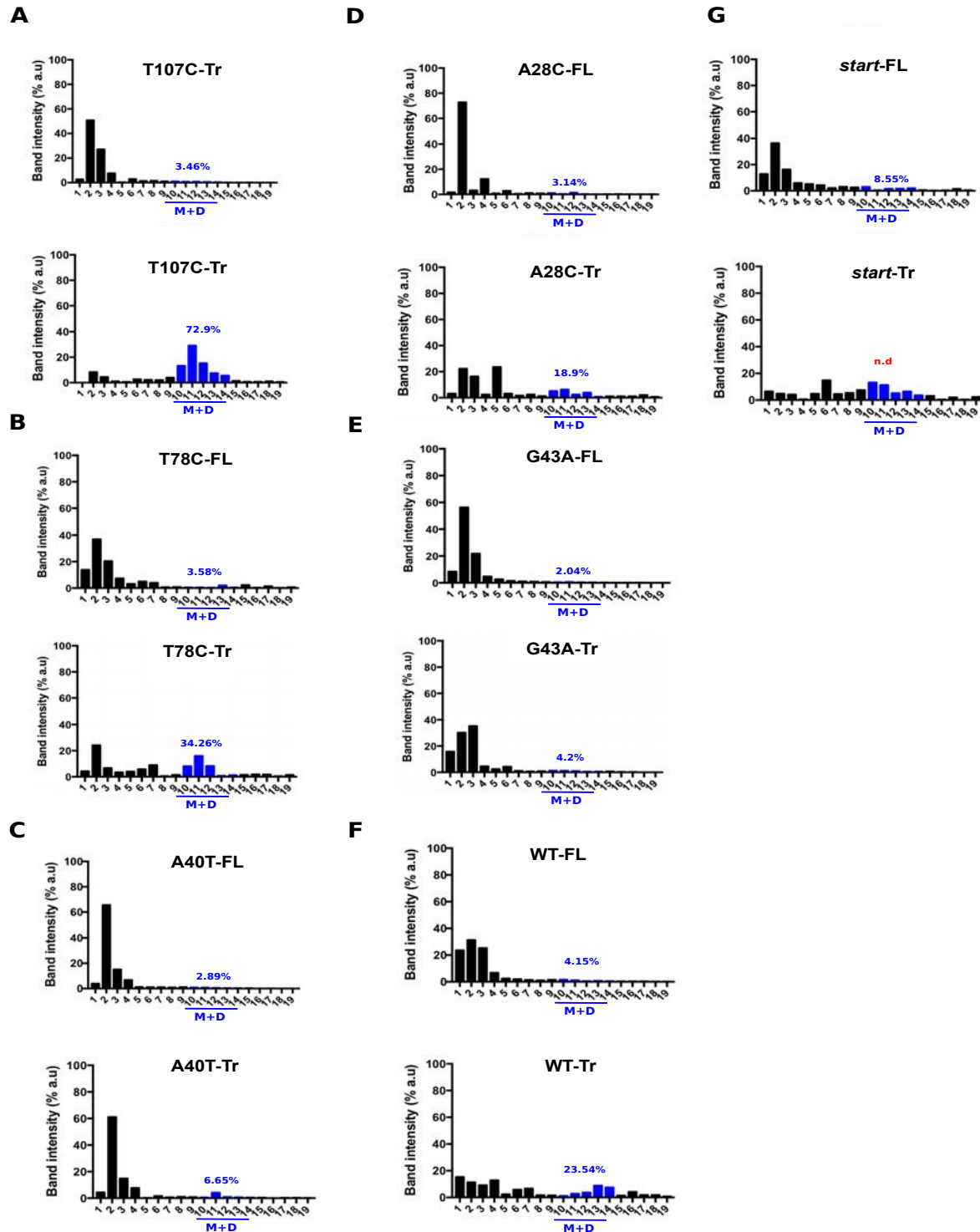


**Figure 8.1.5. Gel analysis of RNase H/ oligonucleotide accessibility assays.** 30 fmol of internally labeled RNA were used. DNA oligonucleotides were used to a final concentration of 0 to 10  $\mu\text{M}$ . Reactions were incubated for 30 min at 30°C in the presence or absence (C, control) of 0.25 U *E. coli* RNase H1. Reactions were stopped with 10  $\mu\text{l}$  of 2X Loading Buffer and products were analyzed in an 8% denaturing PAA gel. See Figure 5B for relative band quantification. P1 and P2 indicate the two RNase H-oligonucleotide-mediated RNA cleavage products.



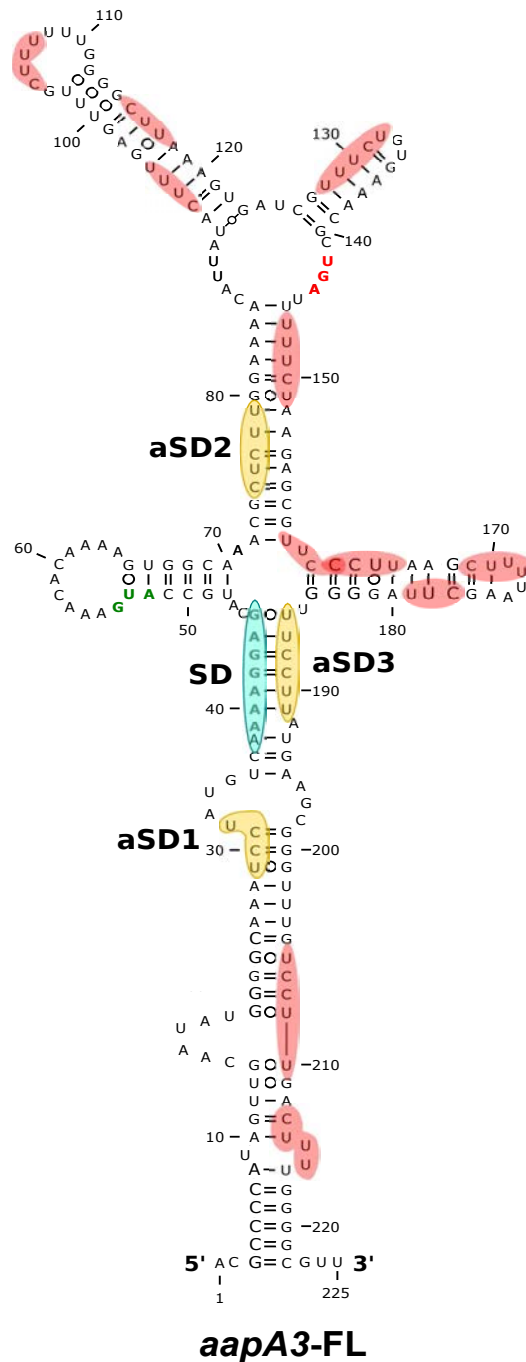


**Figure 8.1.6.** Nucleotide alignment of AapA3 coding region of 49 *Helicobacter pylori* strains. Conserved nucleotides are highlighted in different tones of grey depending on their conservation level. Sequence logo and identity scores are shown on the top. The highly conserved region corresponding to the second SD sequestering-sequence (serine residue at position 9) and the IsoA3 -10 box are highlighted in dark and light blue, respectively. The UM066 strain, only strain containing a CCT Proline codon at position 9, is highlighted in pink. Geneious software 8.1.8 was used for sequence collection and alignment.



**Figure 8.1.7. Quantification of the relative *aapA3* mRNA band intensity from polysome fractionation Northern Blots shown in Figure 4.** Relative band intensity was determined using the ImageLab software. Percentage of band intensity located on 70S fractions was calculated for each *aapA3* transcript (full-length, -FL or truncated, -Tr) by dividing the band intensity in 70S fractions by the total band intensity (addition of intensity values in all fractions). Each panel shows the data for a given WT or mutant transcript. M, monosomes; D, disomes. In the case of the G54T strain, the quantification of the *aapA3*-Tr form present on the M+D fractions was not possible due to a strong mRNA degradation (n.d= non-detectable).





**Figure 8.1.8.** Only three out of the thirteen potential aSD sequenced embedded in *aapA3* mRNA are **functional**. 2D structure predictions were generated with the RNAfold Web Server (Gruber et al., 2008) and VARNA RNA (Darty et al., 2009) was used to perform the drawing. Potential, but not used, anti-Shine-Dalgarno (aSD) sequences (UC-rich motifs) are highlighted in red. The three functional aSD sequences (aSD1, aSD2, aSD3) are shown in gold. Shine-Dalgarno (SD) sequence is shown in turquoise. Translation start and stop codons are shown in green and red, respectively.

## Supplementary Tables

Table 8.1.1. *Helicobacter pylori* strains used in Paper III.

Name	Strain number	Description	Plasmid	Resistance	Reference
B128	B050(H1)	Wild-type B128 strain	none	-	(Farnbacher et al., 2010; McClain et al., 2009)
B128 $\Delta rnc$	H54 (IHP18)	B128 <i>rnc::aphA-3</i>	none	Kan <sup>R</sup>	(Anion et al., 2017)
26695	JR34 (H5)	Wild type 26695 strain, Institut Pasteur collection	none	-	(Tomb et al., 1997b)
26695 <i>rpsL</i> K43R	H158	<i>rpsL</i> gene mutated on the Lys at position 43 to Arg (K43R)	none	Str <sup>R</sup>	this study
26695 $\Delta aapA3$ /IsoA3	H204	$\Delta aapA3$ /IsoA3:: <i>rpsL</i> <sub>CJ</sub> - <i>erm</i> / <i>rpsL</i> K43R	none	Erm <sup>R</sup>	this study
26695 Complemented <i>aapA3/IsoA3</i>	H170	$\Delta aapA3$ /IsoA3 + <i>aapA3/IsoA3</i>	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> start	H171	<i>aapA3</i> start codon mutated to ATT by the single point mutation G54T	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> $\Delta$ T109	H172	<i>aapA3</i> carrying a -1 frameshift mutation (deletion of T at position 109) generating a 23 amino acids-long peptide	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> start/pIsoA3*	H173	<i>aapA3</i> G54T and IsoA3 promoter inactivated by the double point mutation A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> G43A/ pIsoA3*	H247	<i>aapA3</i> SD inactivated by the G43A mutation and IsoA3 promoter A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> T107C/ pIsoA3*	H278	<i>aapA3</i> ORF suppressor T107C (Phe 19 Ser) and IsoA3 promoter A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> A28C/ pIsoA3*	H224	<i>aapA3</i> A28C suppressor mutation and IsoA3 promoter A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> A40T/ pIsoA3*	H225	<i>aapA3</i> A40T and IsoA3 promoter A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> A33T/ A40T/G54T/ pIsoA3*	H257	<i>aapA3</i> A40T/A33T/G54T and IsoA3 promoter A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> T78C/ pIsoA3*	H240	<i>aapA3</i> T78C and IsoA3 promoter A87C/A90G	none	Str <sup>R</sup>	this study
26695 <i>aapA3</i> T78C	H226	<i>aapA3</i> T78C with wild-type IsoA3 expression	none	Str <sup>R</sup>	this study

\* All nucleotide positions are indicated relative to the AapA3 transcriptional start site (TSS, +1).

**Table 8.1.2. Plasmids used in Paper III.**

Name	Description	Origin/ Marker	Reference
pSP60 -2	pSP60 carrying the counter selection cassette <i>rpsL-erm</i>	pSC101* / Amp <sup>R</sup>	(Dailidiene, D. <i>et al.</i> , 2006) (Pernitzsch <i>et al.</i> , 2014)
pA3-Up WT	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus amplified with the FA406/FA386 primer pair	ColE1/ Amp <sup>R</sup>	this study
pA3-Down WT	pGEM-T carrying the downstream fragment of the <i>aapA3</i> /IsoA3 locus amplified with the FA409/FA387 primer pair	ColE1/ Amp <sup>R</sup>	this study
pA3-Down pIsoA3*	pGEM-T carrying the downstream fragment of the <i>aapA3</i> /IsoA3 locus containing IsoA3 -10 box inactivated (A87C/A90G)	ColE1/ Amp <sup>R</sup>	this study
pA3-Up <i>start</i>	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the AapA3 start codon mutation G54T	ColE1/ Amp <sup>R</sup>	this study
pA3-Up A28C	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the suppressor A28C	ColE1/ Amp <sup>R</sup>	this study
pA3-Up A33T	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the A33T mutation	ColE1/ Amp <sup>R</sup>	this study
pA3-Up A40T	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the suppressor A40T	ColE1/ Amp <sup>R</sup>	this study
pA3-Up A40T/A33T	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the A40T and the compensatory mutation A33T	ColE1/ Amp <sup>R</sup>	this study
pA3-Up T78C	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the suppressor T78C	ColE1/ Amp <sup>R</sup>	this study
pA3-Down T78C	pGEM-T carrying the downstream fragment of the <i>aapA3</i> /IsoA3 locus containing the suppressor T78C	ColE1/ Amp <sup>R</sup>	this study
pA3-Down T78C/ pIsoA3*	pGEM-T carrying the downstream fragment of the <i>aapA3</i> /IsoA3 locus containing the suppressor T78C and IsoA3 -10 box inactivated (A87C/A90G)	ColE1/ Amp <sup>R</sup>	this study
pA3-Up G43A	pGEM-T carrying the upstream fragment of the <i>aapA3</i> /IsoA3 locus containing the SD suppressor G43A	ColE1/ Amp <sup>R</sup>	this study

\* All nucleotide positions are indicated relative to the AapA3 transcriptional start site (TSS, +1).

**Table 8.1.3. *Escherichia coli* strains used in Paper III.**

<b>Name</b>	<b>Description/ genotype</b>	<b>Plasmid</b>	<b>Resistance</b>	<b>Reference</b>
TOP10	<i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 deoR recA1 araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>galU galK rpsL endA1 nupG</i>	none	none	Invitrogene
A3-Up WT	TOP10	pA3-Up WT	Amp <sup>R</sup>	this study
A3-Down WT	TOP10	pA3-Do WT	Amp <sup>R</sup>	this study
A3-Down pIsoA3*	TOP10	pA3-Do pIsoA3*	Amp <sup>R</sup>	this study
A3-Up <i>start</i>	TOP10	pA3-Up <i>start</i>	Amp <sup>R</sup>	this study
A3-Up A28C	TOP10	pA3-Up A28C	Amp <sup>R</sup>	this study
A3-Up A33T	TOP10	pA3-Up A33T	Amp <sup>R</sup>	this study
A3-Up A40T	TOP10	pA3-Up A40T	Amp <sup>R</sup>	this study
A3-Up A40T/A33T	TOP10	pA3-Up A40T/A33T	Amp <sup>R</sup>	this study
A3-Up T78C	TOP10	pA3-Up T78C	Amp <sup>R</sup>	this study
A3-Down T78C	TOP10	pA3-Do T78C	Amp <sup>R</sup>	this study
A3-Down T78C/ pIsoA3*	TOP10	pA3-Do T78C/ pIsoA3*	Amp <sup>R</sup>	this study
A3-Up G43A	TOP10	pA3-Up G43A	Amp <sup>R</sup>	this study

\* All nucleotide positions are indicated relative to the AapA3 transcriptional start site (TSS, +1).

**Table 8.1.4. Oligonucleotides used in Paper III.**

	Sequence (5'→3' direction)	Used for
FD11	GAAATTAATACGACTCACTATAGCAAG AGCGTTTGCCACTT	Reverse primers carrying a T7 promoter for IsoA3 amplification for <i>in vitro</i> transcription
FD17	ACGCCCCATAGTTGCAATAT	Forward primer for IsoA3 amplification for <i>in vitro</i> transcription
FD35	TCGGAATGGTTAACTGGGTAGTTCCT	Reverse primer for 5S rRNA mRNA detection by Northern Blot
FD38	GCTCCTTTTGACATAGGATT	Reverse primer for <i>aapA3</i> mRNA detection by Northern Blot
FA110	TGCTTTATAACTATGGATTAAAC	Forward primer for <i>rpsL-erm</i> cassette amplification from pSP60
FA111	TTACTTATTAAATAATTTATAGC	Reverse primer for <i>rpsL-erm</i> cassette amplification from pSP60
FA170	GAAATTAATACGACTCACTATAGGACG CCCCATAGTTGCAATAT	Forward primer carrying a T7 promoter for <i>aapA3</i> <i>in vitro</i> transcription
FA173	AGGAAACCCCTAAGCTTAAAAGC	Reverse primer for <i>aapA3</i> -Tr amplification
FA175	GACCAACGCCCCAAAAGTC	Reverse primer for <i>aapA3</i> full-length amplification
FA281	AGCATGCCATTAAACACAAA	Forward primer for mutagenesis of <i>aapA3</i> 26695 start codon (G54A)
FA282	TTTGTGTTTAATGGCATGCT	Reverse primer for mutagenesis of <i>aapA3</i> 26695 start codon (G54A)
FA283	TGGAAAACCTTGACTTTGAGT	Forward primer for mutagenesis of IsoA3 - 10 box: mutations A87C/A90G
FA284	ACTCAAAGTACAAGGTTTTCCA	Reverse primer for mutagenesis of IsoA3 - 10 box: mutations A87C/A90G
FA386	CCAAGAGCGTTTGCCACTTTTG	Reverse primer for <i>aapA3</i> /IsoA3 locus split cloning in pGEM®T (upstream fragment)
FA387	CACAAAAGTGGCAAACGCTC	Forward primer for <i>aapA3</i> /IsoA3 locus split cloning in pGEM®T (downstream fragment)
FA395	<u>CTTTCCTACACGACGCTCTTCCGATCT</u> CTATCCAATAAAGATAAGC	Forward primer for <i>aapA3</i> 26695 amplification for Illumina paired-end sequencing
FA396	<u>GGAGTTCAGACGTGTGCTCTTCCGATCT</u> GCACTCTATGAGGGGATTTAG	Reverse primer for <i>aapA3</i> 26695 amplification for Illumina paired-end sequencing
FA406	GCATTATAAAATGAAATCC	Forward primer for the amplification of <i>aapA3</i> 26695 fragment Up from <i>hpn</i> -like
FA407	<b>GTTTAATCCATAGTTATAAAGCACAA</b> AAAGAGGGATTTTAAAAG	Reverse primer for the amplification of <i>aapA3</i> 26695 Up fragment to generate the <i>aapA3</i> /IsoA3 locus deletion designed for deep-seq
FA408	<b>GCTATAAATTATTTAATAAGTAACCGC</b> TTGCTCTAGCTTTTTG	Forward for the amplification of <i>aapA3</i> 26695 Down fragment to generate the <i>aapA3</i> /IsoA3 locus deletion designed for deep-seq
FA409	CTAGCCACGCTCTATTAGAG	Reverse for the amplification of <i>aapA3</i> 26695 Down fragment to generate the <i>aapA3</i> /IsoA3 locus deletion designed for deep-seq
FA465	CAATATGGGGCAAcTCCTATGTC	Forward primer for the introduction of the suppressor A28C
FA466	GACATAGGAgtTGCCCCATATTG	Reverse primer for the introduction of the suppressor A28C



FA467	CCTATGTCAAtAGGAGCATG	Forward primer for the introduction of the suppressor A40T
FA468	CATGCTCCTaTTGACATAGG	Reverse primer for the introduction of the suppressor A40T
FA511	CAAAAGTGGCAAACGCTCcTGGAAAACcTTgTACTTTGAGTTTG	Forward primer for the introduction of the suppressor T78C
FA512	GTTTTCCAgGAGCGTTTGCCACTTTTG	Reverse primer for the introduction of the suppressor T78C
FA535	CCTATGTCAAAAGaAGCATGCCATGAAACAC	Forward primer for the introduction of the SD mutation G43A
FA536	GTGTTTCATGGCATGCTtCTTTTGACATAGG	Reverse primer for the introduction of the SD mutation G43A
FA546	GTTGCAATATGGGGCAAATCCTtTGCAAAAGGAGCATGCC	Forward primer for the introduction of A33T mutation (complementation of A40T suppressor)
FA547	GGCATGCTCCTTTTGACAaAGGATTTGCCCATATTGCAAC	Reverse primer for the introduction of A33T mutation (complementation of A40T suppressor)
FA548	GTTGCAATATGGGGCAAATCCTtTGCAAtAGGAGCATGCC	Forward primer for the introduction of A33T mutation in <i>aapA3</i> A40T mutant background
FA549	GGCATGCTCCTaTTGACAaAGGATTTGCCCATATTGCAAC	Reverse primer for the introduction of A33T mutation in <i>aapA3</i> A40T mutant background
FA633	CATGGCATGCTCCTTT	RNaseH/oligonucleotide accessibility assay on WT-FL, WT-Tr and T78C-Tr <i>aapA3</i> mRNAs
FA644	CATAGGATTTGCCCCA	RNaseH/oligonucleotide accessibility assay on A40T-Tr <i>aapA3</i> mRNA
FA651	CAAAGGATTTGCCCCA	RNaseH/oligonucleotide accessibility assay on A33T/A40T-Tr <i>aapA3</i> mRNA
FA652	CATAGGAGTTGCCCCA	RNaseH/oligonucleotide accessibility assay on A28C-Tr <i>aapA3</i> mRNA

\* All nucleotide positions are indicated relative to the AapA3 transcriptional start site (TSS, +1).

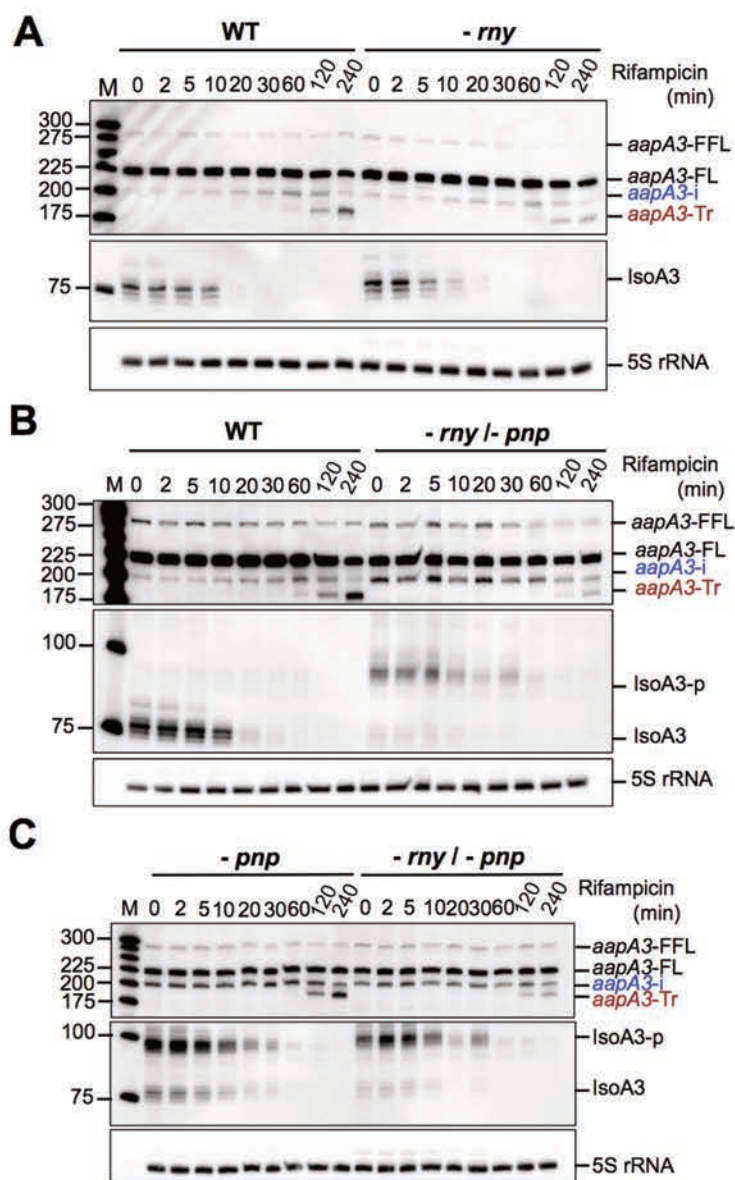
\*\* Sequences highlighted in bold correspond to *rpsL<sub>CJ</sub>-erm* 5'-overhang tails used for assembly PCR during the construction of the *aapA3*/IsoA3 deleted strain.

\*\*\* Underlined sequences correspond to the DNA adaptors used for Illumina paired-end sequencing approach.

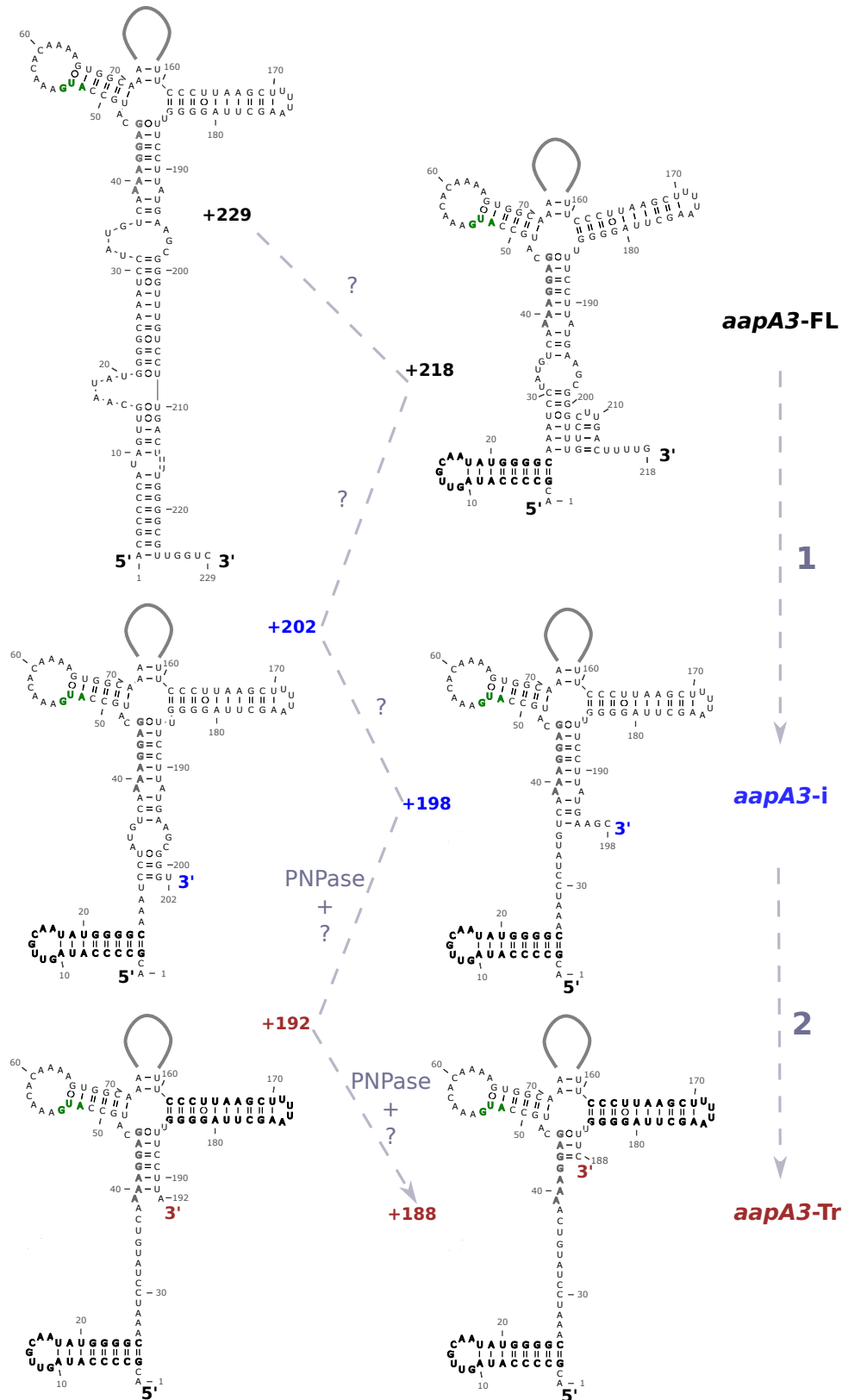
\*\*\*\* Nucleotides in lowercase correspond to the mutations to be introduced by site-directed mutagenesis PCR.

## 8.2. Appendix Paper IV

## Supplementary Figures

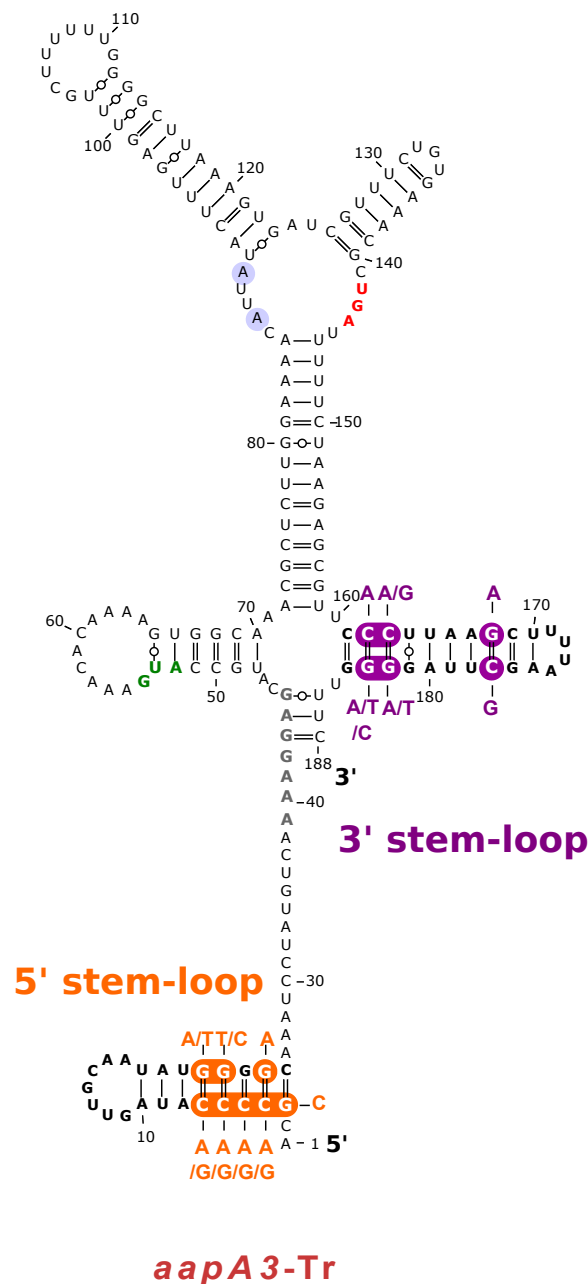


**Figure 8.2.1. PNPase and RNase Y are involved in *aapA3* mRNA and IsoA3 RNA processing.** Total RNA was isolated from the indicated *H. pylori* 26695 strains and time-points upon the addition of rifampicin (80 µg/ ml) once the cultures reached an OD<sub>600</sub> of 1. Northern Blot analysis was performed with 10 µg of total RNA. Membranes were successively probed (after stripping) with <sup>32</sup>P-labeled primers. The oligonucleotide FD38 and IsoA3 riboprobe (<sup>32</sup>P-labeled *aapA3*-Tr) were used to detect *aapA3* mRNA species and IsoA3 RNA, respectively. Proper loading was assessed by the level of the 5S rRNA using the oligonucleotide FD35. A labeled DNA marker (lane M) was used for size estimation. The studied strains were: deletion of RNase Y (-*rny*), PNPase (-*pnp*), double mutant (-*rny* / -*pnp*) and wild type (WT)

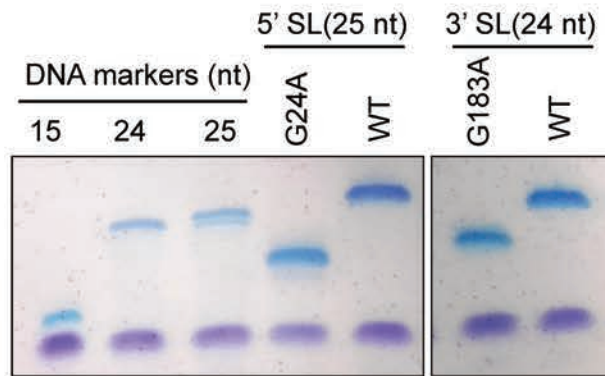


**Figure 8.2.2.** 3' RACE on *aapA3* mRNA revealed a 3' end activation pathway lead 3'-5'-exonucleases and roadblocks. The *aapA3* primary transcript (*aapA3*-FL, 229-218 nt) gets processed by its 3' end until reaching a first roadblock for 3'-exonucleases (double stranded region) leading to the generation of an intermediate species (*aapA3*-i, 202-198 nt). Further 3' end processing takes place until reaching a second roadblock for 3'-exonucleases and generates the translationally active *aapA3* mRNA species (*aapA3*-Tr, 192-188 nt). This species is protected from 5'- and 3'-exonucleases by two highly stable terminal stem-

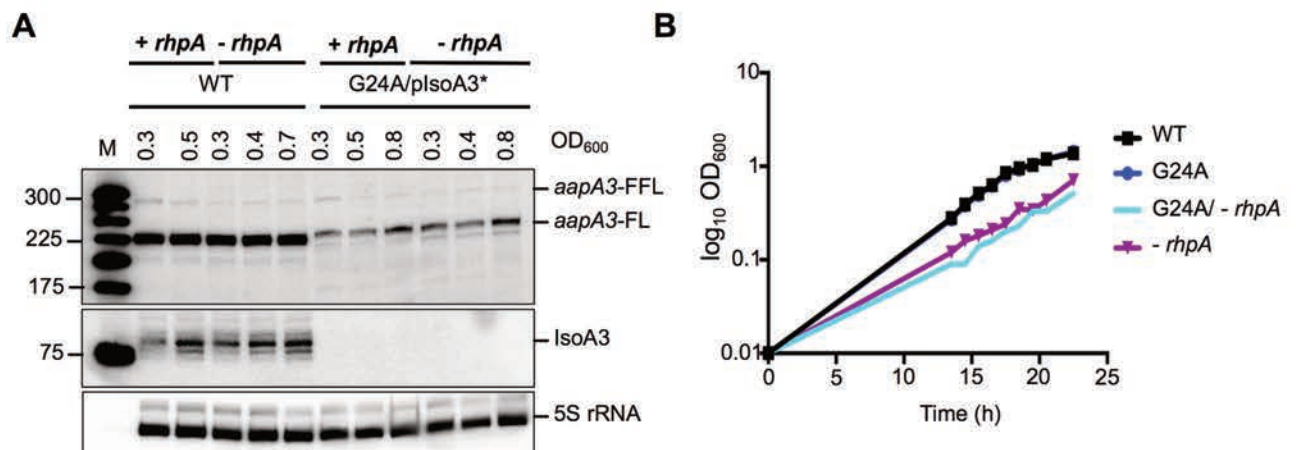
loops (highlighted in bold). 2D structure predictions were generated with the RNAfold Web Server (Gruber et al., 2008) and VARNA RNA (Darty et al., 2009) was used for drawing. Shine-Dalgarno (SD) sequence is highlighted in bold-grey; translation start codons is highlighted in bold-green.



**Figure 8.2.3. Toxicity suppressor mutations cluster symmetrically in the termini stem-loops of the active *aapA3* mRNA.** Single nucleotide substitutions significantly enriched ( $p_{adj} \leq 0.05$ ; DESeq2 ‘nucleotide-specific’ analysis, Masachis *et al.*, in preparation) in absence antitoxin and located close to *aapA3* 5’ and 3’ mRNA ends are shown. *aapA3-Tr* 2D structure predictions were generated with the RNAfold Web Server (Gruber et al., 2008) and VARNA RNA (Darty et al., 2009) was used for drawing. Location of the suppressor mutations in the *aapA3-Tr* mRNA is illustrated. Shine-Dalgarno (SD) sequence is shown in bold-grey; start codon is shown in bold-green; stop codon is shown in bold-red; suppressor mutations located in the 5’-terminal stem-loop are highlighted in orange; suppressor mutations located in the 3’-terminal stem-loop are highlighted in purple. The two single point mutations used to inactivate IsoA3 expression (A87C and A90G) are highlighted in light blue. Most abundant nucleotide substitutions are indicated next to its position.

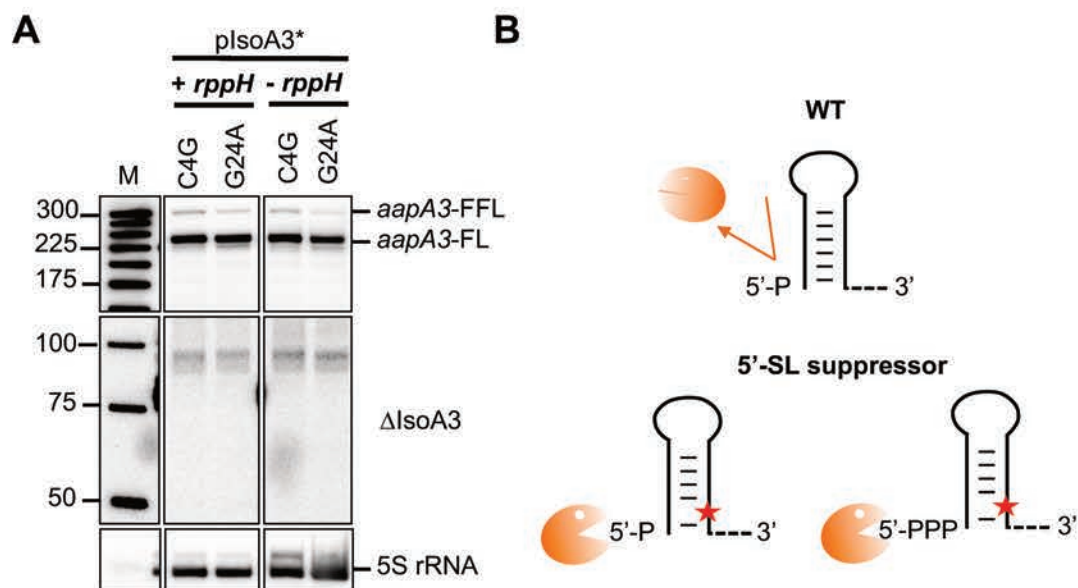


**Figure 8.2.4. Denaturing PAA gel reveals dramatic stability differences of the 5'- and 3'-terminal stem-loops caused by the single nucleotide mutations G24A and G183A.** Chemically synthesized RNAs corresponding to the *aapA3* 5' (25 nt- long)- and 3' (24 nt-long)-terminal stem-loops with wild type or mutated (suppressors G24A and G183A) sequences were run on a 10% PAA containing 7 M urea and visualized by Stains-All. DNA oligonucleotides of 15, 24 and 25 nt-long were used as size markers.

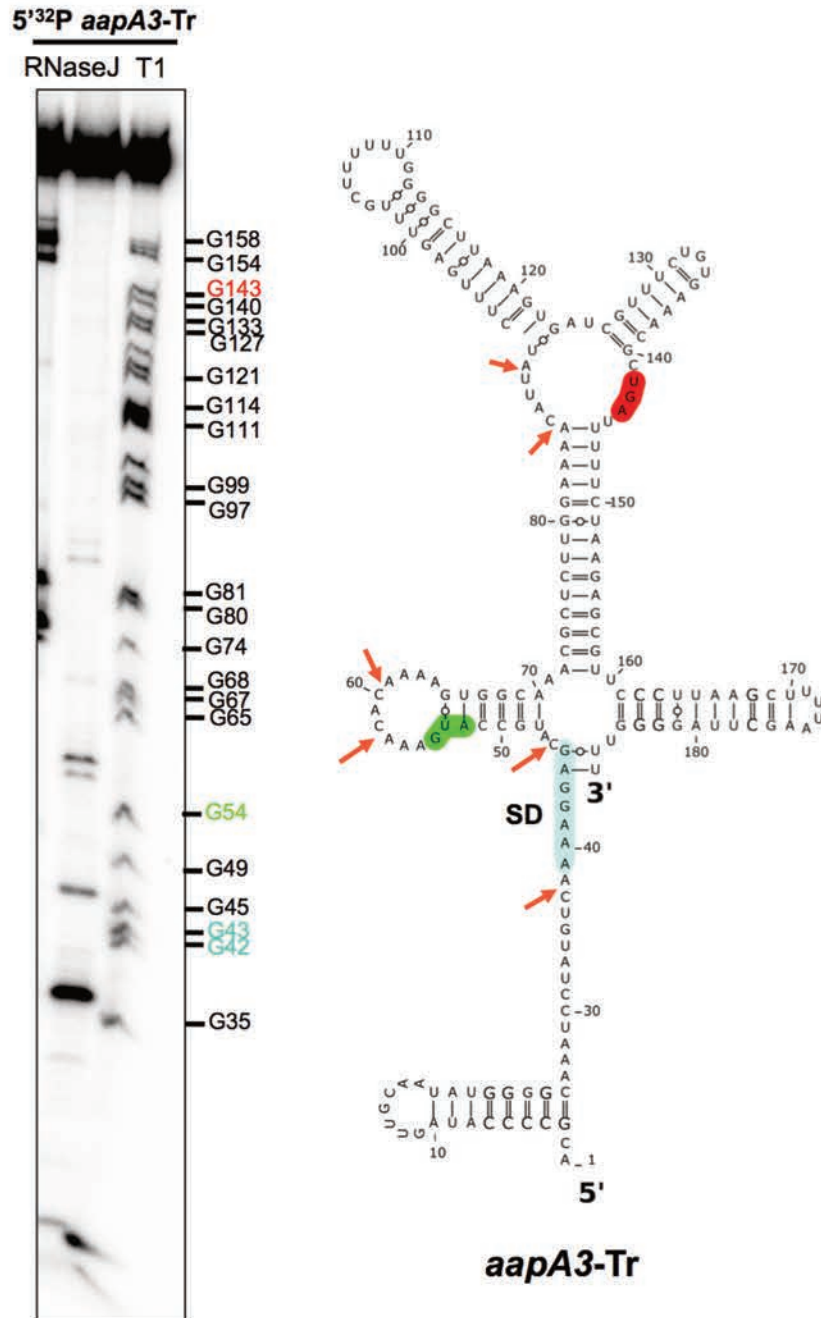


**Figure 8.2.5. The RhpA helicase activity is not required for the *aapA3* G24A-Tr 5'-exonucleolytic degradation.** (A) Northern Blot was performed with 10 µg total RNA isolated from the indicated strains and the specific optical densities (OD<sub>600</sub>). The membrane was successively probed for *aapA3* and IsoA3 RNAs using the <sup>32</sup>P-labeled oligonucleotide FD38 and a riboprobe (<sup>32</sup>P-labeled *aapA3*-Tr), respectively. 5S rRNA level was used to assess proper loading revealed by the oligonucleotide FD35. A labeled DNA marker (lane M) was used for size estimation. (B) The absence of RhpA causes a mild-growth defect. Growth curves of the wild type (WT) or G24A suppressor mutant (G24A, pIsoA3\*) in wild type or deleted *rhpA* background (- *rhpA*) were performed. Data shown are the mean values ± standard deviations of three independent biological replicates.





**Figure 8.2.6. The 5'- exonucleolytic decay of *aapA3* mRNA is independent of the 5' phosphorylation state.** (A) Total RNA was extracted and 10  $\mu$ g was subjected to Northern Blot analysis. The membrane was successively probed with FD38 labeled oligonucleotide and IsoA3 riboprobe to detect *aapA3* and IsoA3 transcripts, respectively. A labeled DNA marker (lane M) was used for size estimation. Proper loading was assessed by the level of 5S rRNA using the labeled oligo probe FD35. Transcripts are annotated as: *aapA3*-FFL (full-full-length,  $\approx$  280 nt), *aapA3*-FL (full-length,  $\approx$  225 nt) and IsoA3 ( $\approx$  80 nt). Same membrane as shown in Figure 3C and D. (B) Schematic representation of the *aapA3*-Tr energetic de-protection due to the single nucleotide suppressors in the 5' stem-loop (5'-SL). In the wild type scenario, the 5'-SL protects the mRNA from degradation even when the 5' end is monophosphorylated. The suppressor mutants dramatically destabilize the stem-loop, leading to the degradation of the *aapA3*-Tr mRNA even in absence of RppH.



**Figure 8.2.7. RNase J acts mainly as endonuclease on free *aapA3-Tr* RNA.** ~0.1 pmol 5'end <sup>32</sup>P-labeled *in vitro* transcribed *aapA3-Tr* was incubated for 10 min at 37°C in the presence of 1 µg Yeast RNA and 28 µM *H. pylori* RNase J in 1X RNase J-endo reaction buffer (20 mM Tris-HCl pH 8.0, 8 mM MgCl<sub>2</sub>, 100 mM NaCl). Reactions were stopped by the addition of 10 µl 2X Loading Buffer II (95% formamide, 18mM EDTA, Xylene Blue and Bromophenol Blue). Cleavage products were run on a 6% denaturing PAA gel and analysed using a Pharos FX phosphorimager. 2D structure predictions and drawing were generated with the RNAfold Web Server (Gruber et al., 2008) and VARNA RNA (Darty et al., 2009), respectively. The Shine-Dalgarno (SD) sequence is highlighted in turquoise; translation start and stop are codons highlighted in green and red, respectively.

## Supplementary Tables

**Table 8.2.1. Minimum Free Energy and padj of the here studied single nucleotide substitutions in the *aapA3* 5'-terminal stem-loop.** Minimum Free Energy values (kcal/mol) were calculated using the RNAfold Web Server (Gruber et al., 2008). Padj values were obtained by nucleotide specific Deseq2 analysis of single nucleotide substitutions appearing in wild type and pIsoA3\* samples in our previously generated Illumina sequencing datasets (Masachis et al., in preparation, see Paper III in Chapter 4 section 4.2).

5'SL (+25)	Minimum Free Energy (Kcal/mol)	Padj
WT	-13.30	-
G3C	-10.30	0.0039 (*)
C4G	-7.70	$2.75 \times 10^{-36}$
G24A	-7.80	$6.87 \times 10^{-150}$
C4T	-10.60	0.64 (Toxic)

**Table 8.2.2. Minimum Free Energy and padj of the here studied tested single nucleotide substitutions in the *aapA3* 3'-terminal stem-loop.** Minimum Free Energy values (kcal/mol) were calculated using the RNAfold Web Server (Gruber et al., 2008). Padj values were obtained by nucleotide specific Deseq2 analysis of single nucleotide substitutions appearing in wild type and pIsoA3\* samples in our previously generated Illumina sequencing datasets (Masachis et al., in preparation, see Paper III in Chapter 4 section 4.2).

3'SL (+16-+184)	Minimum Free Energy (Kcal/mol)	padj
WT	-13.70	-
G183A	-8.60	$2.24 \times 10^{-40}$
G168A	-7.00	0.0334 (Toxic)
G176A	-9.20	0.2083 (Toxic)

**Table 8.2.3. *Helicobacter pylori* strains used in Paper IV**

Name	Strain number	Description	Plasmid	Resistance	Reference
26695	JR34(H5)	Wild-type strain, Institut Pasteur collection	none	-	(Tomb et al., 1997)
26695 <i>rpsL</i> K43R	H158	<i>rpsL</i> gene coding for the ribosomal protein S12 mutated on the amino acid K43 to R conferring resistance to streptomycin	none	Str <sup>R</sup>	This study
26695 $\Delta aapA3/\Delta IsoA3::rpsL-erm$	H204	Strain deleted for the <i>aapA3</i> /IsoA3 locus using the <i>rpsL-erm</i> cassette (in <i>rpsL</i> K43R background)	none	Erm <sup>R</sup>	This study
26695 pAapA3*/IsoA3	H214	<i>aapA3</i> -10 box inactivated by the single nucleotide mutation T-7C (TAGGAT to TAGGAC) (in <i>rpsL</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3</i> start	H171	<i>aapA3</i> start codon mutated by the single nucleotide substitution G54T (from <i>aapA3</i> +1) (in <i>rpsL</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3</i> start/pIsoA3*	H173	<i>aapA3</i> start codon mutated by G54T, and IsoA3 promoter inactivated (double mutation; A87C/A90G from <i>aapA3</i> +1) (in <i>rpsL</i> K43R background)	none	Str <sup>R</sup>	This study
26695 $\Delta rhpA::aphA3$	H287	Strain deleted for the <i>rhpA</i> gene (- <i>rhpA</i> ) using the <i>aphA3</i> ORF	none	Kan <sup>R</sup>	This study
26695 $\Delta rppH::rpsL-erm$	H237	Strain deleted for the <i>rppH</i> gene (- <i>rppH</i> ) using the <i>rpsL-erm</i> cassette (in <i>rpsL</i> K43R background)	none	Erm <sup>R</sup>	This study
26695 $\Delta pnp::rpsL-erm$	H160	Strain deleted for the <i>pnp</i> gene (- <i>pnp</i> ) using the <i>rpsL-erm</i> cassette (in <i>rpsL</i> K43R background)	none	Erm <sup>R</sup>	This study
26695 $\Delta pnp$	H161	Strain deleted for the <i>pnp</i> gene (- <i>pnp</i> ) in <i>rpsL</i> K43R background ( <i>rpsL-erm</i> cassette excised)	none	Str <sup>R</sup>	This study
26695 $\Delta rny::aphA3$	H46	Strain deleted for the <i>rny</i> gene (- <i>rny</i> ) using the <i>aphA3</i> ORF	none	Kan <sup>R</sup>	This study
26695 $\Delta rny::aphA3/\Delta pnp::catCG/$	H213	Strain where <i>rny</i> gene is replaced by the <i>aphA3</i> ORF and the <i>pnp</i> gene by the <i>catCG</i> ORF	none	Kan <sup>R</sup> Cm <sup>R</sup>	This study
26695 $\Delta pnp/\Delta aapA3/\Delta IsoA3::rpsL-erm$	H206	Strain deleted for the <i>pnp</i> gene and the <i>aapA3</i> /IsoA3 locus using the <i>rpsL::erm</i> cassette (in <i>rpsL</i> K43R background)	none	Erm <sup>R</sup>	This study
26695 $\Delta pnp/aapA3$ start	H207	$\Delta pnp/aapA3$ start codon mutated (start, G54T) (in <i>rpsL</i> K43R background)	none	Str <sup>R</sup>	This study
26695 $\Delta pnp/aapA3$ start/pIsoA3*	H208	$\Delta pnp/aapA3$ start codon mutated (start, G54T) and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsL</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3</i> G183A	H209	Suppressor strain carrying the single nucleotide substitution	none	Str <sup>R</sup>	This study

		G182A (in <i>rpsl</i> K43R background)			
26695 <i>Δpnp/aapA3 start</i> G183A/pIsoA3*	H210	<i>Δpnp/aapA3</i> start codon mutated (start, G54T), suppressor mutation G182A and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsl</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3 start</i> G183A/ pIsoA3*	H211	<i>aapA3</i> start codon mutated (start, G54T), suppressor mutation G182A and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsl</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3</i> G24A/pIsoA3*	H212	Suppressor mutation G24A and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsl</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3</i> C4G/pIsoA3*	H241	Suppressor mutation C4G and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsl</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3 start</i> C4T G24A /pIsoA3*	H242	Suppressor mutation G24A, compensatory mutation C4T and start codon mutated (start, G54T) and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsl</i> K43R background)	none	Str <sup>R</sup>	This study
26695 <i>aapA3</i> G3C/pIsoA3*	H259	Suppressor mutation G3C and IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1) (in <i>rpsl</i> K43R background)	none	Str <sup>R</sup>	This study



**Table 8.2.4. *Escherichia coli* strains used in Paper IV.**

<b>Name</b>	<b>Description</b>	<b>Plasmid</b>	<b>Resistance</b>	<b>Reference</b>
TOP10	<i>E. coli</i> strain used for genetic manipulation of plasmids	none	none	Invitrogene
A3-Up WT	TOP10	pGEMT-A3-Up WT	Amp <sup>R</sup>	This study
A3-Down WT	TOP10	pGEMT-A3-Down WT	Amp <sup>R</sup>	This study
A3-Up <i>start</i>	TOP10	pGEMT-A3-Up <i>start</i>	Amp <sup>R</sup>	This study
A3-Down pIsoA3*	TOP10	pGEMT-A3-Down pIsoA3*	Amp <sup>R</sup>	This study
A3-Up G3C	TOP10	pGEMT-A3-Up G3C	Amp <sup>R</sup>	This study
A3-Up C4G	TOP10	pGEMT-A3-Up C4G	Amp <sup>R</sup>	This study
A3-Up C4T	TOP10	pGEMT-A3-Up C4T	Amp <sup>R</sup>	This study
A3-Up G24A	TOP10	pGEMT-A3-Up G24A	Amp <sup>R</sup>	This study
A3-Up G168A	TOP10	pGEMT-A3-Up G168A	Amp <sup>R</sup>	This study
A3-Up G176A	TOP10	pGEMT-A3-Up G176A	Amp <sup>R</sup>	This study
A3-Up G183A	TOP10	pGEMT-A3-Up G183A	Amp <sup>R</sup>	This study

**Table 8.2.5. Plasmids used in Paper IV.**

Name	Description	Resistance	Reference
pSP60	pSP60 carrying the cassette <i>rpsL-erm</i>	Amp <sup>R</sup>	(Dailidienė, D. et al., 2006) (Pernitzsch et al., 2014)
pUC18K2	pUC18 carrying the <i>aphA3</i> gene	Kan <sup>R</sup>	(Skouloubris et al., 1998)
pILL2150	carrying the <i>catCG</i> gene	Cm <sup>R</sup>	(Boneca et al., 2008)
pA3-Up WT	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus (see Figure S2 for upstream references)	Amp <sup>R</sup>	This study
pA3-Up <i>start</i>	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus mutated in the AapA3 start codon (G54A from AapA3 TSS) (see Figure S2 for upstream references)	Amp <sup>R</sup>	(This study)
pA3-Down WT	pGEM-T carrying the downstream fragment of <i>aapA3</i> /IsoA3 locus (see Figure S2 for upstream references)	Amp <sup>R</sup>	This study
pA3-Down pIsoA3*	pGEM-T carrying the downstream fragment of <i>aapA3</i> /IsoA3 locus (see Figure S2 for upstream references) with the IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1)	Amp <sup>R</sup>	This study
pA3-Up G3C	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor G3C from <i>aapA3</i> +1	Amp <sup>R</sup>	This study
pA3-Up C4G	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor C4G from <i>aapA3</i> +1	Amp <sup>R</sup>	This study
pA3-Up C4T	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor C4T from <i>aapA3</i> +1	Amp <sup>R</sup>	This study
pA3-Up G24A	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor G24A from <i>aapA3</i> +1	Amp <sup>R</sup>	This study
pA3-Up C4T/G24A	pGEM-T carrying the upstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor C4T/G24A from <i>aapA3</i> +1	Amp <sup>R</sup>	This study
pA3-Down G168A/ pIsoA3*	pGEM-T carrying the downstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor G168A from <i>aapA3</i> +1 and the IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1)	Amp <sup>R</sup>	This study
pA3-Down G176A/ pIsoA3*	pGEM-T carrying the downstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor G176A from <i>aapA3</i> +1 and the IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1)	Amp <sup>R</sup>	This study
pA3- Down G183A/ pIsoA3*	pGEM-T carrying the downstream fragment of <i>aapA3</i> /IsoA3 locus containing the single nucleotide suppressor G182A from <i>aapA3</i> +1 and the IsoA3 -10 box inactivated (A87C/A90G from <i>aapA3</i> +1)	Amp <sup>R</sup>	This study

**Table 8.2.6. Oligonucleotides used in Paper IV.**

Name	Sequence 5'→3'	Description
FD70	GTACCCGGGTGACTAACTAGG	Forward primer for the amplification of the <i>aphA3</i> gene from pUC18K2
FD255	CAGGTACTAAAACAATTCATCC	Reverse primer for the amplification of the <i>aphA3</i> gene from pUC18K2
FA75	ATGGAGAAAAAATCACTGGA	Forward primer for the amplification of the <i>catCG</i> gene from pILL2150
FD271	GAGAAAAAATCACTGGATATAC	Reverse primer for the amplification of the <i>catCG</i> gene from pILL2150
FD701	<b>TGAGTGGCAGGGCGGGGCGTAACATGGA</b> CAAGTGTCAGTGGT	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene for gene deletion using the <i>catCG</i> cassette
FD700	<b>TATATCCAGTGATTTTTTCTCCATATTTA</b> ATGTTCTTTTTT	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene for gene deletion using the <i>catCG</i> cassette
FA562	GATTTAATATCATCTGCAGAG	Forward primer for the amplification of <i>H. pylori</i> 26695 <i>rhpA</i> gene 'Do' fragment
FA608	GTCAAACCCCTACTATCCATCA	Reverse primer for the amplification of <i>H. pylori</i> 26695 <i>rhpA</i> gene 'Up' fragment
FA610	<b>CCGGTGATATTCTCATTTTAGCCATGGGA</b> GATTCATACCTC	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>rhpA</i> gene 'Down' fragment for gene deletion using the <i>aphA3</i> cassette
FA611	<b>TTTTACTGGATGAATTGTTTTAGAATATTT</b> AAAAAGGAAATTCATG	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>rhpA</i> gene 'Up' fragment for gene deletion using the <i>aphA3</i> cassette
FA484	GAATTGAAAGCGCAATAGCCG	Forward primer for the amplification of <i>H. pylori</i> 26695 <i>rppH</i> gene 'Up' fragment
FA485	<b>GTTTAATCCATAGTTATAAAGCAGTATAC</b> TAGCTTTAAGAATGCC	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>rppH</i> gene 'Up' fragment for gene deletion using the <i>rpsl-erm</i> cassette
FA486	<b>GCTATAAATTATTTAATAAGTAAGATTTG</b> AACAAGCACACGCC	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>rppH</i> gene 'Down' fragment for gene deletion using the <i>rpsl-erm</i> cassette
FA487	CAATGGATTGGATCACGGCG	Reverse primer for the amplification of <i>H. pylori</i> 26695 <i>rppH</i> gene 'Downstream' fragment
FD699	GCTCTGTATTTAGCTCACGCT	Forward primer for the amplification of <i>H. pylori</i> 26695 <i>pnp</i> gene 'Up' fragment
FA109	<b>GTTTAATCCATAGTTATAAAGCAATTTAA</b> TGTTCTTTTTTAA	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene 'Up' fragment for gene deletion using the <i>rpsl-erm</i> cassette

FA112	<b>GCTATAAATTATTTAATAAGTAAACAAGT</b> <b>GTCAAGTGGTTTTAAAAG</b>	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene 'Down' fragment for gene deletion using the <i>rpsl-erm</i> cassette
FA113	<b>CTTTTAAACCACCTTGACACTTGTATTTA</b> <b>ATGTTTCCTTTTTTAAA</b>	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene 'Up' fragment for <i>rpsl-erm</i> cassette excision
FA114	<b>ACAAGTGTCAAGTGGTTTTAAAAG</b>	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene 'Do' fragment for <i>rpsl-erm</i> cassette excision
FD702	<b>ACCAATTCCACCACCAGTGCT</b>	Reverse primer for the amplification of <i>H. pylori</i> 26695 <i>pnp</i> gene 'Downstream' fragment
FD700	<b>TATATCCAGTGATTTTTTTCTCCATATTTA</b> <b>ATGTTTCCTTTTTT</b>	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene 'Up' fragment for gene deletion using the <i>catCG</i> cassette
FD701	<b>TGAGTGGCAGGGCGGGGCGTAACATGGA</b> <b>CAAGTGTCAAGTGGT</b>	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>pnp</i> gene 'Down' fragment for gene deletion using the <i>catCG</i> cassette
FD631	<b>CGTGTGGTTGCCTAAAAGC</b>	Forward primer for the amplification of <i>H. pylori</i> 26695 <i>rny</i> gene 'Up' fragment
FD632	<b>CTAGTTAGTCACCCGGGTACCCCTCTAGC</b> <b>GTAATAGATC</b>	Reverse primer the amplification of the <i>H. pylori</i> 26695 <i>rny</i> gene 'Up' fragment for gene deletion using the <i>aphA3</i> cassette
FD633	<b>GGATGAATTGTTTTAGTACCTGCAAGCTT</b> <b>TAGAAGAGATCGCG</b>	Forward primer the amplification of the <i>H. pylori</i> 26695 <i>rny</i> gene 'Down' fragment for gene deletion using the <i>aphA3</i> cassette
FD634	<b>TGTGGGCCTCTCTCAAGC</b>	Reverse primer for the amplification of <i>H. pylori</i> 26695 <i>rny</i> gene 'Downstream' fragment
FD11	<b>GAAATTAATACGACTCACTATAGCAAGAGC</b> <b>GTTTGCCACTT</b>	Reverse primer carrying a T7 promoter for amplification for IsoA3 <i>in vitro</i> transcription
FD17	<b>ACGCCCCATAGTTGCAATAT</b>	Forward primer for IsoA3 amplification for <i>in vitro</i> transcription
FD35	<b>TCGGAATGGTTAACTGGGTAGTTCCT</b>	Reverse primer for <i>H. pylori</i> 5S rRNA mRNA detection by Northern Blot
FD211	<b>TACGGCGTTTCACTTCTGAGTT</b>	Reverse primer for <i>E. coli</i> 5S rRNA mRNA detection by Northern Blot
FD38	<b>GCTCCTTTTGACATAGGATT</b>	Reverse primer for <i>aapA3</i> mRNA detection by Northern Blot
FA110	<b>TGCTTTATAACTATGGATTAAAC</b>	Forward primer for <i>rpsl-erm</i> cassette amplification from pSP60
FA111	<b>TTACTTATTAAATAATTTATAGC</b>	Reverse primer for <i>rpsl-erm</i> cassette amplification from pSP60
FA170	<b>GAAATTAATACGACTCACTATAGGACGCC</b> <b>CATAGTTGCAATAT</b>	Forward primer carrying a T7 promoter for <i>aapA3</i> <i>in vitro</i> transcription
FA173	<b>AGGAAACCCCTAAGCTTAAAAGC</b>	Reverse primer for <i>aapA3</i> truncated amplification for <i>in vitro</i>

		transcription
FA175	GACCAACGCCCCAAAAGTC	Reverse primer for <i>aapA3</i> full-length amplification for <i>in vitro</i> transcription
FA622	GGGGAGAATTATTCTATCCCCTAAAGAAA ATAAGTTTAC	Reverse primer for <i>aapA3</i> full-full-length amplification for <i>in vitro</i> transcription
FA281	AGCATGCCATTAAACACAAA	Forward primer for mutagenesis of <i>aapA3</i> 26695 start codon (G54A)
FA282	TTTGTGTTTAATGGCATGCT	Reverse primer for mutagenesis of <i>aapA3</i> 26695 start codon (G54A)
FA283	TGGAAAACCTTGACTTTGAGT	Forward primer for mutagenesis of IsoA3 -10 box: mutations A87C/A90G
FA284	ACTCAAAGTACAAGGTTTTCCA	Reverse primer for mutagenesis of IsoA3 -10 box: mutations A87C/A90G
FA386	CCAAGAGCGTTTGCCACTTTTG	Reverse primer for <i>aapA3</i> /IsoA3 locus split cloning in pGEM®T (upstream fragment)
FA387	CACAAAAGTGGCAAACGCTC	Forward primer for <i>aapA3</i> /IsoA3 locus split cloning in pGEM®T (downstream fragment)
FA406	GCATTATAAAATGAAATCC	Forward primer for the amplification of <i>aapA3</i> 26695 fragment 'Up' from <i>hpn</i> -like
FA407	<b>GTTTAATCCATAGTTATAAAGCACAAAAA</b> GAGGGATTTTAAAAG	Reverse primer for the amplification of <i>aapA3</i> 26695 Up fragment to generate the <i>aapA3</i> /IsoA3 locus deletion designed for deep-seq
FA408	<b>GCTATAAATTATTTAATAAGTAACCGCTT</b> GCTCTAGCTTTTTG	Forward for the amplification of <i>aapA3</i> 26695 'Down' fragment to generate the <i>aapA3</i> /IsoA3 locus deletion
FA409	CTAGCCACGCTCTATTAGAG	Reverse for the amplification of <i>aapA3</i> 26695 'Down' fragment to generate the <i>aapA3</i> /IsoA3 locus deletion
FA461	GATTTTCATCACGtCCCATAGTTG	Forward primer for C4T site-directed mutagenesis PCR
FA462	CAACTATGGGgCGTGATGAAATC	Reverse primer for C4T site-directed mutagenesis PCR
FA463	GCAATATGGGgCAAATCCTATG	Forward primer for G24A site-directed mutagenesis PCR
FA464	CATAGGATTTGtCCCATATTGC	Reverse primer for G24A site-directed mutagenesis PCR
FA515	GATTTTCATCACGgCCCATAGTTG	Forward primer for C4G site-directed mutagenesis PCR
FA516	CAACTATGGGcCGTGATGAAATC	Reverse primer for C4G site-directed mutagenesis PCR
FA540	GTTGTAGGATTTTCATCACcCCCATAGTTGC AATATGGGGC	Forward primer for G3C site-directed mutagenesis PCR
FA541	GCCCCATATTGCAACTATGGGGgGTGATGA AATCCTACAAC	Reverse primer for G3C site-directed mutagenesis PCR
FA559	CTAAGAGCGTTCCTTAAaCTTTTAAGCTTA GGGGTTCC	Forward primer for G168A site-directed mutagenesis PCR
FA560	GGAAACCCCTAAGCTTAAAAGtTAAAGGA ACGCTCTAG	Reverse primer for G168A site-directed mutagenesis PCR
FA471	GCTTTTAAaCTTAGGGGTTTCC	Forward primer for G176A site-



		directed mutagenesis PCR
FA472	GGAAACCCCTAAGtTTAAAAGC	Reverse primer for G176A site-directed mutagenesis PCR
FA609	[PHO]NNNNNNNTTCACTGTTCTTAGCGGCCG CATGCTC[AmC6]	3' DNA adapter for 3'RACE PCR (5'P and 3'NH <sub>2</sub> in C6)
FA624	TGCTCCTTTTGACATAGGATTTG	Reverse primer for <i>aapA3</i> 5'-terminal 47 nt amplification for <i>in vitro</i> transcription
FA680	AGGAAACtCCTAAGCTTAAAAGC	Reverse primer for the amplification of <i>aapA3</i> -Tr G183A for <i>in vitro</i> transcription

\* Nucleotides in lowercase correspond to the mutations to be introduced by site-directed mutagenesis PCR.

\*\* Sequences highlighted in bold correspond to *rpsL-erm*, *aphA3* or *catCG* 5'-overhang tails used for assembly PCR during the construction of the different deletion strains.

### 8.3. Related paper I: Mechanistic insights into type I toxin antitoxin systems in *Helicobacter pylori*: the importance of mRNA folding in controlling toxin expression

*Nucleic Acids Research*, 2017 **1–14**  
doi: 10.1093/nar/gkw1343

## Mechanistic insights into type I toxin antitoxin systems in *Helicobacter pylori*: the importance of mRNA folding in controlling toxin expression

Hélène Arnion, Dursun Nizam Korkut, Sara Masachis Gelo, Sandrine Chabas, Jérémy Reignier, Isabelle Iost and Fabien Darfeuille\*

INSERM U1212, CNRS UMR5320, Univ. Bordeaux, ARNA Laboratory, 146 rue Léo Saignat, F-33076 Bordeaux, France

Received July 08, 2016; Revised December 20, 2016; Editorial Decision December 21, 2016; Accepted December 22, 2016

#### ABSTRACT

**Type I toxin-antitoxin (TA) systems have been identified in a wide range of bacterial genomes. Here, we report the characterization of a new type I TA system present on the chromosome of the major human gastric pathogen, *Helicobacter pylori*. We show that the *aapA1* gene encodes a 30 amino acid peptide whose artificial expression in *H. pylori* induces cell death. The synthesis of this toxin is prevented by the transcription of an antitoxin RNA, named *IsoA1*, expressed on the opposite strand of the toxin gene. We further reveal additional layers of post-transcriptional regulation that control toxin expression: (i) transcription of the *aapA1* gene generates a full-length transcript whose folding impedes translation (ii) a 3' end processing of this message generates a shorter transcript that, after a structural rearrangement, becomes translatable (iii) but this rearrangement also leads to the formation of two stem-loop structures allowing formation of an extended duplex with *IsoA1* via kissing-loop interactions. This interaction ensures both the translation inhibition of the *AapA1* active message and its rapid degradation by RNase III, thus preventing toxin synthesis under normal growth conditions. Finally, a search for homologous mRNA structures identifies similar TA systems in a large number of *Helicobacter* and *Campylobacter* genomes.**

#### INTRODUCTION

The bacterial pathogen *Helicobacter pylori* is the etiologic agent of chronic gastritis and peptic ulcers and plays a major role in the genesis of gastric cancer (1). About half of the human population is infected by this bacterium, which is responsible for about 700 000 deaths worldwide every year

(2). To chronically survive and multiply in the human stomach, *H. pylori* has developed original strategies to modulate its gene expression in response to various stresses. Riboregulation, which has emerged as a major level of regulation in bacteria, was also proposed to play an important role in the adaptive response of *H. pylori* (3). However apart from housekeeping RNAs, transfer-messenger RNA, signal recognition particle RNA, 6S RNA and M1 RNA (RNase P), none of the enterobacterial small non-coding RNAs (sRNAs) are conserved in this bacterium. A combination of bioinformatics and genome wide RNA-seq analysis allowed us to characterize the *H. pylori* transcriptome and to reveal the existence of more than 60 new sRNAs in *H. pylori* strain 26695 (4). Regulator of polymeric G repeats (RepG) was identified as the first example of a *trans*-acting sRNA in *H. pylori*, repressing the expression of TlpB, a chemotaxis receptor (5). Although several of the RNA-seq identified sRNAs in *H. pylori* are putative regulators, their mechanisms and functions are still unknown.

Among the sRNA with the highest level of expression in *H. pylori* strain 26695 was an intriguing family of six homologous *cis*-encoded antisense RNAs (named *IsoA1* to *IsoA6*) that are expressed on the opposite strand of a novel class of small mRNAs (4). Using *in vitro* translation, we previously showed that each of the small mRNA of the A family expresses a short peptide (30 amino acids), designated *AapA* (Antisense-associated peptide family A). We also showed that the *AapA* and *IsoA* transcripts are both constitutively expressed *in vivo* during exponential growth, defining a small expression module, repeated many times at six different chromosomal loci (I–VI) (4). *In vitro* translation of *AapA1* and *AapA3* mRNAs was specifically inhibited by their cognate *IsoA1* and *IsoA3* antisense RNA, respectively. Due to the gene organization of these loci, it was hypothesized that these expression modules might constitute a new family of chromosomally encoded type I toxin-antitoxin (TA) systems. The TAs systems are categorized into six types based on their genetic organization and the

\*To whom correspondence should be addressed. Tel: +33 557574565; Email: fabien.darfeuille@inserm.fr

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

nature of the antitoxin (6,7). In the type I, the toxin is down-regulated by base-pairing of the antitoxin sRNA with the stable mRNA of the toxin (8). These systems were initially discovered on plasmids, where they play a key role in their stabilization during bacterial cell division, a phenomenon also known as post-segregational killing (9). When present on the chromosome, the identification of their function is less intuitive. A few of them have been reported to play important roles in adaptive responses to stress, including phenomena such as bacterial persistence (10).

In the present study, we characterize the *aapA1*/IsoA1 locus of *H. pylori* and demonstrate that it belongs to a new family of type I TA system. By using an artificial expression system, we show that the *aapA1* gene encodes a small peptide whose expression leads to toxicity. The synthesis of the toxin is prevented by IsoA1 sRNA which thus acts as an antitoxin. Surprisingly, the use of rifampicin during RNA decay measurements reveals the existence of a transcript generated from a 3' processing of the highly stable *AapA1* full-length (FL) mRNA. By using *in vitro* translation assays and footprinting experiments, we further demonstrate that, in contrast to the FL mRNA, the processed *AapA1* mRNA can be translated due to a structural rearrangement of the 5' untranslated region (UTR). This truncated transcript binds IsoA1, creating an extended duplex that prevents ribosome binding and that is targeted for degradation by RNase III. This degradation prevents the accumulation of the active message, and, together with the particular folding of the FL mRNA allow *H. pylori* growth despite the presence of a toxic gene in its genome. Finally, we take advantage of the strong conservation of the mRNA folding properties of this new TA system to identify many homologs in other *Helicobacter* and *Campylobacter* species. Interestingly, they are not only present on the chromosome but also associated with mobile genetic elements (MGE) such as plasmids, prophages and integrative and conjugative elements (ICE).

## MATERIALS AND METHODS

### Molecular techniques

Molecular biology experiments were performed according to standard procedures and the supplier (NEB) recommendations. QIAprep Spin Miniprep Kit (Qiagen), PureLink® HiPure Plasmid Maxiprep Kit (Thermo Fisher Scientific) and QIAamp DNA Mini Kit (Qiagen) were used for plasmids preparations and *H. pylori* genomic DNA extractions, respectively. PCR were performed either with Taq Core DNA polymerase (MP Biomedicals), or with Phusion Hot Start DNA polymerase (Finnzymes) when the product required high fidelity polymerase.

### *H. pylori* strains and culture conditions

The *H. pylori* strains used in this study (Supplementary Table S1) were 26 695 (11) B128 (12,13) and X47-2AL (14). Strains were grown on Columbia agar plates supplemented with 7% horse blood and Dent selective supplement (Oxoid, Basingstoke, UK) for 24–48 h depending on the strain. Liquid cultures were performed in brain-heart infusion medium (Oxoid) supplemented with 10% fetal bovine

serum and Dent. *H. pylori* plates and liquid cultures were incubated at 37°C under microaerobic conditions (10% CO<sub>2</sub>, 6% O<sub>2</sub>, 84% N<sub>2</sub>) using an Anoxomat (MART microbiology) atmosphere generator. For liquid cultures, bacteria harvested from plates were inoculated at an optical density at 600 nm of 0.05 (OD<sub>600</sub> = 0.05) into 5 ml (tubes, shaking at 175 rpm) brain-heart infusion medium supplemented with 10% fetal bovine serum and Dent supplement. After 12–24 h, pre-cultures were diluted to an OD<sub>600</sub> of 0.05 into 25 ml (flasks, shaking at 125 rpm). Plasmids used for cloning were amplified in *Escherichia coli* strain JM109, which was grown in Luria–Bertani medium, supplemented either with kanamycin (50 µg.ml<sup>-1</sup>) or chloramphenicol (30 µg.ml<sup>-1</sup>). For *H. pylori* mutant selection and culture, antibiotics were used at the following final concentrations 20 µg.ml<sup>-1</sup> kanamycin (Sigma) and 8 µg.ml<sup>-1</sup> chloramphenicol (Sigma).

### Construction of *H. pylori* mutant strains

Chromosomal mutants of *H. pylori* (Supplementary Table S1) were obtained by natural transformation as previously described (15). The  $\Delta aapA1$ /IsoA1,  $\Delta rnc$ , *aapA1* <sup>$\Delta$ -10 box</sup> mutants and X47-2AL *aapA1*<sup>26695</sup> or *aapA1*<sup>B128</sup> complementation strains were constructed by homologous recombination using a PCR cassette carrying an antibiotic resistance gene (*aphA-3* or *catGC* gene conferring kanamycin or chloramphenicol resistance, respectively), flanked by approximately 500 base-pairs (bp) regions upstream and downstream of the gene of interest, as previously described (16). For the *aapA1*/IsoA1 deletion, the region between nt 1 245 653 to 1 245 866 (encompassing the Shine–Dalgarno (SD) sequence and the ATG and TAG codons) was removed, placing *aphA-3* (from its own SD to its own stop codon) under the control of the *aapA1* promoter and upstream of a transcriptional terminator (Supplementary Figure S8). The B128 and X47-2AL  $\Delta rnc$  strains were constructed by replacing the *rnc* Open Reading Frame (ORF) (from the ATG to 24 nt before its stop codon) by the *aphA-3* gene amplified from its own ATG to its own TAG (Supplementary Figure S11). The *aapA1* <sup>$\Delta$ -10 box</sup> mutant was constructed by removing the two last nt of the *aapA1* -10 box sequence (TAAAT) (Supplementary Figure S9). To construct the X47-2AL *aapA1*<sup>26695</sup> / *aapA1*<sup>B128</sup> complementation strains, the *aapA1* locus (nt 1 245 624 to 1 245 986 for 26695 strain and 287 085 to 286 691 for B128 strain) with its own promoter (either from strain 26695 or B128) was fused to the *catGC* resistance gene and inserted into the *rdxA* locus of the X47-2AL strain. RdxA is a non-essential gene routinely used for complementation in *H. pylori* (17). The X47-2AL has a non-homologous copy of the *aapA1*/IsoA1 module (Supplementary Figure S10). The genomic DNA of *H. pylori* strain 26695, B128, P12 *repG::aphA-3* (5) and the vectors pUC18K2 (kanamycin resistance gene) and pILL2150 (chloramphenicol resistance gene) were used as template for all PCR amplifications (see plasmids and primers list, Supplementary Table S2 and S3, respectively).

### Plasmids constructions

Three plasmids carrying different isoforms of the *aapA1*/IsoA1 locus (from the start codon to the 3' UTR) were generated for this study. The pA1-IsoA1 plasmid contains the wild-type *aapA1*/IsoA1 sequence; the pA1 plasmid carries two mutations that inactivate the *isoA1* promoter without changing the coding sequence of the peptide and pA1\* is a derivative of pA1 containing an additional mutation in the start codon of AapA1 (ATG → ATT). The pA1-IsoA1 and pA1 plasmids were obtained by PCR amplification of genomic DNA from the 26695 strain, with the primer pairs FD213/FD180 and FD212/FD180, respectively. These products were cloned into pILL2157bis (18) between the *NdeI* and *BamHI* restriction sites (Supplementary Table S2). The pA1\* plasmid was obtained by site-directed mutagenesis of the pA1 plasmid using primers FD608/FD609. The resulting plasmids were introduced into *H. pylori* strain B128 by mobilization, as described in Backert *et al.* (19).

### Total RNA extraction

For RNA extraction, bacterial growth was stopped at the desired OD<sub>600</sub> by adding 1.25 ml cold stop solution (95% ethanol, 5% acidic phenol) to 10 ml of culture, which was placed on ice. Cells were then centrifuged for 10 min at 3500 rpm at 4°C, and the pellets were stored at -80°C. Cell pellets were resuspended in 600 µl lysis solution (20 mM NaAc pH 5.2, 0.5% sodium dodecyl sulphate (SDS), 1 mM ethylenediaminetetraacetic acid (EDTA)) and added to 600 µl hot acidic phenol. After 6–10 min incubation at 65°C, the mixture was centrifuged for 10 min at 13 000 rpm at room temperature. Next the aqueous phase was transferred to a phase locked gel tube (Eppendorf) with an equal volume of chloroform and centrifuged 10 min at 13 000 rpm at room temperature. Total RNA was then precipitated from the aqueous phase by adding 2.5 volumes of EtOH 100% and 10% 3 M NaOAc pH 5.2. For RNA half-life determinations, rifampicin (Sigma, prepared at 34 mg.ml<sup>-1</sup> in methanol) was added to the culture at a final concentration of 80 µg.ml<sup>-1</sup> and cells were harvested after 0, 2, 5, 10, 20, 30, 60, 120 and 240 min of incubation. A culture where rifampicin was replaced by the same volume of methanol served as a non-treated control.

### Northern blot

For Northern blot analysis, 5–20 µg of total RNA were separated on an 8% polyacrylamide (PAA), 7 M urea, 1X Tris Borate EDTA (TBE) gel. RNA was transferred to a nylon membrane (Hybond<sup>TM</sup>-N, GE Healthcare Life Science) by electroblotting in TBE 1X at 8V overnight. Then RNA was cross-linked to the membrane by UV irradiation and hybridized with 5'-labeled ( $\gamma$ -<sup>32</sup>P) oligodeoxynucleotides (see Supplementary Table S3) in a modified Church buffer (1 mM EDTA, 0.5 M NaPO<sub>4</sub> pH 7.2, 7% SDS) overnight at 42°C. Membranes were washed two times for 5 min in 2X SSC, 0.1% SDS and revealed using a Pharos FX phosphorimager (Biorad).

For riboprobe synthesis, a DNA template containing a T7 promoter sequence was amplified by PCR with primers

FD671/FD672 from the 26695 strain genomic DNA as template. The *AapA1* riboprobe was prepared as described in the Maxiscript Kit (Ambion) and purified on a Sephadex G50 column. Hybridization was performed in the modified Church buffer at 64°C and the membrane was washed 2 times for 5 min in 2X SSC, 0.1% SDS at 64°C.

### In vitro transcription and translation assays

For *in vitro* synthesis of the *AapA* and IsoA RNAs, DNA templates were amplified from *H. pylori* 26695 genomic DNA using the primer pairs FD54/FD452 (*AapA1-FL*), FD54/FD55 (*AapA1-Tr1*), FD9 /FD15 (IsoA1 wt), FD205/FD234 (IsoA1L1L2), each forward primer carrying a T7 promoter sequence (see Supplementary Table S3). FD205/FD234 primers carry both three mutations resulting in the synthesis of the IsoA1L1L2 RNA that contains three mutations in each loop. *In vitro* transcription was carried out using the MEGAscript<sup>®</sup> T7 Transcription Kit (Ambion #AM1334) according to the manufacturer's protocol. After phenol:chloroform extraction followed by isopropanol precipitation, the RNA samples were desalted by gel filtration using a sephadex G-50 or G-25 (GE Healthcare) column, depending on the RNA size. For *in vitro* translation of the AapA1.FL and AapA1.Tr1 mRNAs, 0.05 to 1 µg of RNA was added to the *E. coli* S30 kit (Promega, #L1030) as previously described (4).

### In vitro structure probing

*AapA1* transcripts were first dephosphorylated by the CIP alkaline phosphatase (NEB) and then labeled with 10 pmol of  $\gamma$ -<sup>32</sup>P-ATP and T4 PNK enzyme (NEB). Labeled RNA was purified on an 8% PAA containing 7 M urea 1X TBE and eluted overnight at 4°C under shaking in 750 µl elution buffer (0.1 M NaOAc pH 5.2, 0.1% SDS). Then, RNA was extracted by phenol:chloroform, desalted and concentrated by ethanol precipitation, pellets were resuspended in 50 µl H<sub>2</sub>O and stored at -20°C. Before use, each *in vitro* transcribed RNA was denatured by incubation at 90°C for 2 min in the absence of magnesium and salt, then chilled on ice for 1 min, followed by a renaturation step at room temperature for 15 min in 1X Structure Buffer (10 mM Tris-HCl pH 7.0, 10 mM MgCl<sub>2</sub>, 100 mM KCl).

Structures probing analyses were performed as described previously (4,20), using 0.1 pmol of labeled *AapA1* RNA. To determine the secondary structure of RNA in native conditions (N), 1 µl RNase T1 (0.01 U.µl<sup>-1</sup>; Ambion, #AM2283), 1 µl RNase TA (0.005 µg.ml<sup>-1</sup>; Ambion, #AM2275) and 1 µl of RNase V1 (0.0005 U.µl<sup>-1</sup>; Ambion, #AM2275) were added to the labeled RNA and incubated in 1X Structure Buffer for 1 to 2 min at 37°C. For the denaturing conditions (D), 1X Sequencing Buffer (20 mM Sodium Citrate, pH 5.0, 1 mM EDTA, 7M Urea) was used, with the same RNase T1 concentration and 1 µl of RNase TA (0.01U.µl<sup>-1</sup>) and incubation was performed at 37°C for 5 min. We performed lead acetate (5 mM final concentration) digestions on both *aapA1-Tr1* and *aapA1-FL* in the absence or in the presence of a 2- or 10-fold excess of wild-type or mutated IsoA1 RNAs. All the reactions were stopped by adding 10 µl of 2X Loading Buffer (95% formamide, 18 mM EDTA, Xylene Blue and Bromophenol

Blue) and stored at  $-20^{\circ}\text{C}$ . Cleaved fragments were then analyzed on an 8% denaturing PAA gel containing 7 M urea and 1X TBE. Gels were dried 45 min at  $80^{\circ}\text{C}$ , and revealed using a Pharos FX phosphorimager (Biorad).

### Sequence conservation analysis

A sequence similarity search for *AapA* mRNA homologs was carried out on the *Helicobacter* and *Campylobacter* genomes using the Geneious program on an iterative process (Geneious version 8.1.3 <http://www.geneious.com> (21)). Genome sequences used are listed in Supplementary Table S4. A sequence similarity with a threshold of 60% was used as a first restraint with the *H. pylori* 26695 *aapA*/IsoA nucleotide (nt) sequence (see Supplementary Figure S12 for more details). Once identified, the secondary structure of the *AapA* and IsoA RNAs were analyzed via the RNA fold plugin (22) implemented in Geneious. Finally, each locus was manually inspected for the presence of a promoter for both *aapA* and IsoA RNAs, as well as the presence of a SD sequence upstream of the *AapA* ORF. Every new result was added to the query database and the process was completed when no new result was found (Supplementary Figure S13).

## RESULTS

### The expression of AapA1 peptide is toxic to *H. pylori*

In the present study, our first goal was to establish whether the *aapA1*/IsoA1 module located at the locus I of the *H. pylori* strain 26695 encodes a type I TA system (Figure 1). Indeed, we previously reported that *in vitro* translation of the *AapA1* mRNA leads to the synthesis of a small peptide of 30 amino acids whose expression is repressed *in vitro* by a small antisense RNA named IsoA1 (4). To this end, we cloned the coding sequence of this peptide into an *E. coli/H. pylori* shuttle vector (18) under the control of an isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG)-inducible promoter to generate plasmid pA1-IsoA1 (Figure 2A). Two mutated variants of this plasmid were also constructed. The first one carries two mutations in the IsoA1 antisense promoter (TATAAT  $\rightarrow$  TACAAG) that prevents the expression of the antisense RNA without changing the peptide sequence (pA1 plasmid). The second one contains an additional mutation in the *AapA1* mRNA start codon (ATG  $\rightarrow$  ATT), which prevents the expression of the *AapA1* peptide (plasmid pA1\*) (Figure 2A). Each plasmid was transformed into *H. pylori* B128 strain deleted for the chromosomal copy of *aapA1*/IsoA1. The transcripts expressed from the different constructs were analyzed by Northern blot (Supplementary Figure S1A). We confirmed the inducible expression of the *AapA1* transcripts in presence of IPTG despite some leakiness of the promoter in absence of IPTG. We also showed that the mutated *isoA1* promoter is inactive, as shown by the absence of IsoA1 (Supplementary Figure S1A).

We next analyzed the effects of the *AapA1* peptide expression on both growth rate (Figure 2B) and cell viability (Figure 2C). In the strain carrying plasmid pA1, induction of *AapA1* expression caused an immediate growth arrest and a  $10^4$ -fold decrease in cell number (Figure 2B and C). When IPTG was washed out from the cultures after 8

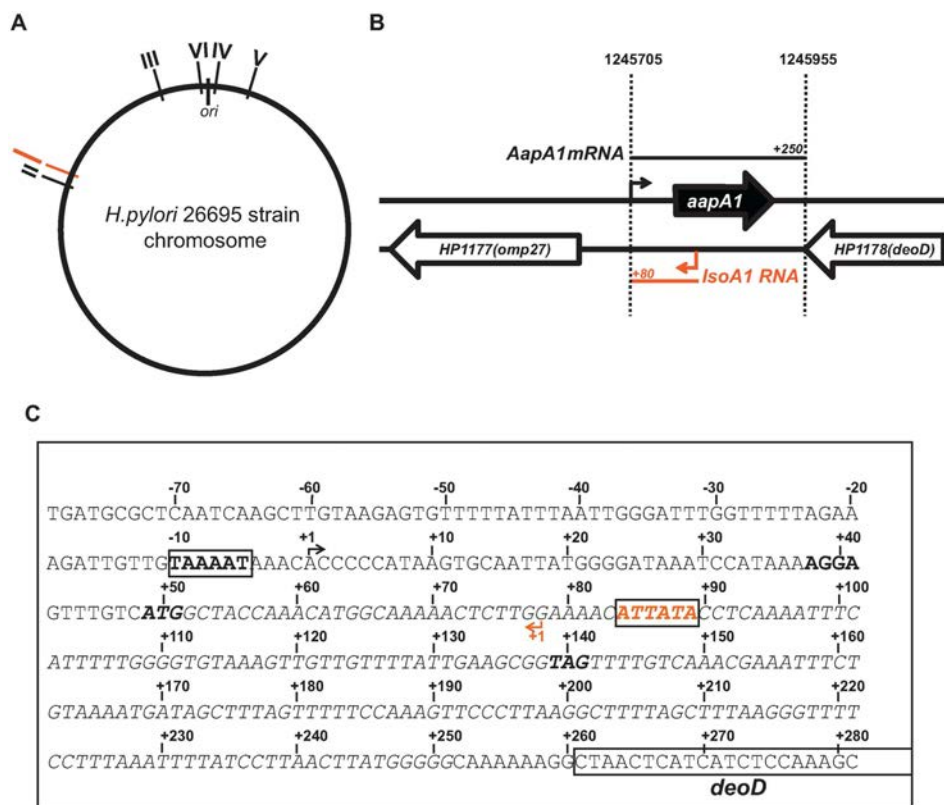
h of induction, cells expressing *AapA1* were not able to recover growth (data not shown). Thus, expression of *AapA1* is strongly toxic and bactericidal, and the onset of toxicity is fast relative to the generation time (3.8 h for pA1 expressing strain compared to 3.6 h for pA1-IsoA1 and pA1\*). Of note, no cell lysis was observed upon expression of the toxin (Figure 2B). Interestingly, this lethality was correlated with a dramatic degradation of the 23S and 16S ribosomal RNAs while the toxin-encoding mRNA was still present (Figure 2D and Supplementary Figure S1B). The toxicity was not observed in the strain carrying the pA1-IsoA1 plasmid, demonstrating that transcription of IsoA1 prevents the toxicity of *AapA1* and thus indeed acts as an RNA antitoxin. In addition, the absence of toxicity of the pA1\* plasmid (in which the translation of the *AapA1* mRNA is abolished) demonstrates that the toxicity is specifically due to the *AapA1* peptide expression and not to the *AapA1*-encoding mRNA. The same results were obtained when using the wild-type background instead of the  $\Delta aapA1$  B128 strain, indicating that the chromosomal copy of IsoA1 cannot neutralize the ectopic expression of *AapA1* (Supplementary Figure S1B).

Altogether these experiments demonstrate that the *aapA1*/IsoA1 module functions like a type I TA system in which the expression of a toxic peptide is prevented by the expression of an RNA antitoxin.

### 3'-processing of *AapA1* mRNA triggers translation activation

Type I TA systems are often characterized by a differential stability between the toxin-encoding mRNA and the antitoxin. To analyze the stability of both *AapA1* and *IsoA1* transcripts, cultures of *H. pylori* 26695 strain were treated with rifampicin to block transcription, and total RNA was extracted at different time points and analyzed by Northern Blot (Figure 3A). An oligonucleotide probe (FD47) targeting the *AapA1* coding region detected several species from 175 to 250 nucleotides (nt) (Figure 3A, left panel). A similar pattern of expression was confirmed with a riboprobe directed against the first 225 nt of the mRNA (Supplementary Figure S2). None of these transcripts were detected in a strain deleted for the *aapA1* promoter (Figure 3A right panel) showing that they are all produced from the same transcription start site at the position 1245705 (4). The strongest signal corresponds to the FL transcript of 250 nt (designated *AapA1*.FL), which displays a very slow decay with a half-life longer than 3 h. In contrast, the IsoA1 RNA has a relative short half-life of 15 min (Figure 3B). Most interestingly, 120 min after rifampicin addition, a transcript of around 225 nt appears. Since transcription is blocked, this transcript is the result of *AapA1*.FL processing and was consequently designated as *AapA1*.Tr1 (for truncated) (Figure 3A). Different probes targeting either the 5' or 3' end of the *aapA1*.FL revealed that the truncation occurs around 25 nt upstream of the 3' end of the FL message (Supplementary Figure S3). Another transcript of 172 nt, named *AapA1*.Tr2 was detected with FD47 but not with FA115 and FA116 probes (corresponding to the 5' and 3' ends of *AapA1*) indicating that this transcript is the result of cleavages of the FL message at both ends. Another faint band corresponding to a transcript of  $\sim 200$  nt was detected. However, since



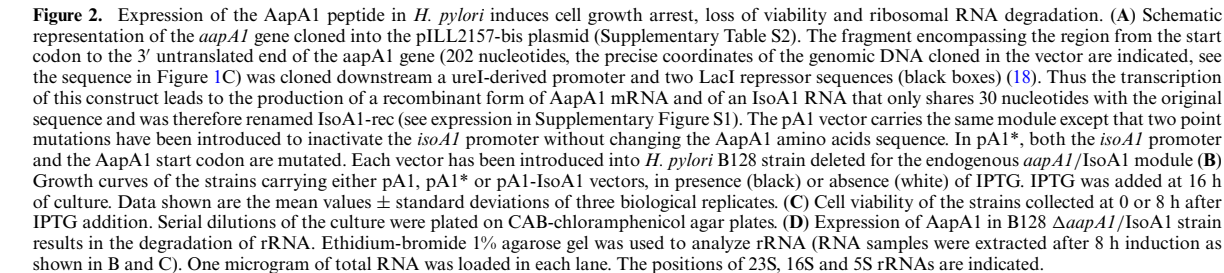


**Figure 1.** Genetic organization of the *aapA1*/IsoA1 locus in the *H. pylori* 26695 strain. (A) Localization of the six different loci (I–VI) containing an homolog of the IsoA sRNA on the chromosome of *H. pylori* 26695 strain. The locus I containing the AapA1/IsoA1 module is shown in red. All loci are in the same orientation on the chromosome, the *aapA* gene being always on the positive strand (forward). (B) Genomic organization of the locus I. The open reading frame encoding the 30 amino acids peptide AapA1 (in black) and IsoA1 small RNA (in red) are transcribed from the intergenic region between the *deoD* gene (HP1178, purine nucleotide phosphorylase) and the *omp27* gene (HP1177, outer membrane protein). Arrows indicate the respective transcriptional start sites of each transcript and black and red bars their approximate length determined by RNA-seq analysis (4). (C) The sequence of the intergenic region containing the *aapA1*/IsoA1 module in 26695 *H. pylori* strain is shown. The AapA1 fragment cloned into the pILL2157 vector is shown in italic. Promoter sequences (boxes) of *aapA1* and *isoA1* genes are shown in black and red, respectively. The Shine–Dalgarno sequence (SD), the start and stop codons are indicated by bold letters.

it was not observed in other *H. pylori* strains tested with the riboprobe (Supplementary Figure S2), we did not explore it in the present study. During the course of the rifampicin experiment we observed that *AapA1*\_*Tr1* appears at the same time that IsoA1 disappears. Since *AapA1* mRNA expression leads to toxicity when the antitoxin is absent (Figure 2), we propose that *AapA1*\_*Tr1* is the active message. To test this hypothesis, the translatability of *AapA1*\_*FL* and *AapA1*\_*Tr1* transcripts was assessed by *in vitro* translation assays (Figure 3C). Whereas a very faint signal corresponding to AapA1 peptide was observed with the FL transcript, a strong signal was detected with the Tr1 species (Figure 3C and 3D). These results demonstrate that a 3'-end processing is required for a fully efficient *in vitro* translation of the AapA1 peptide.

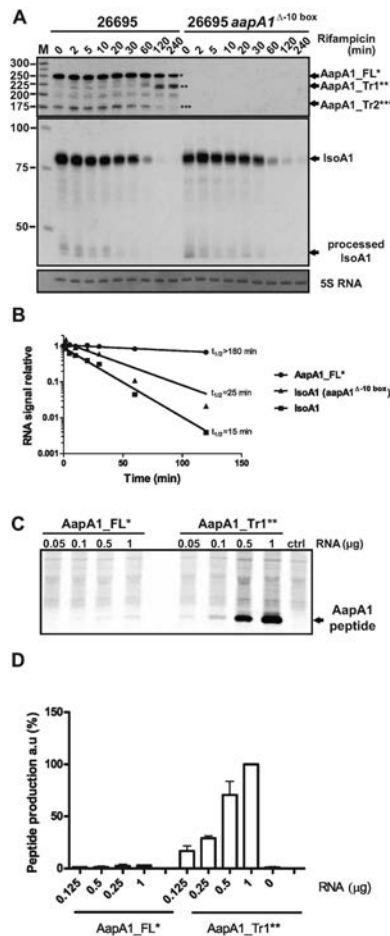
#### The SD sequence is accessible in the truncated mRNA and sequestered in the full-length transcript

To investigate the reasons for the differential translation efficiency of the two *AapA1* transcripts (FL and Tr1), we performed *in vitro* structure probing by combining chemical and enzymatic footprinting (Figure 4). RNAstructure software 5.2 (22) was used to generate secondary structure predictions of both transcripts according to the experimental data obtained (see the structures Figure 5). Lead cleavage showed a strong structural rearrangement of the 5' UTR between the two transcripts. Indeed while the cleavage pattern of *AapA1*\_*FL* (Figure 4A, lane 7) fits very well with the formation of two internal loops (IL I and IL II, Figure 5), this region folds as two apical loops (AL I and AL II, Figure 5) connected by a 14 nt linker in *AapA1*\_*Tr1* (Figure 4B, lane 7). *AapA1*\_*FL* folding involves a long distance interaction between the 5' and 3' ends of the transcript where the SD sequence is sequestered (Figure 5). This structure predic-



tion explains the high stability of the transcript (few single stranded nucleotides accessible to RNases) but also why it is poorly translated. On the contrary, the 3'-end truncation of *AapA1\_FL* leads to an important structural rearrangement in the 5' UTR in which the SD sequence becomes accessible to ribosomes (Figure 5). In addition we also observed increased lead cleavages in the residues 66–75 in Tr1 versus the FL transcript, indicating that the second stem-loop might be more accessible to ribosomes in the truncated transcript than in the full length mRNA.

To characterize the interaction between *AapA.Tr1* and *IsoA1* RNA, we performed lead probing on 5'-end-labeled *AapA.Tr1* transcript in absence or presence of *IsoA1* RNA (Figure 4B, lanes 7, 8 and 10). A comparison of the cleavage patterns showed that a 2-fold excess of *IsoA1* resulted in a complete protection of the mRNA region that is complementary to *IsoA1* (Figure 4B, lane 8). Indeed, lead cleavage was strongly reduced in the 5' UTR encompassing AL I, the 14 nt linker and AL II and this was even more pronounced when *IsoA1* was added in a 10-fold excess (Figure 4B, lane 10). Moreover, other structural rearrangements (highlighted with a star) were observed indicating a refolding of the *AapA.Tr1* transcript upon the formation of an extended duplex with *IsoA1* (Figure 4B). Under the



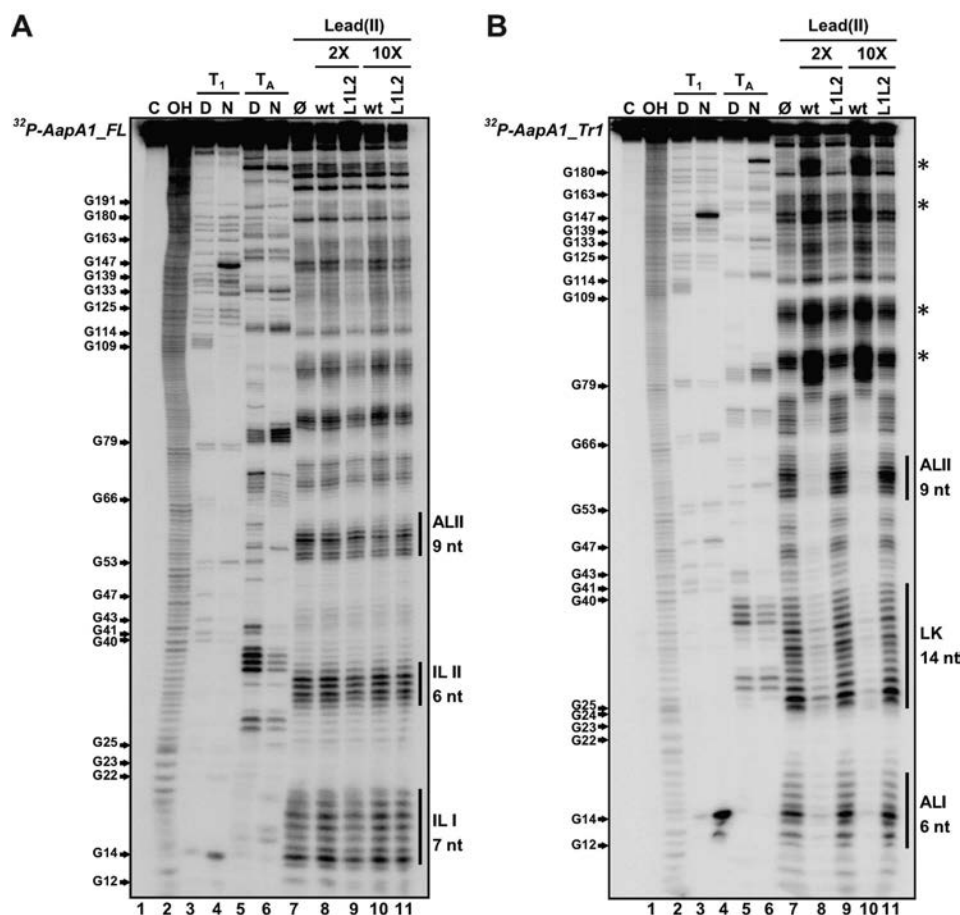
**Figure 3.** A truncated *AapA1* mRNA species revealed by rifampicin treatment is efficiently translated *in vitro* (A) *AapA1* and IsoA1 transcripts half-lives were determined in the 26695 *H. pylori* strain (left panel) or in an isogenic mutant deleted for *aapA1* promoter ( $\Delta$ -10 box, right panel). After rifampicin addition (at OD<sub>600</sub> = 1.7), aliquots of cultures were collected at several time points (as indicated on top of the gel), RNA extracted and subjected to Northern blot analysis. The same membrane was successively probed with FD47 and FD198 labeled oligonucleotides to detect either *AapA1* or IsoA1 transcripts, respectively. Arrows indicate the different transcripts: *AapA1\_FL* (one star), *AapA1\_Tr1* (two stars), *AapA1\_Tr2* (three stars) and IsoA1 full-length and processed transcripts. A labeled DNA marker (lane M) was used for size estimation. Proper loading was assessed by staining the membrane with methylene blue (level of 5S rRNA is shown). The Northern blot shown here is a representative of more than three independent experiments using several biological replicates. (B) RNA decay was determined by plotting normalized intensities (RNA signal relative to time 0) of band corresponding to *AapA1\_FL*\* and IsoA1 transcripts, either in the wild-type or  $\Delta$ -10 box mutated strains, as function of time after rifampicin addition. Approximate half-lives ( $t_{1/2}$ , in min) are indicated for each transcript analyzed. (C) *In vitro* translations assays of *AapA1\_FL* (left) and *AapA1\_Tr1* (right). Increasing amounts of *in vitro* synthesized mRNAs ( $\mu$ g) were added to *E. coli* S30 extracts in presence of <sup>35</sup>S methionine. Control lane (Ctrl) shows the translation background obtained without exogenous mRNA. Here is a representative experiment out of several independent assays. (D) The amount of AapA1 peptide produced from both transcripts (*AapA1\_FL* and *Tr1*) was quantified and the intensity of the band at the highest *AapA1\_Tr1* quantity (1  $\mu$ g) was set to 100%. Data are the average and standard deviations of three independent experiments.

same conditions, no interaction could be detected between *AapA1\_FL* and IsoA1 (Figure 4A, lanes 8 and 10), indicating that the antisense RNA specifically targets the 3'-end-processed mRNA. Additionally, structure mapping and secondary structure predictions performed on IsoA1 RNA (Supplementary Figure S4) showed that the *AapA1\_Tr1* 5' UTR and IsoA1 both fold into two complementary stem-loops that could favor the formation of loop-loop interactions.

To determine whether such kissing-loop complexes could represent the initial step of the *AapA1*/IsoA1 hybrid formation, we performed *in vitro* footprinting experiments in presence of neomycin. This antibiotic of the aminoglycoside family is known to stabilize kissing complexes and to prevent the formation of extended duplexes (23). In presence of increasing concentration of neomycin, the LK region of *AapA1\_Tr1* was no longer base-paired whereas the ALI and ALII loops were still protected (Supplementary Figure S5A), confirming that loop-loop interactions are promoting the interaction between the two transcripts. In addition, we could map neomycin's binding regions at the boundaries of the helices (Supplementary Figure S5A, lanes 6 and 7), explaining why this antibiotic prevents the *AapA1\_Tr1*/IsoA1 hybrid to extend into a full duplex (Supplementary Figure S5A lanes 10–11). No interaction was observed with IsoA3 which displays a similar fold than IsoA1 but a different sequence in the loop (Supplementary Figure S4B).

To further confirm that the initial interaction is mediated via loop-loop interaction, we constructed an IsoA1 variant mutated in its two apical loops L1 and L2 (named IsoA1 L1L2) and tested its ability to bind the *AapA1\_Tr1* transcript (Figure 4 and Supplementary Figure S6). We first verified that the three nt mutations introduced in each loop of IsoA1 did not alter its folding (Supplementary Figure S6). In contrast to the wild-type (wt) IsoA1, the addition of IsoA1 L1L2 did not change the digestion profile of *AapA1\_Tr1* (Figure 4B, lanes 9 and 11) indicating that this mutant cannot bind despite a strong sequence complementarity. This result confirms that the interaction between *AapA1\_Tr1* and IsoA1 is mediated via a loop-loop interaction. Finally, we performed the reverse experiment, i.e. lead cleavage of 5' end-labeled IsoA1 upon *AapA1\_Tr1* addition. The results confirmed the duplex formation between IsoA1 and the first 76 nt of *AapA1\_Tr1* mRNA (Supplementary Figure S6A). This very long hybrid is masking the ribosome binding site (RBS) of *AapA1\_Tr1* (Supplementary Figure S5B) suggesting that the primary role of the antitoxin is to inhibit the translation of the toxin.

Taken together these data indicate that IsoA1 antisense RNA binds specifically to the *AapA1\_Tr1* active transcript via kissing-loop interactions (Supplementary Figure S5B). These short interactions are key determinants in the specificity of the interaction and explain why each IsoA RNA only represses the translation of its cognate mRNA, as shown previously (4).



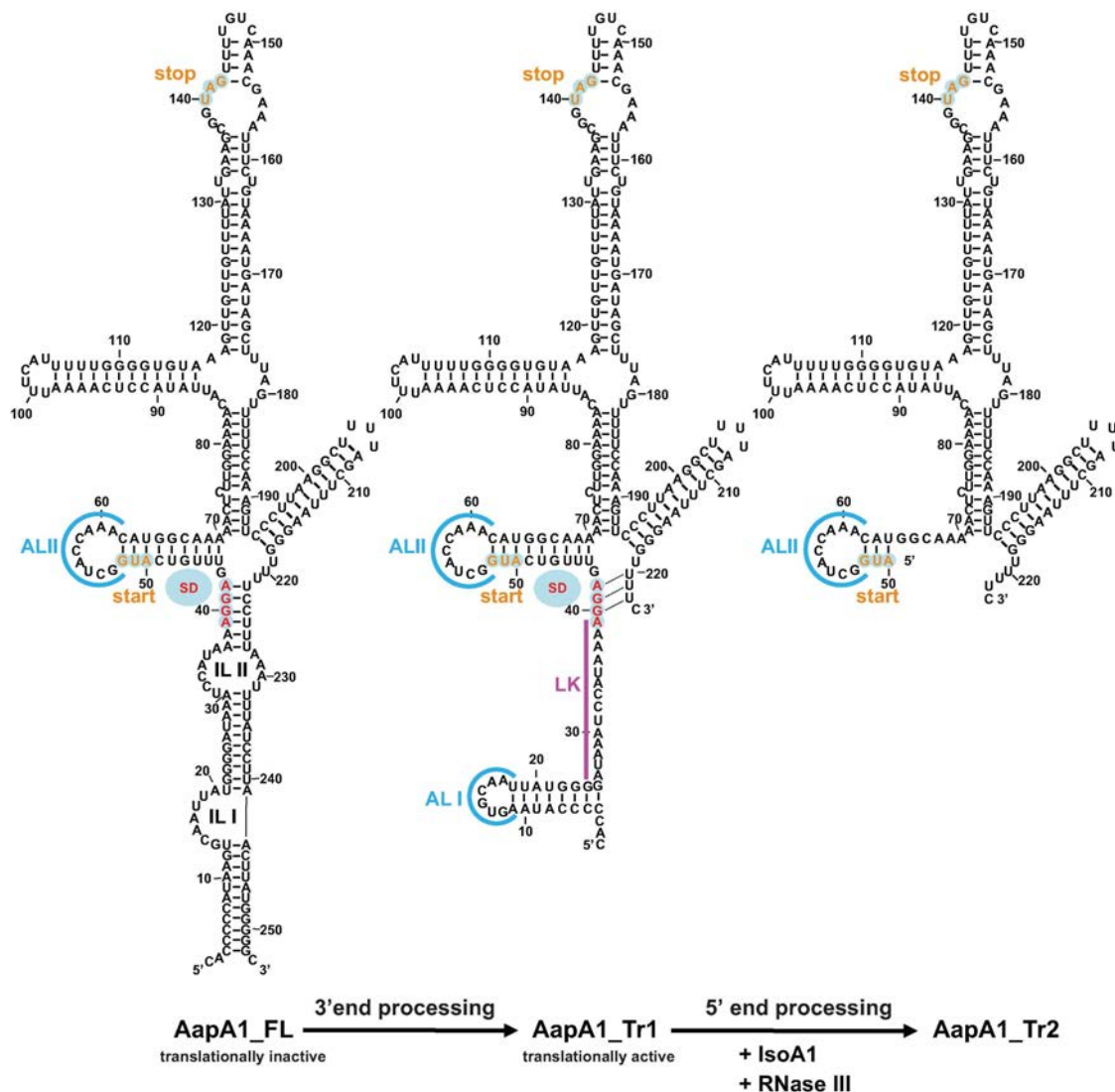
**Figure 4.** Structure probing of (A) *AapA1-FL* and (B) *AapA1-Tr1* RNAs in presence or absence of IsoA1. The secondary structure of each *in vitro* transcribed RNA was probed by submitting ~0.1 pmol  $^{32}$ P- labeled RNA to partial enzymatic digestion either under native (N) or denaturing conditions (D) (RNase T1 cleaving single stranded G residues, lanes 3 and 4; RNase TA cleaving single stranded A residues, lanes 5 and 6). The interaction between *AapA1* and IsoA1 was mapped using lead probing (lanes 7–11) in the absence (lane 7) or presence of either 2 or 10 times excess of wild-type IsoA1 (wt, lanes 8 and 10) or IsoA1 mutated in its apical loop (L1L2, lanes 9 and 11). Structure mapping, as well as secondary structures of IsoA1 wt and L1L2, are shown in Supplementary Figure S4. Untreated RNA (lane 1, denoted C) and partially alkali digested RNA (lane 2, denoted OH) served as control and ladder, respectively. Cleaved fragments were analyzed on an 8% denaturing polyacrylamide gel. Positions of all G residues are indicated relative to the transcription start site of the *aapA1* gene (left of the gel). Single stranded regions involved in the 5' structural rearrangement are indicated by vertical black bars (right of the gel). The internal loops I and II (IL I, IL II) present in *AapA1-FL* are replaced by an apical loop (AL II) and a 14-nucleotides linker (LK) in *AapA1-Tr1*. Black stars on the right of the gel indicate other structural rearrangements observed following IsoA1 RNA binding.

#### RNase III cleavage ensures rapid turnover of the translationally active message

Base-pairing between *AapA1-Tr1* and IsoA1 creates a long duplex of 76 bp (Supplementary Figure S5B), which could be a good substrate for the double-stranded specific ribonuclease RNase III. To investigate the role of this ribonuclease in *aapA1*/IsoA1 regulation *in vivo*, we compared both transcripts in wt and RNase III-deficient *H. pylori* strains. Since we could not delete the *rnc* (HP0662) gene encoding RNase III in the 26695 strain, we used two other strains, B128 and X47-2AL, for which the deletion of *rnc* could be obtained. We inserted the complete *aapA1*/IsoA1 module of either the 26695 or the B128 strain into the *rdxA*

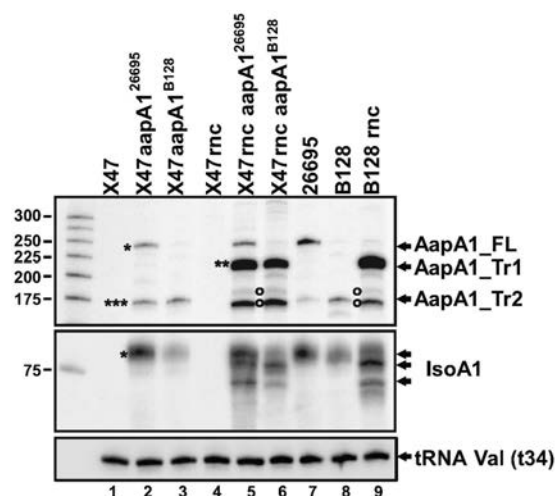
gene of the wt and  $\Delta rnc$  X47-2AL strains (see Supplementary Figure S10 for details). We then analyzed the expression of *AapA1* and IsoA1 RNAs by Northern blot (Figure 6). The introduction of either *aapA1*/IsoA1 module in the X47-2AL background resulted in the expression pattern observed for the parental strain (Figure 6, compare lanes 2 and 3 with lanes 7 and 8). Thus, transcription of the *aapA1*/IsoA1 module is not influenced by the recipient strain under these conditions. In both the X47-2AL and B128 backgrounds, the *rnc* deletion led to a strong accumulation of the *AapA1-Tr1* truncated form (two stars) while the levels of *AapA1-FL* (one star) were unaffected (Figure 6, lanes 5, 6 and 9). Same samples were run longer for a better separation of the different species (Supplementary Fig-





ure S7). In addition *AapA1-Tr2* (three stars) was absent in the  $\Delta rnc$  strains, indicating that this transcript results from RNase III cleavage of *AapA1-Tr1* (Supplementary Figures S2 and S7). In contrast, the amount of the full-length IsoA1 RNA seemed unaffected in absence of RNase III. We have shown that, in wild-type cells, IsoA1 RNA is in large excess over *AapA1-Tr1*. Therefore, the proportion of IsoA1





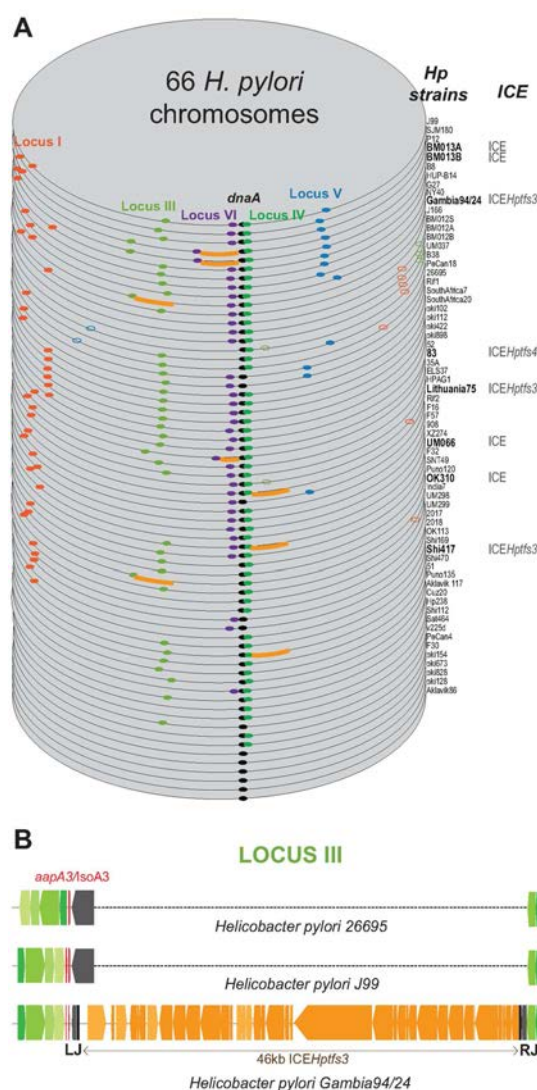
**Figure 6.** RNase III prevents accumulation of the active *AapA1.Tr1* mRNA. Expression of *AapA1* and *IsoA1* RNAs in wild type and isogenic  $\Delta rnc$  *H. pylori* strains (X47-2AL, 26695 and B128) were analyzed by Northern blot. The *aapA1/IsoA1* locus of the 26695 strain (*AapA1*<sup>26695</sup>) or of the B128 strain (*AapA1*<sup>B128</sup>) was inserted in the X47-2AL strain and its  $\Delta rnc$  derivative. Samples from each strain were collected when culture reached OD<sub>600</sub>  $\approx$  1. tRNA Val (t34) served as a loading control and was probed with the FD499 radiolabeled oligonucleotide. Upper panel: *AapA1* transcripts revealed with FD47 probe; middle panel: *IsoA1* transcripts revealed with FD198 probe.

alternative degradation of *AapA1.Tr1* mRNA by unknown ribonucleases.

Overall these data indicate that RNase III is the major enzyme responsible for the *AapA1.Tr1/IsoA1* duplex degradation thereby ensuring the rapid turnover of the *AapA1* active message.

#### Identification of new *aapA/IsoA* systems in other *Helicobacter* and *Campylobacter* species through structural conservation

We analyzed the conservation of the newly characterized type I *aapA/IsoA* TA system among the *H. pylori* strains. A PSI-BLAST search to detect small peptides (as described in (24)) resulted in very few positive hits because few ORFs encoding small proteins have been annotated. We next conducted a TBLASTN search against nucleotide database. This search revealed the presence of *aapA/IsoA* homologs in almost every *H. pylori* genome and in the closely related genomes of *Helicobacter acinonychis* and *Helicobacter cetorum* (Supplementary Table S4). Moreover, the synteny associated with the different loci was strongly conserved between the various strains (Supplementary Figure S12). The *aapA/IsoA* loci were thus classified according to their position at the 5 loci named I, III, IV, V and VI (Figure 7A). Although *IsoA2* at locus II is conserved, no functional promoter upstream of a putative *AapA2* mRNA nor a conserved small ORF longer than four amino acids were detected. Thus, this locus seems not functional and therefore was no longer considered in this study. Among the 66 *H. py-*



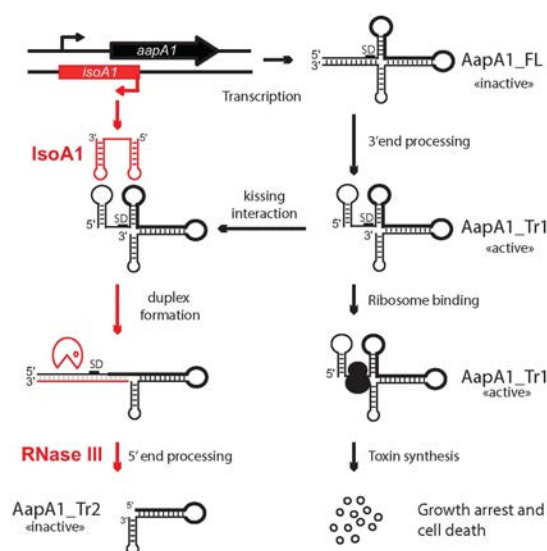
**Figure 7.** Conservation of *aapA/IsoA* TA systems in 66 *Helicobacter pylori* strains. (A) Each genome is stacked according to the decreasing number of systems present at each locus and aligned to the origin of replication (next to the *dnaA* gene). Each locus containing at least one copy of a functional *aapA/IsoA* module (presence of both promoters and of a small peptide with the consensus length of 30 amino acids) is represented according to the following color code: locus I (red), III (light green), IV (dark green), V (blue) and VI (purple). Most systems have the mRNA on the positive strand; the few ones on the minus strand are shown as open dots. A thick orange line shows the presence of an integrative and conjugative element (ICE) next to an *aapA/IsoA* module. (B) Schematic representation of the locus III alignment in *H. pylori* strains 26695, J99 and Gambia94/24. Blue and green colored genes represent the 5' and 3' side of the locus III, respectively. The *aapA* gene is in red. All the ORF contained in the ICE Hptfs3 are shown in orange. Left (LJ) and right (RJ) junctions of the ICE are indicated.

*lori* complete genomes analyzed, the *aapA*/IsoA modules were present in one or multiple copies at the 5 conserved loci in all but 6 genomes that are free of intact *aapA*/IsoA systems (Figure 7A and Supplementary Table S4). We then hypothesized that, given the importance of mRNA folding in the control of the AapA1 toxin expression, this folding should be conserved. We thus carried out a search for structural mRNA homologs against all available *Helicobacter* complete genomes and other closely related species such as *Campylobacter* (Supplementary Figure S13). This search identified many new loci that were all manually inspected for features such as mRNA folding, presence of a conserved SD sequence, start and stop codons, as well as a promoter for the antisense RNA (Supplementary Figure S13). This manual inspection allowed us to dismiss a considerable fraction of detected loci. Each new positive locus was implemented for the next search in an iterative manner (see the detail of this search in Supplementary Figure S13). Interestingly, this new search identified many *AapA* mRNA homologs in other *Helicobacter* species that were not found with the TBLASTN or PSI-BLAST searches. Since the synteny was not conserved in these *Helicobacter* species, these systems were named *aapA*/IsoA without any specific number associated (Supplementary Table S4). Surprisingly, this search revealed the presence of this system near prophage encoding genes in *H. felis* ATC 49179 and *H. bizzozeronii* CIII-1 or near a transposon in *H. cetorum* MIT-99 5656 (Supplementary Figure S14). A closer look at the *H. pylori* loci revealed that some of them were associated with another type of MGE such as the ICE recently characterized in *H. pylori* genomes (25,26). For instance in Gambia94/24, a complete ICE of 46 kb named ICEHptfs3 was present in the intergenic region corresponding to the locus III, creating a much larger genomic region (Figure 7B). Finally, although no *aapA*/IsoA system was detected in *Helicobacter* plasmids, one was identified on a *Campylobacter jejuni* plasmid as well on the chromosome of some *C. jejuni* strains in a region involved in plasmid stabilization (Supplementary Figure S15).

Altogether these results indicate that while these TA systems have been discovered initially on the chromosome of *H. pylori*, their localization near MGEs such as ICE, transposon and prophage suggests that they might have been acquired via horizontal gene transfer and played a role in stabilizing these mobile elements.

## DISCUSSION

In this article we characterize for the first time a type I toxin-antitoxin system in *H. pylori*. We have shown that the *aapA1* gene encodes a small peptide whose expression is toxic to *H. pylori* and found that several layers of post-transcriptional regulation are preventing the expression of this toxin, as described in our working model in Figure 8. The transcription of the *aapA1* gene generates a full-length mRNA of 250 nt, denoted *AapA1\_FL* that is translationally inactive. This inert form is constitutively processed at its 3' end to generate a truncated transcript of 225 nt which is translatable. This active form denoted *AapA1\_Tr1*, base pairs with the IsoA1 antisense RNA to form an extended duplex that is rapidly degraded by RNase III to generate



**Figure 8.** Model of the *aapA1*/IsoA1 TA system regulation. This model is based on all experimental data presented and is described in the text (see discussion). Simplified structures of *AapA1* transcripts (FL, Tr1 and Tr2) are based on Figure 5. The region encoding the AapA1 ORF in the transcript is shown with a thick black line. The Shine-Dalgarno sequence is indicated as SD. The translationally active and inactive states of the different mRNA species are also indicated.

*AapA1\_Tr2*. In absence of IsoA1, the synthesis of AapA1 toxin can occur leading to a growth arrest and cell death.

## Metastable structures prevent premature toxin expression during mRNA synthesis

While most TA systems including type II, III, IV and V, have adopted a specific operon organization allowing the production of the antitoxin before that of the toxin, type I TA systems are using an alternative strategy to prevent toxin production. Most type I toxin encoding mRNAs have their SD sequence sequestered in a stable double-stranded region that prevents expression of the toxin while the RNA is transcribed (20,27–31). In some cases, this sequestration involves a long distance base-pairing between the 5'- and 3'-UTRs (31,32). However due to transcription-translation coupling that occurs in bacteria, the RBS could be, in principle, available before the 3' region is synthesized. This particular mode of regulation implies that metastable structures sequestering the RBS are formed during transcription in order to prevent translation initiation on the nascent mRNA (31,32). To predict such transient secondary structures that are formed during transcription, we used the Kinefold stochastic simulations that predict co-transcriptional folding paths of functional RNAs (33). This simulation showed that three successive structures masking the RBS can potentially form at various stages of mRNA transcription (Supplementary Figure S16 and Video SM1 for *AapA1\_FL* mRNA, data not shown for the other loci). A more stable structure replaces the first two metastable struc-

tures when transcription reaches the end of the message. Besides sequestration of the RBS, the 5' to 3' end-pairing of the full-length mRNA forms an 11 bp helix including 6 GC pairs (Figure 5) that induces an efficient transcription termination at this position. Another putative terminator stem-loop was previously annotated at the position +225 (4) but our work shows that it is not able to induce transcription termination very efficiently. In addition, the IsoA1 could, in principle, bind to the nascent AapA1 mRNA, before the RNA polymerase reaches the end of the message. This interaction could then promote the production of the Tr1 transcript. However, we believe this scenario quite unlikely because the *AapA1-Tr1* is only visible after rifampicin treatment at late stages when IsoA1 is completely degraded. In addition we would never observed accumulation of the full length mRNA. Altogether our results clearly indicate that IsoA1 is required for the degradation of Tr1 but not for its production.

#### ***AapA1* mRNA requires a 3' end processing to be translated**

For type I toxin mRNAs having their SD sequence sequestered, the mechanism by which the toxin mRNA gains activation is still poorly documented. Our results show that for *AapA1*, a 3' end-processing event is required to refold the mRNA and to unmask the SD (Figure 5). Until now, two examples of type I TA systems for which the toxin is produced after the cleavage of the mRNA have been described (either in 5' for TisB or 3' end for Hok (20,34)). In contrast to the 3' end processing involved in the maturation of the *Hok* mRNA, the 3' to 5' exonucleolytic activity of the polynucleotide phosphorylase and of the ribonuclease II is not sufficient to process the 3' end of the *AapA1-FL* mRNA (H. Arnion, unpublished results). As for TisB (20), the enzymatic activity responsible for the processing remains to be identified.

#### **IsoA1 inhibits toxin expression primarily at the translational level**

In order to avoid toxin expression, IsoA1 is constitutively synthesized in a large excess over *AapA1-Tr1*. The double stem-loop structure in the 5'UTR of *AapA1-Tr1* mRNA mediates a specific recognition by the IsoA1 RNA via loop-loop interactions. *In vitro* this hybrid is then extended into a full duplex that is cleaved by the double-stranded specific RNase III. *In vivo* this degradation is very rapid since the *AapA1-Tr1* does not accumulate. However, RNase III is not essential in the B128 strain despite the presence of several *aapA*/IsoA systems in its genome, suggesting that, at least in this strain, RNase III cleavage is not mandatory for the repression of the toxin. The binding of IsoA1 to *AapA1-Tr1* is completely masking the RBS and in absence of RNase III, the stability of the long duplex formed (76 bp) is probably sufficient to prevent ribosome binding. Thus, the repression of AapA1 toxin expression by IsoA1 RNA may primarily occur at the translational level, as previously shown for other type I toxins (i.e. Hok (35), TisB (20), Fst (29) and BsrG (28)). In contrast, in *B. subtilis*, the degradation of the TxpA and YonT toxin mRNAs by RNase III is essential to prevent toxin synthesis (27,36).

#### **Analogy with other type I TA systems**

This new system identified in *epsilon* proteobacter species shares many similarities with other type I TA previously identified in enterobacteria. For instance the *aapA1*/IsoA1 system share common features with the overlapping *cis*-encoded antisense RNA regulating TA systems characterized in *E. coli* (37). Similar to *hok*/Sok and *ldr*/Rdl TA systems, this system encodes a small toxic protein (size < 50 aa) that is synthesized from a highly stable mRNA and which production is repressed by a small unstable antisense RNA. In addition, as for *hok*/Sok, the *aapA1*/IsoA1 requires a 3' processing to get activated (34). However, in contrast to these two TA systems, its translation activation does not seem to require the translation of an overlapping small ORF (like the mok and ldrX peptides of the *hok*/Sok and *ldr*/Rdl systems, respectively) (38,39). Another major difference with other TA systems relies on the fact that the IsoA1 antitoxin is directly targeting the translation initiation region of the toxin gene. This specificity is due to the genomic organization of this TA system, where the antitoxin is fully complementary to the 5' UTR of the toxin-encoding mRNA, as also shown for the *symE*/SymR TA system in *E. coli* (40).

#### **Conservation of *aapA*/IsoA TA systems in *Helicobacter* and *Campylobacter***

Few reports have highlighted the difficulties in predicting the presence of type I TA system in bacterial genomes (24,41). Indeed the large diversity of toxin sequences as well as their small nucleotide size complicates the finding of new type I TA systems. In this report we present a new strategy taking advantage of the specific folding adopted by the AapA mRNA to identify more representatives of these systems. Our search revealed the presence of a large number of these systems in *Helicobacter* genomes, larger than anticipated (4). With the exception of a few strains that only have ghost copies of *aapA*/IsoA (i.e. with either no functional promoter and/or no SD and/or no ORF longer than 5–6 amino acids), most *H. pylori* strains contain one or multiple copies of these systems at conserved loci on the chromosome. At a given locus, they can be present in single or multiple copies, depending on the strains. The locus IV, located close to the replication origin is the most conserved one, as it is found in all the strains analyzed. Our data show that these TA systems are sometimes inactivated by point mutations, genomic rearrangements or insertion of an IS (insertion sequence) element. This is similar to what was observed for the *hok*/Sok systems present on the chromosome of *E. coli* (41). However, we also found that in some cases, the *aapA*/IsoA systems are associated with MGE. For instance our search reveals their presence on *Campylobacter* plasmids and within genomic islands coding either for an ICE element (strains shown in bold in Figure 7A) or prophage genes (*Helicobacter felis* ATCC49179 and *Helicobacter bizzozeroni* MIT 99–5656). In these cases, these newly identified TA systems may be part of a plasmid stabilization and/or addiction system as was reported for several TA systems (42). This phenomenon, also known as post-segregational killing, was first described for *hok*/Sok system (43). If a plasmid bearing a TA module is not transmitted to



a daughter cell, the unstable antitoxin is degraded while the stable toxin acts on cellular targets and kill the plasmid-free cells. For instance type II systems have already been shown to promote the maintenance of ICE element and conjugative plasmids (44,45). Whether a similar role could be played by the *aapA*/IsoA systems remains to be explored.

## CONCLUSION

To conclude, the characterization of a new TA system in *H. pylori* revealed the importance of mRNA folding in type I TA regulation. We propose that the *AapA* mRNA requires two structural RNA switches to gain activation. The first one consists of a spontaneous conformational rearrangement, in which metastable 5'-end structures are disrupted in favor of a more stable 5' end/3' end interaction. The second switch is induced after a processing event done by a yet unknown RNase, which renders the mRNA translationally active. This sophisticated mechanism may be conserved for other TA systems. Another interesting characteristic of the *aapA*/IsoA system, is the high stability of the mRNA versus the instability of the antitoxin that makes it a particularly efficient addiction module. Indeed, the cell should not be able to get rid of such TA system unless a non-sense mutation is present in the toxin gene. Finally our bioinformatic analysis revealed that the *aapA*/IsoA system belongs to a novel large family of type I TA system that is not only present on the chromosome but also associated with MGEs. Although our data provide some hints for a role in stabilizing MGEs, these systems might have diverse functions depending on their genetic localization or on the organisms that host them.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENT

The authors thank Denis Dupuy, Axel Innis and Cathy Staedel for critical reading of the manuscript and all present and past members of the ARNA laboratory for helpful discussions.

## FUNDING

INSERM ; Université de Bordeaux; Agence Nationale de la Recherche (<http://www.agence-nationale-recherche.fr/>) [ANR-12-BSV5-0025-Bactox1, ANR-12-BSV6-007-asSUPYCO]; European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement [642738]. Funding for open access charge: INSERM.

*Conflict of interest statement.* None declared.

## REFERENCES

- Cover, T.L. and Blaser, M.J. (2009) *Helicobacter pylori* in health and disease. *Gastroenterology*, **136**, 1863–1873.
- Ferlay, J., Shin, H.-R., Bray, F., Forman, D., Mathers, C. and Parkin, D.M. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer*, **127**, 2893–2917.
- Pernitzsch, S.R. and Sharma, C.M. (2012) Transcriptome complexity and riboregulation in the human pathogen *Helicobacter pylori*. *Front. Cell. Infect. Microbiol.*, **2**, 14.
- Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiß, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R. *et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, **464**, 250–255.
- Pernitzsch, S.R., Tirier, S.M., Beier, D. and Sharma, C.M. (2014) A variable homopolymeric G-repeat defines small RNA-mediated posttranscriptional regulation of a chemotaxis receptor in *Helicobacter pylori*. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E501–510.
- Goeders, N. and Van Melderen, L. (2014) Toxin-antitoxin systems as multilevel interaction systems. *Toxins*, **6**, 304–324.
- Wen, J. and Fozo, E.M. (2014) sRNA antitoxins: more than one way to repress a toxin. *Toxins*, **6**, 2310–2335.
- Brantl, S. and Jahn, N. (2015) sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.*, **39**, 413–427.
- Gerdes, K., Gulyaev, A.P., Franch, T., Pedersen, K. and Mikkelsen, N.D. (1997) Antisense RNA-regulated programmed cell death. *Annu. Rev. Genet.*, **31**, 1–31.
- Gerdes, K. and Maisonneuve, E. (2012) Bacterial persistence and toxin-antitoxin loci. *Annu. Rev. Microbiol.*, **66**, 103–123.
- Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A. *et al.* (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*, **388**, 539–547.
- Farnbacher, M., Jahns, T., Willrodt, D., Daniel, R., Haas, R., Goemann, A., Kurtz, S. and Rieder, G. (2010) Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. *BMC Genomics*, **11**, 335.
- McClain, M.S., Shaffer, C.L., Israel, D.A., Peek, R.M. and Cover, T.L. (2009) Genome sequence analysis of *Helicobacter pylori* strains associated with gastric ulceration and gastric cancer. *BMC Genomics*, **10**, 3.
- Ermak, T.H., Giannasca, P.J., Nichols, R., Myers, G.A., Nedrud, J., Weltzin, R., Lee, C.K., Kleanthous, H. and Monath, T.P. (1998) Immunization of mice with urease vaccine affords protection against *Helicobacter pylori* infection in the absence of antibodies and is mediated by MHC class II-restricted responses. *J. Exp. Med.*, **188**, 2277–2288.
- Skouloubris, S., Thiberge, J.M., Labigne, A. and De Reuse, H. (1998) The *Helicobacter pylori* UreI protein is not involved in urease activity but is essential for bacterial survival in vivo. *Infect. Immun.*, **66**, 4517–4521.
- Stingl, K., Brandt, S., Uhlemann, E.-M., Schmid, R., Altendorf, K., Zeilinger, C., Ecobichon, C., Labigne, A., Bakker, E.P. and de Reuse, H. (2007) Channel-mediated potassium uptake in *Helicobacter pylori* is essential for gastric colonization. *EMBO J.*, **26**, 232–241.
- Goodwin, A., Kersulyte, D., Sisson, G., Veldhuyzen van Zanten, S.J., Berg, D.E. and Hoffman, P.S. (1998) Metronidazole resistance in *Helicobacter pylori* is due to null mutations in a gene (*rdxA*) that encodes an oxygen-insensitive NADPH nitroreductase. *Mol. Microbiol.*, **28**, 383–393.
- Boneca, I.G., Ecobichon, C., Chaput, C., Mathieu, A., Guadagnini, S., Prévost, M.-C., Colland, F., Labigne, A. and de Reuse, H. (2008) Development of inducible systems to engineer conditional mutants of essential genes of *Helicobacter pylori*. *Appl. Environ. Microbiol.*, **74**, 2095–2102.
- Backert, S., Kwok, T. and König, W. (2005) Conjugative plasmid DNA transfer in *Helicobacter pylori* mediated by chromosomally encoded relaxase and TraG-like proteins. *Microbiol. Read. Engl.*, **151**, 3493–3503.
- Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E.G.H. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell*, **26**, 381–392.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C. *et al.* (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.
- Reuter, J.S. and Mathews, D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 129.

## 14 Nucleic Acids Research, 2017

23. Bernacchi, S., Freisz, S., Maechling, C., Spiess, B., Marquet, R., Dumas, P. and Ennifar, E. (2007) Aminoglycoside binding to the HIV-1 RNA dimerization initiation site: thermodynamics and effect on the kissing-loop to duplex conversion. *Nucleic Acids Res.*, **35**, 7128–7139.
24. Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V. and Storz, G. (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res.*, **38**, 3743–3759.
25. Fischer, W., Breithaupt, U., Kern, B., Smith, S.I., Spicher, C. and Haas, R. (2014) A comprehensive analysis of *Helicobacter pylori* plasticity zones reveals that they are integrating conjugative elements with intermediate integration specificity. *BMC Genomics*, **15**, 310.
26. Kersulyte, D., Lee, W., Subramaniam, D., Anant, S., Herrera, P., Cabrera, L., Balqui, J., Barabas, O., Kalia, A., Gilman, R.H. *et al.* (2009) *Helicobacter Pylori*'s plasticity zones are novel transposable elements. *PLoS One*, **4**, e6859.
27. Durand, S., Jahn, N., Condon, C. and Brantl, S. (2012) Type I toxin-antitoxin systems in *Bacillus subtilis*. *RNA Biol.*, **9**, 1491–1497.
28. Jahn, N. and Brantl, S. (2013) One antitoxin—two functions: SR4 controls toxin mRNA decay and translation. *Nucleic Acids Res.*, **41**, 9870–9880.
29. Shokeen, S., Patel, S., Greenfield, T.J., Brinkman, C. and Weaver, K.E. (2008) Translational regulation by an intramolecular stem-loop is required for intermolecular RNA regulation of the par addiction module. *J. Bacteriol.*, **190**, 6076–6083.
30. Wen, J., Won, D. and Fozo, E.M. (2014) The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin. *Nucleic Acids Res.*, **42**, 1930–1946.
31. Møller-Jensen, J., Franch, T. and Gerdes, K. (2001) Temporal translational control by a metastable RNA structure. *J. Biol. Chem.*, **276**, 35707–35713.
32. Nagel, J.H., Gultyaev, A.P., Gerdes, K. and Pleij, C.W. (1999) Metastable structures and refolding kinetics in hok mRNA of plasmid R1. *RNA*, **5**, 1408–1418.
33. Xayaphoummine, A., Bucher, T. and Isambert, H. (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.*, **33**, W605–W610.
34. Thisted, T., Nielsen, A.K. and Gerdes, K. (1994) Mechanism of post-segregational killing: translation of Hok, SrnB and Pnd mRNAs of plasmids R1, F and R483 is activated by 3'-end processing. *EMBO J.*, **13**, 1950–1959.
35. Gerdes, K., Nielsen, A., Thorsted, P. and Wagner, E.G. (1992) Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable hok, srnB and pndA effector messenger RNAs. *J. Mol. Biol.*, **226**, 637–649.
36. Durand, S., Gilet, L. and Condon, C. (2012) The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet.*, **8**, e1003181.
37. Kawano, M. (2012) Divergently overlapping cis-encoded antisense RNA regulating toxin-antitoxin systems from *E. coli*: hok/sok, ldr/rdl, symE/symR. *RNA Biol.*, **9**, 1520–1527.
38. Thisted, T. and Gerdes, K. (1992) Mechanism of post-segregational killing by the hok/sok system of plasmid R1. Sok antisense RNA regulates hok gene expression indirectly through the overlapping mok gene. *J. Mol. Biol.*, **223**, 41–54.
39. Kawano, M., Oshima, T., Kasai, H. and Mori, H. (2002) Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*. *Mol. Microbiol.*, **45**, 333–349.
40. Kawano, M., Aravind, L. and Storz, G. (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol. Microbiol.*, **64**, 738–754.
41. Pedersen, K. and Gerdes, K. (1999) Multiple hok genes on the chromosome of *Escherichia coli*. *Mol. Microbiol.*, **32**, 1090–1102.
42. Van Melderen, L. and Saavedra De Bast, M. (2009) Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet.*, **5**, e1000437.
43. Gerdes, K., Thisted, T. and Martinussen, J. (1990) Mechanism of post-segregational killing by the hok/sok system of plasmid R1: sok antisense RNA regulates formation of a hok mRNA species correlated with killing of plasmid-free cells. *Mol. Microbiol.*, **4**, 1807–1818.
44. Carraro, N., Poulin, D. and Burrus, V. (2015) Replication and active partition of integrative and conjugative elements (ICEs) of the SXT/R391 family: the line between ICEs and conjugative plasmids is getting thinner. *PLoS Genet.*, **11**, e1005298.
45. Wozniak, R.A.F. and Waldor, M.K. (2009) A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genet.*, **5**, e1000439.
46. Darty, K., Denise, A. and Ponty, Y. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.



# Curriculum vitae

## Education

- 2010-2014 **Bachelor Degree** in Biology, University of Cordoba, Spain. Final grade 8.84 (0-10)
- 2014-2015 **Master Degree** in Cellular and Molecular Biotechnology and Genetics, University of Cordoba, Spain. Final grade 9.80 (0-10) Honour mention
- 2015-2018 **Ph.D. in Biochemistry**, University of Bordeaux, France



## Research Experience

### Current Position

From Oct. 2015 Marie-Curie Horizon 2020 ITN Ph.D. student. Supervisor: **Dr. Fabien Darfeuille**. Thesis: **Understanding RNA-based gene expression regulation in bacteria: lessons from type I toxin-antitoxin systems** University of Bordeaux/INSERM U1212, Bordeaux, France

### Previous Positions

- 2011-2015 Internship student at the Molecular Genetics of Fungal Pathogenicity Group. Supervisor: **Pr. Dr. Antonio Di Pietro**
- 2013-2014 Bachelor student, supervisor: **Pr. Dr. Antonio Di Pietro**. Thesis: *Role of alkalizing secreted effectors during fungal infection*. Score: 9.2 (0-10). Mention: excellent
- 2014-2015 Master student, supervisor: **Pr. Dr. Antonio Di Pietro**. Thesis: *Role of F-RALF, an effector protein that promotes virulence in the plant pathogenic fungus F. oxysporum* Score: 10 (0-10). Mention: excellent with honours

## Awards & Honours

- 2012-2014 **Merit-based Undergraduate Research Fellowship**. Funding bachelor thesis. Supervisor: Pr. Dr. Antonio Di Pietro. Financial backer: University of Cordoba
- 2014-2015 **Merit-based Undergraduate Research Fellowship**. Funding master thesis project. Supervisor: Pr. Dr. Antonio Di Pietro, UCO. Financial backer: Government of Spain
- 2015 **Bachelor Extraordinary award**. Best student of the promotion 2010/2014, UCO
- 2016 **Master Extraordinary award**. Best student of the promotion 2014/2015, UCO
- 2015-2018 **Innovative Training Network (ITN) PhD grant "MetaRNA"**. Thesis director: Dr. Fabien Darfeuille. Institution: University of Bordeaux/INSERM, Bordeaux, France. Financial backer: European Union's Horizon 2020 under the Marie Skłodowska-Curie Grant agreement N°642738

## Training activities & formation (selected courses)

- 2017 Workshop: Scientific writing (10h), University of Copenhagen
- 2017 Intense PhD-course: Science-Based Entrepreneurship (5 ECTS ≈ 125 h), DTU
- 2017 Workshop: Advanced light microscopy (1 ECTS ≈ 25 h), University of Groningen
- 2017 Workshop: Industrial product and process development/working in Industry (1 ECTS ≈ 25 h), University of Delft, Netherlands

---

## Publication List

- I. **Masachis S**, Segorbe D, Turrà D, Leon-Ruiz M, Fürst U, El Ghalid M, Leonard G, López-Berges MS, Richards TA, Felix G, Di Pietro (2016) *A fungal pathogen secretes plant alkalizing peptides to increase infection*. **Nature Microbiology** 11;1(6):16043
- II. Arnion H, Korkut DN, **Masachis Gelo S**, Chabas S, Reignier J, Iost I, Darfeuille F (2017) *Mechanistic insights into type I toxin-antitoxin systems in Helicobacter pylori: the importance of mRNA folding in controlling toxin expression*. **Nucleic Acids Research** 5;45(8):4782-4795
- III. **Masachis S**, Darfeuille F (2018) *Type I toxin-antitoxin systems: regulating toxin expression via Shine-Dalgarno sequestration and small RNA binding*. **Microbiol Spectr.** 2018 Jul;6(4). doi: 10.1128/microbiolspec.RWR-0030-2018.
- IV. **Masachis S**, Tourasse NJ, Chabas S, Bouchez O, Darfeuille F (2018) *FASTBAC-Seq: Functional Analysis of Toxin-antitoxin systems in BACteria by deep-sequencing*. **Methods in Enzymology: High-Density Sequencing Applications in Microbial Molecular Genetics, Vol. 612**. Edited by **A.J. Carpoussis**. <https://doi.org/10.1016/bs.mie.2018.08.033>
- V. **Masachis S**, Tourasse NJ, Lays C, Faucher M, Chabas S, Iost I, Darfeuille F. *A genetic selection of toxicity suppressors reveals functional metastable structures embedded in a toxin-encoding mRNA* (in preparation)
- VI. **Masachis S**, Arnion H, Chabas S, Boissier F, Darfeuille F. *When mRNA folding rules decay: lessons from a type I toxin-antitoxin system* (in preparation)

## References

### Ph.D. thesis supervisor

**Dr. Fabien Darfeuille**  
ARNA Laboratory  
INSERM U1212/CNRS 5320

University of Bordeaux, Bat. 3A, 1<sup>er</sup> ét.  
33076 Bordeaux (France)  
146 rue Léo Saignat  
[fabien.darfeuille@inserm.fr](mailto:fabien.darfeuille@inserm.fr)  
(+33) (0) 55757 4565

### Bachelor/ Master thesis supervisor

**Pr. Dr. Antonio Di Pietro**  
Molecular Genetics of Fungal  
Pathogenicity Group  
Department of Genetics

University of Cordoba  
Campus Rabanales, Edf. Gregor Mendel  
14071 Cordoba (Spain)  
Carretera Nacional IV Km 396  
[ge2dipia@uco.es](mailto:ge2dipia@uco.es)  
(+34) 957 21 89 81