



HAL
open science

Étude du mécanisme de synchronisation de la réplication des deux chromosomes de *Vibrio cholerae*

Théophile Niault

► **To cite this version:**

Théophile Niault. Étude du mécanisme de synchronisation de la réplication des deux chromosomes de *Vibrio cholerae*. Biochimie [q-bio.BM]. Sorbonne Université, 2023. Français. NNT : 2023SORUS121 . tel-04142724

HAL Id: tel-04142724

<https://theses.hal.science/tel-04142724>

Submitted on 27 Jun 2023

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.



Sorbonne Université

École doctorale – Complexité du Vivant
Institut Pasteur – Unité Plasticité du Génome Bactérien

**Initiation de la réplication du
chromosome secondaire de
*Vibrio cholerae***

Par Théophile Niault
Thèse de doctorat de Génétique Bactérienne

Dirigée par Marie-Ève Val et Didier Mazel

Présentée et soutenue le 21 février 2023

Devant le jury composé de :

Pr. Guennadi Sezonov
Pr. Justine Collier
Dr. Jean-Yves Bouet
Dr. Géraldine Laloux
Dr. Mathilde Guzzo
Dr. François Xavier Barre
Dr. Marie-Eve Val
Pr. Didier Mazel

Président du jury
Rapporteuse
Rapporteur
Examinatrice
Examinatrice
Examinateur
Co-directrice de thèse
Co-directeur de thèse

Résumé :

Contrairement à la grande majorité des bactéries qui possèdent un chromosome unique, *Vibrio cholerae*, l'agent pathogène responsable du choléra, contient un génome divisé en deux réplicons : un chromosome principal (Chr1) et un chromosome secondaire (Chr2). Les chromosomes secondaires sont apparus indépendamment dans de nombreux phylum bactériens et sont souvent associés à des espèces pathogènes ou symbiotiques (par exemple *Leptospira*, *Agrobacterium*, *Burkholderia*). Les chromosomes secondaires résultent à priori de l'acquisition et de la domestication de mégaplasmites. En effet, certains "mégaplasmites" ont cohabité avec leur hôte au cours d'une longue période évolutive et sont devenus un composant obligatoire du génome de leur hôte, essentiel à leur croissance. Au cours de leur domestication, les chromosomes secondaires ont acquis différentes caractéristiques permettant leur maintien à travers les divisions cellulaires assurant leur transfert vertical aux générations suivantes. L'étude des bactéries à génome multipartite permet de s'interroger sur le bénéfice sélectif d'un organisme à posséder plusieurs chromosomes. Ou encore sur comment un élément génétique acquis horizontalement peut-il se connecter à un réseau génétique préétabli.

L'une des principales étapes de la domestication des mégaplasmites est leur intégration dans le cycle cellulaire de l'hôte. Chez *V.cholerae*, la réplication de Chr2 est coordonnée avec celle du Chr1. Chr1, qui est trois fois plus grand que Chr2, est initié en premier et lorsque les 2/3 de la molécule a été répliquée, un signal d'initiation déclenche la réplication du Chr2 de telle sorte que les deux chromosomes terminent leur réplication de façon synchrone. Ce signal d'initiation est émis lorsque la fourche de réplication passe par un locus non codant présent sur le Chr1 nommé *crtS* (Chr2 Replication Triggering Site). Cependant, le mécanisme moléculaire à l'origine de ce signal d'initiation reste encore à élucider. Ce travail de thèse vise à comprendre le mécanisme de ce nouveau checkpoint bactérien.

Sommaire :

Table des matières

Résumé :	3
Sommaire :	5
Introduction	8
I. Organisation du génome bactérien	10
1. Chromosome : le réplicon principal de la cellule bactérienne.....	11
2. Plasmides et mégaplasmides : des réplicons non essentiels	12
3. Chromosomes secondaires et chromides : des réplicons porteurs de gènes de ménages	13
4. Théorie de la formation des chromides : une domestication de mégaplasmide	13
II. La réplication du chromosome bactérien	15
1. Cycle cellulaire et taux de croissance	16
2. La masse d'initiation ou volume critique	18
3. Découverte de DnaA l'initiateur universel bactérien et de l'origine de réplication <i>oriC</i>	18
4. DnaA : une protéine de la famille AAA+ ATPase.....	20
5. Ouverture de l'ADN au niveau de l' <i>oriC</i>	21
6. Inactivation d' <i>oriC</i> post-initiation : séquestration par SeqA.....	23
7. Inactivation de DnaA : RIDA et DDAH	25
8. Régénération de DnaA-ATP	27
9. Chargement du replisome lors de l'initiation de la réplication à <i>oriC</i>	29
10. Étape d'élongation lors de la réplication de l'ADN	31
11. Terminaison de la réplication : le système Ter/Tus d' <i>E.coli</i>	32
III. Réplication des chromosomes secondaires bactériens	34
1. Organisation des origines de type RepABC.....	35
2. La réplication des origines à itérons	37
3. Structure de l'origine d'initiation de la réplication	37
4. Comparaison de l'ouverture de l'ADN entre chromosome et plasmides à itérons	39
5. Structure des protéines de la famille Rep.....	40
6. <i>Handcuffing</i> : mécanisme principal de la régulation négative de la réplication.....	42
IV- <i>Vibrio cholerae</i> : un modèle d'étude des bactéries à génome multipartite	46
1. Un génome bipartite conservé chez les <i>Vibrionaceae</i>	47
2. Initiation du Chr1 par un mécanisme <i>oriC-DnaA like</i>	49
3. Le Chr1 et le Chr2 suivent une chorégraphie d'initiation	51
4. Activation de l'initiation du Chr2	52
5. Contrôle par la méthylation de la réplication de Chr2.....	54
6. Régulation négative de l'origine du chromosome 2	56
7. Coordination de la terminaison de la réplication du Chr1 et du Chr2	59
8. <i>crtS</i> et <i>ori2</i> : à proximité dans la cellule	61
9. Implication du régulateur global Lrp dans l'activité de <i>crtS</i>	62
10. RctB : une protéine Rep avec deux domaines supplémentaires	63
11. Similarité structurelle entre Rep et RctB	65
12. Comparaison entre la régulation de la réplication du Chr2 et des plasmides à itérons.....	66
13. Chargement du replisome, élongation et terminaison de la réplication des deux chromosomes..	68
Questions et hypothèses du projet de thèse	71
Résultats	75
I. Premier article: "The coordinated replication of <i>Vibrio cholerae</i>'s two chromosomes required the acquisition of a unique domain by the RctB initiator"	75

II. Second article: “Genome-wide binding of secondary chromosome initiator reveals mechanism for replication coordination in <i>Vibrio cholerae</i>”	91
III. Résultats non publiés : étude du rôle de Lrp dans le mécanisme <i>crtS-ori2</i>.....	123
1. Introduction : Lrp un régulateur global.....	123
2. Rôle de Lrp (Leucine responsive Regulatory Protein) dans le mécanisme <i>crtS-ori2</i>	125
3. Lrp et RctB sur <i>crtS</i> : une fixation coopérative.....	127
4. CryoEM : observations du complexe RctB-Lrp- <i>crtS</i>	128
5. ChIPseq de RctB dans un contexte Δlrp	130
6. ChIPseq de Lrp chez <i>Vibrio cholerae</i>	131
7. Effet de la phase de croissance sur le pattern de fixation de Lrp	133
8. ChIPseq de Lrp en population synchronisée	134
9. L'état d'oligomérisation de Lrp dans l'action de <i>crtS</i>	137
10. Lrp régule un tiers des gènes de <i>Vibrio cholerae</i>	138
IV. Résultats non publiés : RctB et modification post-traductionnelle	142
V. Résultats non publiés : outil de synchronisation de la réplication chez <i>Vibrio cholerae</i>.....	147
<i>Conclusions générales</i>.....	157
<i>Discussion</i>	161
<i>crtS-ori2</i> : une action directe et locale ?	161
Quel rôle pour le passage de la fourche sur <i>crtS</i> ?	163
Organisation du complexe d'inhibition de l' <i>ori2</i>	164
Comment dissocier le complexe d'inhibition?.....	166
A la recherche de l'ancêtre du Chr2 : mégaplasmides portant des initiateurs RctB-like.....	167
Terminaison synchrone : un mécanisme commun aux chromides ?.....	168
<i>Références</i>	170

INTRODUCTION

Introduction

L'hérédité est le processus par lequel les caractères génétiques sont transmis d'une génération à l'autre. Cela permet aux organismes vivants de transmettre leurs traits à leurs descendants et de maintenir la diversité de la vie sur Terre. La théorie de l'hérédité a été formulée pour la première fois par Charles Darwin dans son ouvrage "De l'origine des espèces", publié en 1859. Darwin propose alors que les organismes évoluent au fil des générations grâce au processus de sélection naturelle. Selon cette théorie, les organismes qui sont mieux adaptés à leur environnement ont plus de chances de survivre et de se reproduire, ce qui permet à leurs traits de se propager dans la population.

En 1865, le moine austro-hongrois Gregor Mendel a publié des résultats de recherches sur l'hérédité chez les petits pois qui ont permis de mieux comprendre comment ces traits sont transmis. Mendel a montré que les traits sont transmis par des unités fonctionnelles appelées gènes. La question du support de l'hérédité commence alors à émerger.

Au début des années 1900, des chercheurs tels que Theodor Boveri et Walter Flemming découvrent que les chromosomes des cellules animales, facilement observables au microscope, sont constitués d'ADN. Puis, dans les années 1920, le microbiologiste américain Fred Griffith démontre que des cellules bactériennes peuvent acquérir de nouvelles informations génétiques à partir de l'environnement en mélangeant des souches virulentes et des souches non virulentes de *Streptococcus pneumoniae*.

Dans les années 1940, les scientifiques canadiens Oswald Avery, Colin MacLeod et Maclyn McCarty ont réalisé des expériences démontrant que c'est l'ADN qui constitue le matériel génétique des bactéries. Toujours avec *Streptococcus pneumoniae*, ils ont démontré que lorsqu'uniquement l'ADN purifié de la souche virulente était donné à la souche non virulente, les cellules résultantes devenaient elles-mêmes virulentes.

Au début des années 1950, Watson et Crick commencent à travailler sur le problème de la structure de l'ADN. Ils se sont basés sur les travaux précédents de Rosalind Franklin et Maurice Wilkins, qui avaient utilisé la diffraction des rayons X pour étudier l'ADN. Franklin et Wilkins avaient obtenu des images de cristaux d'ADN cruciales pour comprendre sa structure. En utilisant ces images, Watson et Crick proposent une structure en double hélice pour l'ADN, dans un article désormais célèbre publié dans Nature en 1953. Selon cette structure, l'ADN

est formé par deux brins de nucléotides enroulés l'un autour de l'autre. Chaque brin est composé d'une chaîne de nucléotides, ces brins sont reliés entre eux par des liaisons de base complémentaires. Cette découverte permet d'élucider la base moléculaire de l'hérédité, car elle permet de stocker et de transmettre l'information génétique. Rapidement on comprend que la molécule d'ADN est répliquée de manière semi-conservative, un des brins sert de matrice et est recopié en respectant la complémentarité des bases.

En 1963, John Cairns publie une autoradiographie d'une lyse cellulaire de *E. coli*, en marquant au préalable l'ADN avec de la thymidine tritiée (Figure 1). On observe ainsi pour la première fois que ce génome bactérien semble être constitué d'un unique chromosome circulaire. De plus, il apparaît qu'il est vraisemblablement dupliqué à partir d'un unique point.

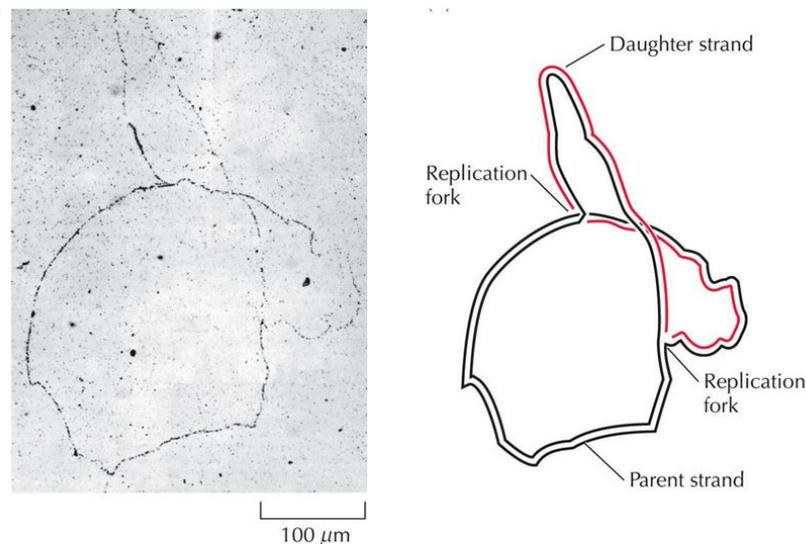


Figure 1: Le génome bactérien est composé d'un unique chromosome circulaire. Autoradiographie d'une lyse d'*E.coli* avec l'ADN marqué radioactivement réalisé par Cairns en 1963.

I. Organisation du génome bactérien

Après l'observation en 1963 du chromosome circulaire d'*E. coli* (Cairns, 1963) suivent d'autres études pour d'autres espèces bactériennes. Un dogme semblait alors se construire pour l'organisation du génome bactérien : un unique chromosome circulaire qui porte l'ensemble des gènes de ménage. Ce chromosome pouvait éventuellement être accompagné dans la cellule d'un autre type de réplicon bien plus petit en taille, également circulaire et encodant uniquement des gènes accessoires : les plasmides.

Cependant ce paradigme : génome bactérien = 1 chromosome (+ plasmides) va être progressivement remis en question. Tout d'abord avec la découverte de réplicons bactériens linéaires comme le chromosome linéaire de *Borrelia burgdorferi* (Baril et al., 1989). Puis par la découverte de réplicon secondaire de grande taille, notamment un plasmide de 460kpb chez *Sinorhizobium meliloti* (Rosenberg et al., 1981) alors nommé « mégaplasmide ». Enfin, en 1989 l'étude du génome de *Rhodobacter sphaeroides* constitue la première preuve d'une bactérie à deux chromosomes (Suwanto and Kaplan, 1989).

Suite à l'explosion des méthodes de séquençages à haut débit dans les années 2000 et l'accumulation de génomes bactériens entiers, on en arrive à la conclusion qu'environ 10% des espèces bactériennes présente un génome multipartite, c'est à dire composé de plusieurs chromosomes (Harrison et al., 2010). De manière intéressante, ces espèces bactériennes occupent généralement des niches écologiques particulières, on y retrouve des symbiotes et pathogènes de plantes, comme *Rhizobium* et *Agrobacterium*, également des pathogènes des animaux ou de l'homme, comme *Brucella*, *Vibrio* et *Burkholderia* (Figure 2).

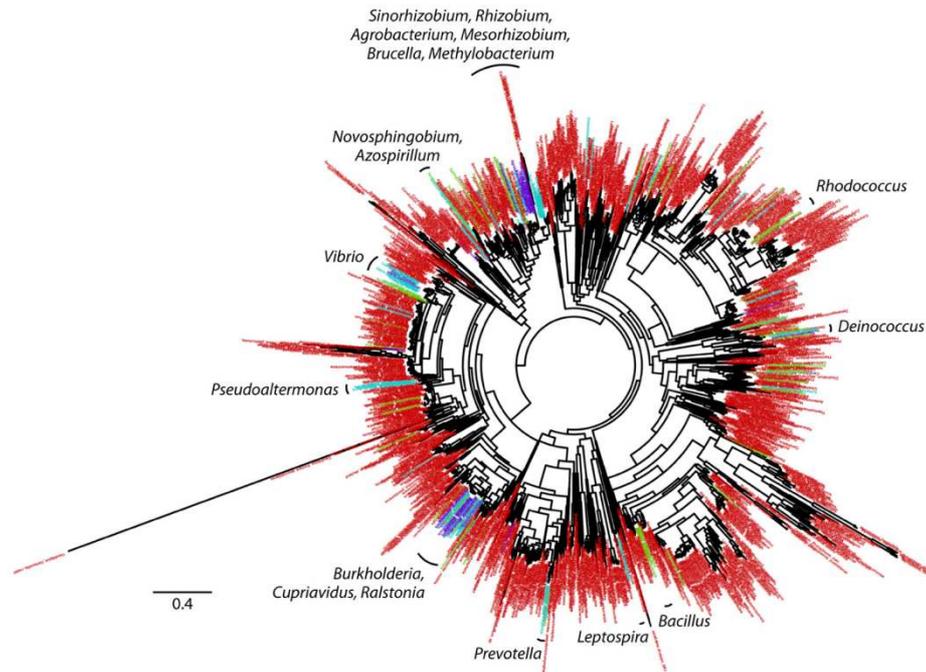


Figure 2. Distribution des génomes multipartites dans la phylogénie bactérienne. Une distribution phylogénétique de 1.708 espèces bactériennes. Les noms des taxons sont colorés en fonction de la structure du génome, avec le rouge pour les espèces sans mégaplasme ou chromosome, le vert pour les espèces avec un ou plusieurs mégaplasmes mais pas de chromosome, le bleu pour les espèces avec un ou plusieurs chromosomes mais pas de mégaplasme, et le violet pour les espèces avec à la fois un ou plusieurs mégaplasmes et un ou plusieurs chromosomes (diCenzo and Finan, 2017).

1. Chromosome : le réplicon principal de la cellule bactérienne

L'organisation du chromosome bactérien diffère grandement de celle des chromosomes eucaryotes. Taille : les chromosomes bactériens sont généralement beaucoup plus petits que les chromosomes trouvés chez les eucaryotes, avec une taille moyenne d'environ 3,5 million de paires de bases. Forme : les chromosomes bactériens sont très généralement circulaires, plutôt que linéaires comme pour les chromosomes eucaryotes. Organisation : le matériel génétique d'un chromosome bactérien est organisé en une seule molécule continue, plutôt que d'être séparé en chromosomes distincts comme chez les eucaryotes. La densité génétique : les chromosomes bactériens ont une densité génétique élevée, ce qui signifie qu'ils contiennent un grand nombre de gènes dans une petite quantité d'ADN. L'évolution : les chromosomes bactériens et eucaryotes sont soumis à diverses forces évolutives, notamment la mutation ou la recombinaison. Cependant chez les bactéries le chromosome est principalement influencé par le transfert horizontal de gènes, qui permet l'acquisition de nouveaux caractères et dirige l'évolution des espèces bactériennes.

Le chromosome contient généralement la totalité des gènes de ménage, c'est-à-dire les gènes nécessaires au maintien des fonctions cellulaires essentielles : réplication –

transcription – traduction – métabolisme, essentiels dans toutes les conditions de culture (Harrison *et al.*, 2010). Une réplication fidèle et une transmission verticale aux cellules filles est donc primordiale.

Le chromosome bactérien est défini comme le réplicon principal de la cellule, il est en effet toujours le plus grand réplicon de la cellule. Le plus grand chromosome séquencé est une souche de *Sorangium cellulosum*, à 14 782 125 pb, et contient 11 599 gènes (Han *et al.*, 2013). Le plus petit génome bactérien séquencé est celui d'une souche de *Candidatus Nasuia deltocephalinicola* ; ce génome ne code que 137 protéines et ne fait que 112 091 pb (Bennett *et Moran* 2013). La taille médiane des chromosomes bactériens est de 3,46 Mpb. Ainsi le chromosome contient la quasi-totalité du matériel génétique. Cependant, ce n'est pas toujours le cas. Les chromosomes de *Sinorhizobium meliloti* 1021 et *Burkholderia xenovorans* LB400 ne représentent respectivement que 54,6 % et 50,3 % de leurs génomes (Galibert *et al.*, 2001).

Parmi les 1708 espèces bactériennes dont le génome complet est disponible dans la base de données NCBI (étude réalisée en mars 2016), 1017 (~59,5%) contiennent uniquement un chromosome mais aucun réplicon secondaire (chromide, mégaplasmide ou plasmide), tandis que 192 (~11%) contiennent un chromide et/ou un mégaplasmide (diCenzo and Finan, 2017).

2. Plasmides et mégaplasmides : des réplicons non essentiels

Les plasmides et les mégaplasmides ne portent aucun gène de ménage et sont par conséquent non essentiels pour la viabilité des cellules pour une croissance en milieu riche. La majorité des gènes de ces réplicons ont été acquis lors d'évènements récents de transferts horizontaux. Leur contenu en GC diffère significativement de celui du chromosome de leur hôte (Harrison *et al.*, 2010). La distinction entre plasmide et mégaplasmide est actuellement arbitraire car uniquement basée sur la taille du réplicon.

En 2017, un seuil de 350kpb a été proposé pour séparer les plasmides et les mégaplasmides correspondant à 10% de la taille du chromosome moyen (diCenzo and Finan, 2017). En utilisant cette limite pour distinguer les deux types de réplicons, les auteurs ont alors trouvé une taille médiane de 46,2kpb pour les plasmides et de 558kpb pour les mégaplasmides. En termes de différences biologiques : les mégaplasmides sont

généralement présents en une seule copie dans la cellule et encodent leurs propres systèmes de partition (Hall et al., 2022).

A noter qu'il existe un seuil plus naturel pour séparer les petits et les grands plasmides qui ne repose pas sur une taille arbitrairement choisie (comme c'est le cas pour 350kpb). Celui-ci est la présence obligatoire d'au moins un système de partition de type ParAB, ParRM ou StbA dans tous les plasmides supérieurs à 25kpb (Planchenault et al., 2020).

3. Chromosomes secondaires et chromides : des réplicons porteurs de gènes de ménages

En 2010, Harrison *et al.* proposent le terme " chromide ", une combinaison des termes chromosome et plasmide, pour désigner un réplicon intermédiaire entre un plasmide et un chromosome (Harrison *et al.*, 2010). Les systèmes de réplication des chromides ressemblent à ceux des plasmides et des mégaplasmites. Cependant, contrairement à ceux-ci, les chromides portent au moins un gène de ménage essentiel à la viabilité cellulaire dans tous les milieux de culture. Ainsi les chromides sont particulièrement compliqués à identifier à partir de l'analyse de séquences sur les bases de données de génomes bactériens. Seule une preuve expérimentale de leur essentialité permet de départager un mégaplasmitide d'un chromide.

Néanmoins, les chromides présentent généralement des signatures génomiques, comme le taux de GC ou l'usage de codon, qui ressemblent davantage à celles du chromosome. De plus, ils sont généralement plus importants en taille, leur taille médiane étant de 1,26Mpb soit environ deux fois plus que celle des mégaplasmites (0,56Mpb) (diCenzo and Finan, 2017).

4. Théorie de la formation des chromides : une domestication de mégaplasmites

La théorie principale permettant d'expliquer l'apparition des chromides est la domestication d'un mégaplasmitide préalablement acquis par transfert horizontal. Selon cette théorie, une coévolution prolongée d'un mégaplasmitide avec un chromosome dans une même cellule bactérienne entraîne la régression de la signature génomique du mégaplasmitide (usage de codons, taux de GC) vers celle du chromosome. Pendant ce processus, un ou plusieurs gènes de ménages sont transmis depuis le chromosome vers le mégaplasmitide par

duplication. Enfin ces mêmes gènes sont perdus par délétion sur le chromosome rendant ainsi le mégaplasme essentiel.

Plusieurs arguments viennent appuyer ce modèle de domestication de mégaplasme (diCenzo and Finan, 2017):

- Les chromides présentent toujours des traces de leurs « ancêtres plasmidiques », notamment leur mécanisme de réplication et de partition ressemblent fortement à ceux des mégaplasmes.
- Les gènes de ménage sont fortement sous-représentés sur les chromides suggérant un événement d'acquisition du chromide par transfert horizontal.
- Il y a un biais constant dans l'annotation fonctionnelle des gènes présents sur les chromides par rapport aux gènes présents sur les chromosomes suggérant une histoire évolutive différentes pour les deux réplicons.
- Certaines analyses phylogénétiques démontrent l'évolution des chromides à partir de mégaplasmes (notamment à partir de la protéine de partition RepA dans les réplicons chez les Alphaproteobacteries).
- On observe également une perte des systèmes de conjugaison dans les chromides, attestant de leur transfert vertical stable (Smillie et al., 2010).

II. La réplication du chromosome bactérien

En 1963, Jacob *et al.* sont les premiers à introduire le concept de réplicon, ce modèle suppose que pour initier la réplication d'un chromosome circulaire, une molécule initiatrice (protéine ou ARN) à action positive interagit avec un site spécifique, l'origine de réplication (Figure 3). Ce modèle a probablement été conçu à partir de la découverte alors récente que l'arrêt de la synthèse des protéines et/ou de l'ARN entraîne l'inhibition de l'initiation mais permet la fin des cycles de réplication en cours (Maaloe and Hanawalt, 1961).

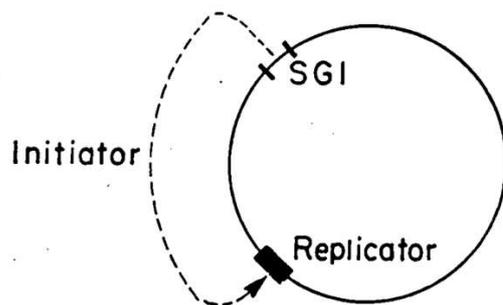


Figure 3. Modèle du réplicon proposé par François Jacob. Le génome bactérien est constitué d'un chromosome circulaire, il encode une molécule initiatrice qui interagit avec une origine de réplication.

Dans ce modèle du réplicon Jacob *et al.* décrivent la temporalité des événements de la réplication. Dans un premier temps, l'initiateur doit être produit pour interagir avec l'origine afin d'initier la réplication. Puis lors de l'initiation, de nouvelles origines sont synthétisées et l'initiateur doit être détruit ou du moins empêché de se reformer aux origines nouvellement synthétisées. Enfin, de nouvelles structures initiatrices doivent ensuite être réformées juste avant que l'initiation puisse avoir lieu dans le cycle cellulaire suivant. Au moment où a été présenté ce modèle, la base moléculaire de la structure initiatrice était complètement inconnue.

Soixante années sont passées depuis l'élaboration du modèle du réplicon, l'étude intensive de la réplication particulièrement chez l'organisme modèle *E.coli* a permis de répondre à de nombreuses interrogations sur son fonctionnement ainsi que sur ses mécanismes de régulation. Nous savons désormais que la réplication du chromosome bactérien se fait via des enzymes appelées ADN polymérases (principalement PolIII), en commençant à partir d'un unique locus situé sur le chromosome (*oriC*) et son démarrage est conditionné par une protéine initiatrice (DnaA). L'initiation est déclenchée à partir d'une seule origine de réplication. La réplication se déroule ensuite de manière bidirectionnelle le long des deux bras du chromosome jusqu'à ce que l'ensemble de la molécule ait été répliquée et que le processus s'arrête.

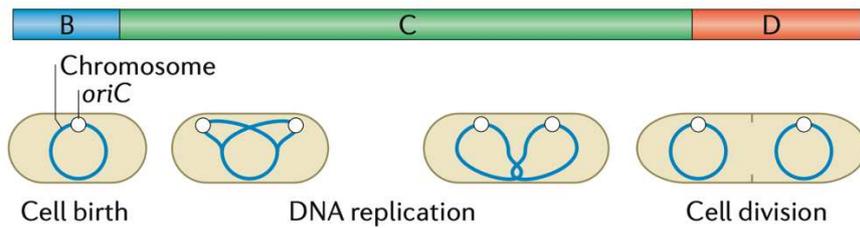
1. Cycle cellulaire et taux de croissance

Plusieurs modèles ont été conçus pour expliquer le couplage entre le cycle cellulaire et la croissance de *E.coli*. En effet, cet organisme peut se développer à différents taux selon le milieu de croissance. Le temps de génération dans un milieu riche, tel que le LB, est d'environ 20 min à 37°C alors que les temps de génération dans des milieux minimaux peuvent être de plusieurs heures (Schaechter et al., 1958).

La caractérisation fine du cycle cellulaire bactérien a été rendue possible grâce à l'invention de la « machine à cellules filles » ou *Baby-machine* (Helmstetter and Cummings, 1963). Dans ce dispositif, des cellules d'*E.coli* sont fixées à la surface d'une membrane en nitrocellulose. Un milieu de croissance en flux constant traverse cette membrane, les cellules fixées se développent et se divisent. Chaque cellule qui se divise donne naissance à deux cellules filles : une encore fixée à la membrane et une non fixée. Les cellules filles non fixées sont alors évacuées du système et récupérées dans un milieu froid où la croissance est stoppée. Les populations de cellules filles ainsi récupérées sont très synchrones, et leur analyse après reprise de la croissance a permis d'améliorer grandement la compréhension du cycle cellulaire bactérien (Bogan and Helmstetter, 1996).

En 1968, Helmstetter et Cooper propose leur modèle d'organisation du cycle cellulaire bactérien (Figure 4). Les cellules d'*E.coli* avec un temps de doublement compris entre 20 et 60 minutes répliquent complètement leur ADN en 40min (période C). Après la fin de la terminaison de la réplication de l'ADN, la cellule a besoin de 20 minutes supplémentaires (période D) pour se diviser en deux cellules filles. Ainsi dans cet organisme modèle le temps de génération peut être plus court (20min) que le temps nécessaire à la réplication de son chromosome (40min).

Non-overlapping cell cycles (slow growth)



Overlapping cell cycles (fast growth)

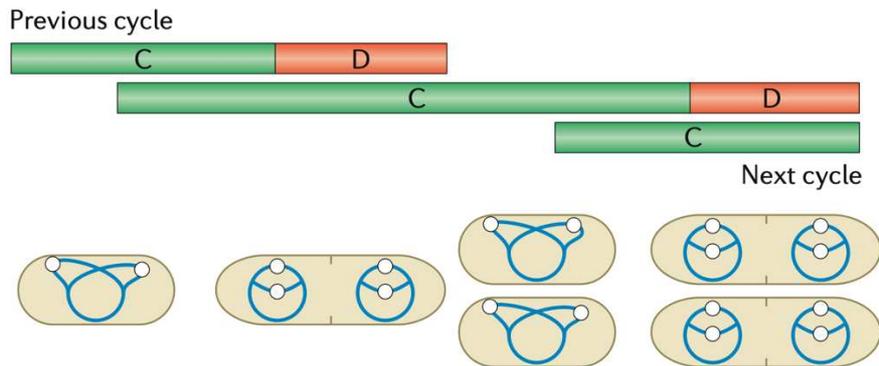


Figure 4. Les étapes du cycle cellulaire en relation avec le taux de croissance. Au cours d'une croissance lente, le temps de génération est plus long que les périodes C et D combinées, ce qui entraîne une réplication complète de l'initiation à la terminaison dans un seul cycle cellulaire. Au cours d'une croissance rapide, le temps de génération est plus court que les périodes C et D combinées, ce qui entraîne des événements multiples d'initiation de la réplication de l'ADN et des cycles de réplication qui se chevauchent dans chaque cellule ; dans ce cas, de multiples origines de réplication initient la réplication de manière synchrone. Les cellules dont les cycles de réplication se chevauchent n'ont pas de période B (Reyes-Lamothe and Sherratt, 2019).

Or, alors même que le temps de génération est plus court que le temps de réplication, le temps combiné des périodes C et D reste constant et indépendant du taux de croissance (Helmstetter, 1968). La solution trouvée par l'évolution pour parvenir à des temps de génération aussi court est la mise en place de cycles de réplication chevauchants. On estime que lorsque les cellules d'*E.coli* se développent avec un temps de génération de 20min (en milieu riche), elles auront par conséquent une réplication de l'ADN à fourches multiples et jusqu'à trois cycles cellulaires qui se chevauchent. Après terminaison d'un cycle de réplication, la cellule mère se divise et les cellules filles héritent de chromosomes partiellement répliqués à la naissance.

Par la suite, il a été démontré que toutes les bactéries à croissance rapide, comme *Escherichia coli*, *Bacillus subtilis* ou encore *Vibrio cholerae*, ont évolué pour coordonner des cycles de réplication chevauchant avec la croissance cellulaire, la ségrégation des chromosomes et la division cellulaire. Il y a notamment un biais dans l'organisation des gènes lié aux machineries de réplication, de transcription et de traduction qui sont localisés proches

de l'origine profitant ainsi d'une augmentation de *gene dosage* lors de la réplication multi-fourches (Couturier and Rocha, 2006).

En revanche, les espèces bactériennes à croissance plus lente (par exemple *Caulobacter crescentus* ou *Rhodobacter sphaeroides*), ainsi que les bactéries à croissance rapide poussant dans des conditions limitées en nutriments, terminent la réplication dans la même génération qu'elle a été initiée (Dubarry et al., 2019), (Collier, 2012).

2. La masse d'initiation ou volume critique

Toujours dans les années soixante, d'autres théories sur le cycle cellulaire bactérien ont été proposées. Il a été observé que les cellules à croissance rapide sont d'une taille plus importante que les cellules à croissance lente. En effet, il s'avère que le volume des cellules bactériennes est proportionnel au taux de croissance. Celui-ci augmente exponentiellement au fur et à mesure de la croissance (Schaechter et al., 1959).

D'autre part, il a été remarqué que l'initiation de la réplication de l'ADN dépend d'un unique paramètre : le volume cellulaire par origine (Donachie, 1968). Des mesures très minutieuses de l'ADN et de la masse cellulaire à différents taux de croissance de *Salmonella typhimurium* (Schaechter et al., 1958) ont conduit à la conclusion que l'initiation de la réplication se fait lorsqu'une certaine masse est atteinte par *oriC* : on nomme ce seuil masse d'initiation ou volume critique. Par la suite, il a été prouvé chez *E.coli* que la masse d'initiation par origine est toujours la même à différents taux de croissance. Lorsque cette masse par origine est atteinte, toutes les origines de réplication présente dans la cellule sont initiées simultanément (Skarstad and Lobner-Olesen, 2003).

Ainsi l'initiation de la réplication apparaît comme un point de réinitialisation dans le cycle cellulaire bactérien, c'est l'évènement central qui façonne tout le déroulement du cycle.

3. Découverte de DnaA l'initiateur universel bactérien et de l'origine de réplication *oriC*

En 1963, dans le laboratoire de François Jacob, des mutants de *E.coli* dont la réplication de l'ADN est thermosensible ont été isolés après une mutagénèse aléatoire. Certains de ces mutants étaient capables de terminer leur réplication à une température de croissance non permissive mais ils étaient incapables de re-initier leur initiation. Une fois

cartographiées génétiquement, ces mutations se sont révélées être situées à proximité du locus *ilv* dans le gène *dnaA* (Hirota et al., 1968). Il sera ensuite démontré que DnaA est la protéine initiateur universelle présente dans tous les chromosomes bactériens (Katayama et al., 2010).

L'origine de la réplication a été identifiée également à proximité des gènes *ilv* via plusieurs études au cours des années 1970 (Bird et al., 1972). Le symbole *oriC* a finalement été introduit pour : origine de réplication du chromosome de *E. coli* (Hiraga, 1976). Une analyse par délétion successive a permis de délimiter la région *oriC* à une taille de 245 pb (Tabata et al., 1983). Bien que tous les chromosomes bactériens possèdent une région *oriC*, elles sont étonnamment diverses en termes de taille (allant de 250 pb à 2 kbp), de séquence et d'organisation. On peut citer par exemple le cas d'*Helicobacter pylori* qui possède une origine bipartite (*oriC1*, *oriC2*) qui entoure de part et d'autre le gène *dnaA* (Donczew et al., 2012).

Néanmoins, la capacité du système DnaA-*oriC* à déclencher la réplication suit toujours le même processus. La fixation et l'oligomérisation de DnaA à *oriC* entraînent l'ouverture des deux brins d'ADN d'une région riche en AT et le recrutement de deux complexes de réplication nommés réplisomes. La réplication du chromosome est bidirectionnelle, les deux réplisomes progressent et répliquent l'ADN le long des deux bras du chromosome nommés réplichores (Figure 5).

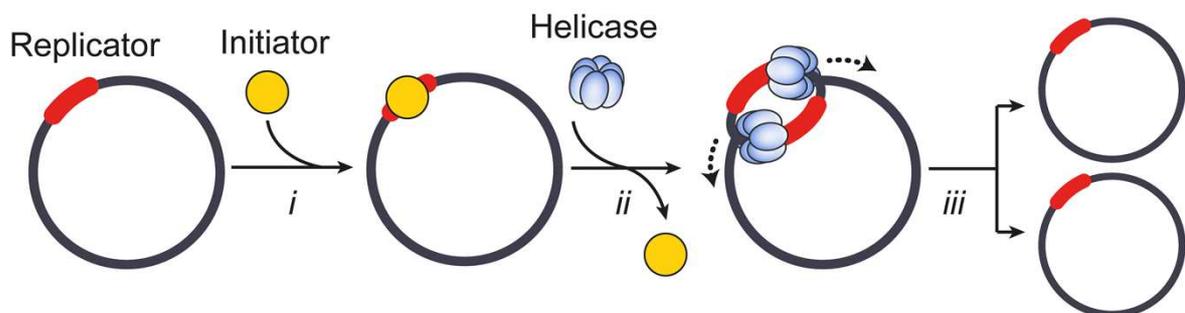


Figure 5. Modèle général du principe de l'initiation de la réplication bactérienne. Les chromosomes bactériens circulaires contiennent un élément agissant en cis, le réplicateur, qui est situé au niveau ou à proximité des origines de réplication. i) Le réplicateur recrute les protéines initiateur d'une manière spécifique à la séquence d'ADN, ce qui entraîne la fusion de l'hélice d'ADN et le chargement de l'hélicase répliquative sur chacun des brins d'ADN simples (ii). iii) Les réplicateurs assemblés répliquent l'ADN de manière bidirectionnelle pour produire deux copies du chromosome bactérien (Ekundayo and Bleichert, 2019).

4. DnaA : une protéine de la famille AAA+ ATPase

DnaA est la protéine clé dans l'initiation de la réplication de l'ADN chromosomique conservée et essentielle chez toutes les bactéries. Elle déclenche le processus de réplication en se liant sur l'origine à des séquences spécifiques de l'ADN appelées boîtes DnaA. Lorsque DnaA se lie à ces séquences, elle déclenche l'ouverture de la double hélice et la séparation des deux brins d'ADN. Ce processus permet la formation des fourches de réplication où de nouveaux brins d'ADN pourront être synthétisés.

A ce jour aucune structure complète de la protéine DnaA n'a été obtenue, la protéine semble néanmoins être organisée en quatre domaines structuraux aux fonctions distinctes (Figure 6).

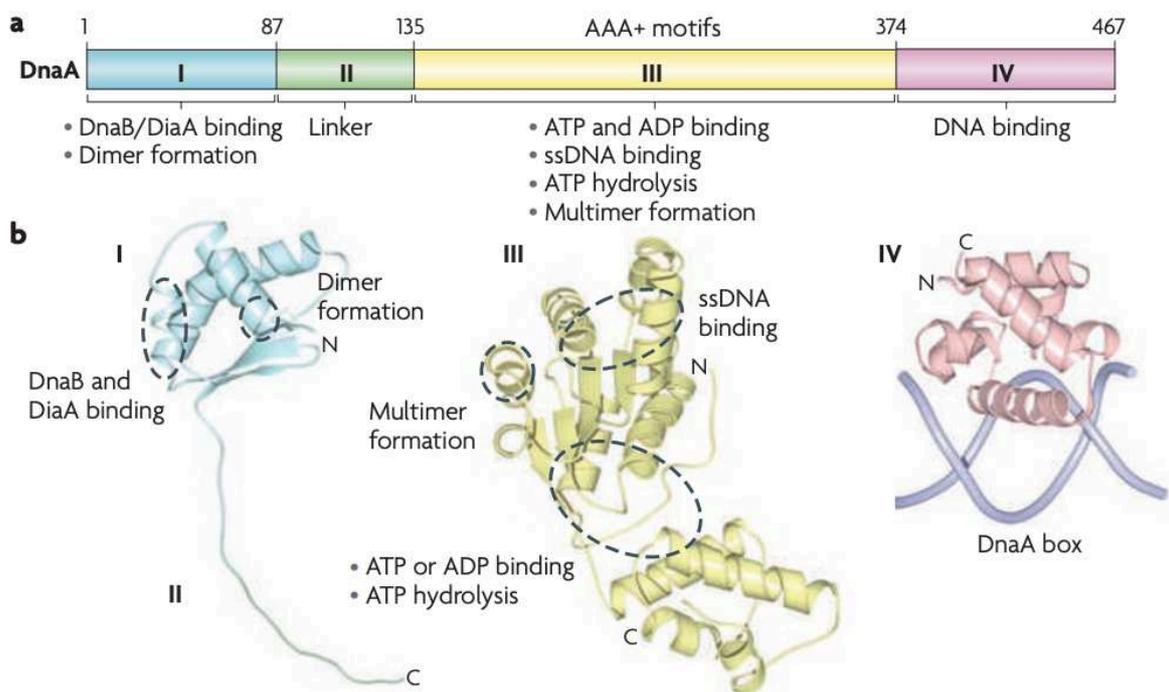


Figure 6. L'initiateur DnaA est organisé en quatre domaines structuraux. Structures tertiaires des domaines I à IV de la DnaA de *E. coli*. Plusieurs régions fonctionnelles dans chaque domaine sont indiquées par des contours en pointillés (Katayama *et al.*, 2010).

Le domaine I permet les interactions protéine-protéine. Il permet la dimérisation de DnaA qui est essentielle à l'initiation de la réplication. Il permet également l'interaction avec d'autres protéines : avec l'hélicase répllicative DnaB (Seitz *et al.*, 2000) et avec la protéine auxiliaire DiaA chez *E. coli* (Ishida *et al.*, 2004).

Le domaine II est considéré comme étant un linker flexible sans fonction régulatrice particulière permettant la liaison entre les domaines I et III.

Le domaine III contient un motif AAA+ : ATPases associées à diverses activités. Via ce motif, DnaA peut se lier soit à l'ATP et soit à l'ADP avec des affinités similaires, à noter que l'ATP est majoritaire dans la cellule (Sekimizu et al., 1987). La forme DnaA-ATP est la forme active pour l'initiation de la réplication pour plusieurs raisons. D'abord elle facilite la liaison coopérative de DnaA aux boîtes DnaA dégénérées de l'origine (McGarry et al., 2004). Elle est nécessaire pour l'ouverture de l'ADN au niveau de la région riche en AT ou DUE (Duplex Unwinding Element) à l'*oriC* (Sekimizu *et al.*, 1987), ainsi que pour la liaison à l'ADN simple brin (ss) après ouverture (Ozaki and Katayama, 2012).

Le domaine IV quant à lui permet l'interaction avec l'ADN double brin sur les boîtes DnaA. Cette liaison à l'ADN est médiée par un motif hélice-tour-hélice (HTH) (Fujikawa et al., 2003).

5. Ouverture de l'ADN au niveau de l'*oriC* chez *E. coli*

L'origine *oriC* est le site spécifique sur le chromosome bactérien où la réplication de l'ADN commence. Elle contient trois éléments fonctionnels qui permettent l'initiation: des répétitions d'ADN conservées qui sont spécifiquement reconnues par DnaA (appelées boîtes DnaA), une région d'ADN riche en nucléotides A/T (aussi appelée DUE pour DNA Unwinding Element) et des sites de fixation pour des protéines secondaires qui aident à réguler l'initiation de la réplication (IHF, Fis).

L'initiateur DnaA interagit à la fois avec l'ADN double brin (db) au niveau des boîtes DnaA et l'ADN simple brin (sb) sur des séquences de 13bp dans la région riche en AT. L'*oriC* contient quatre types de sites de fixation double brin pour DnaA. Ceux-ci diffèrent dans leurs affinités pour DnaA et leurs dépendances au cofacteur ATP. Les boîtes DnaA R1, R2 et R4 sont des sites de haute affinité qui sont liés par le domaine HTH de DnaA indépendamment de l'état de liaison de l'ATP (Weigel et al., 1997). En revanche, les sites I, τ et C, qui sont intercalés entre les sites R, sont des boîtes DnaA de faible affinité et s'associent préférentiellement à la DnaA liée à l'ATP (McGarry *et al.*, 2004).

La liaison de plusieurs protéines DnaA via leur domaine HTH aux boîtes de haute et de basse affinité favorise l'oligomérisation d'autres protéines DnaA de manière ATP dépendante. L'ensemble de ces protéines DnaA forment alors un filament qui enroule le duplex

d'ADN autour de sa surface externe. Cet enroulement génère une torsion super-hélicoïdale entraînant la dissociation des deux brins d'ADN au niveau de la région riche en AT (Zorman et al., 2012). Cette ouverture est favorisée car il y a deux interactions hydrogènes dans la paire A/T contre 3 pour la paire G/C, ainsi l'énergie libre (ΔG) nécessaire pour séparer les deux brins d'ADN est beaucoup plus faible dans cette région que dans les séquences adjacentes.

La séparation des brins d'ADN est également facilitée par les interactions directes du domaine AAA+ avec des sites contenus dans la région riche en AT, appelées DnaA-trios. L'interaction du filament de DnaA avec les DnaA-trios étire l'ADN simple brin et stabilise la bulle d'initiation en empêchant la reformation du duplex d'ADN (Richardson et al., 2016).

La structure exacte du complexe d'ouverture après oligomérisation de DnaA reste incertaine. Deux modèles ont été proposés pour expliquer l'organisation du filament de DnaA le long de l'*oriC*. Le premier nommé modèle à deux états propose qu'un filament continu de DnaA se forme avec d'abord un mode de liaison à l'ADN_{db} (le complexe organisateur) puis à un mode de liaison à l'ADN_{sb} dans la DUE (le complexe de séparation) (Figure 7) (Duderstadt et al., 2010).

Le second nommé modèle à boucle propose que l'ADN de l'*oriC* soit plié en deux via l'action d'IHF. Ainsi, la région DUE se replie sur le filament initiateur de sorte que les protomères de DnaA se lient simultanément aux régions d'ADN double et simple brin (Figure 7) (Ozaki and Katayama, 2012).

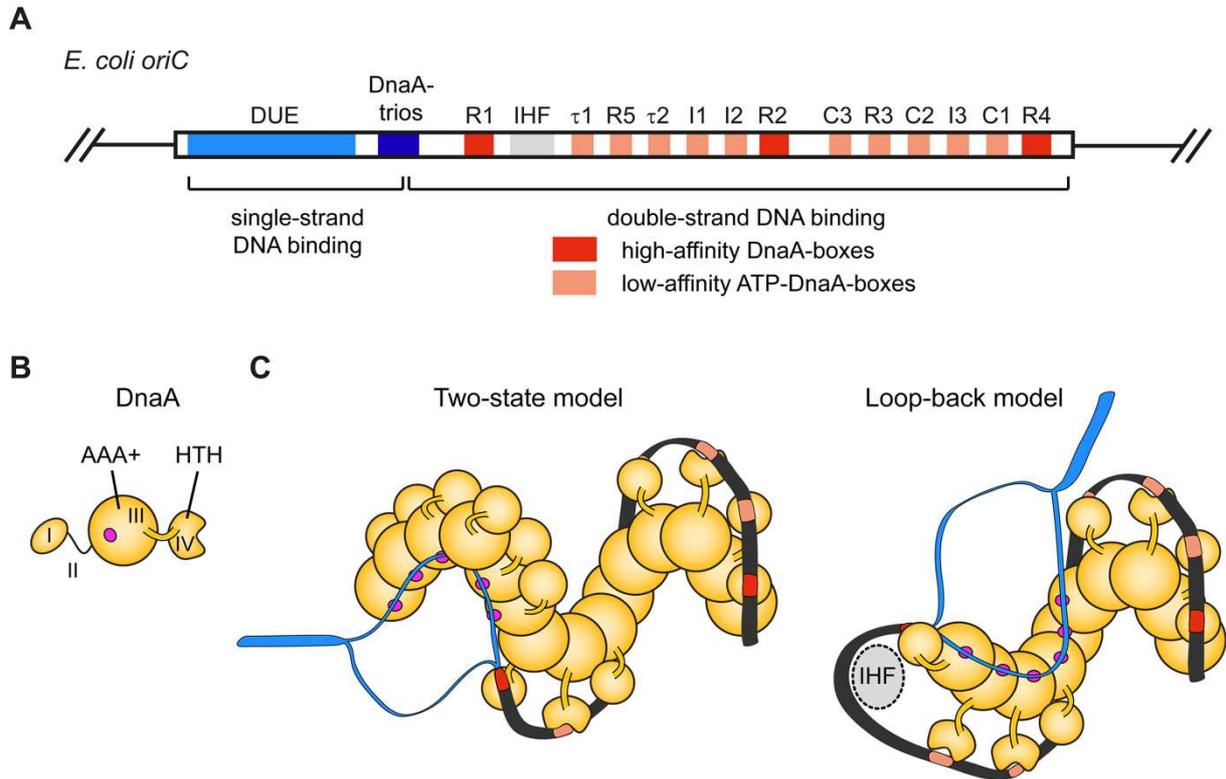


Figure 7. Organisation et initiation de l'origine chez les bactéries. A) Schéma de l'architecture de l'origine *oriC* d'*E. coli*. La région DUE est flanquée d'un côté par plusieurs boîtes DnaA à haute et faible affinité B) Organisation des domaines de l'initiateur DnaA d'*E. coli*. Le cercle magenta indique le site de liaison à l'ADN simple brin. C) Modèles pour la reconnaissance de l'origine et la fusion par DnaA. Dans le modèle à deux états (panneau de gauche), les protomères de DnaA passent d'un mode de liaison à l'ADNdb (médié par les domaines HTH reconnaissant les boîtes de DnaA) à un mode de liaison à l'ADNsb (médié par les domaines AAA+). Dans le modèle en boucle, l'ADN est fortement courbé vers l'arrière sur le filament de DnaA (facilité par la protéine régulatrice IHF) de sorte qu'un seul protomère se lie à la fois aux régions duplex et simple brin. Dans les deux cas, le filament DnaA fait séparer le duplex d'ADN et stabilise la bulle d'initiation avant le chargement de l'hélicase répliquative (Ekundayo and Bleichert, 2019).

6. Inactivation d'*oriC* post-initiation chez *E.coli* : séquestration par SeqA

Après l'initiation de la réplication, l'origine doit être inactivée pendant une certaine période pour éviter un phénomène de sur-initiation, chez *E.coli* cette inactivation est en parti assurée par un mécanisme de séquestration. Il y a sur l'*oriC* une sur-représentation de motifs GATC et notamment dans les boîtes DnaA. Ces sites GATC sont reconnues et méthylés par la Dam méthylase au niveau de l'adénine sur les deux brins d'ADN. Après le passage de la fourche de réplication sur un site GATC, celui-ci passe à un état hémiméthylé c'est-à-dire qu'un seul des deux brins est méthylé (Figure 8).

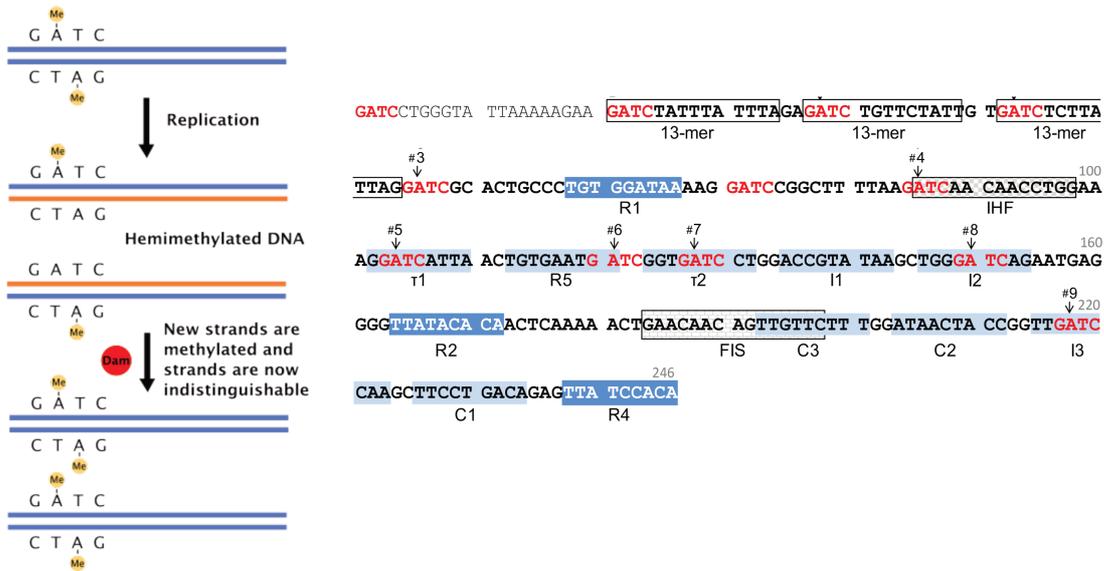


Figure 8. Enrichissement d'*oriC* en motif GATC. A gauche : principe de l'hémiméthylation des motifs GATC par la Dam méthylase, après le passage de la fourche de réplication les motifs GATC sont transitoirement hémiméthylés, uniquement sur le brin matrice. A droite : séquence d'*oriC* qui inclut la région minimale requise pour la fonction d'origine. La région comprend plusieurs sites GATC (indiqués en rouge) (Jha and Chattoraj, 2016).

Les sites GATC hémiméthylés sont spécifiquement reconnus par la protéine SeqA. Post-initiation, SeqA se lie à l'*oriC* hémiméthylée empêchant la ré-initiation aux origines nouvellement répliquées (Figure 9). Ce phénomène de séquestration est maintenu pendant environ un tiers du cycle cellulaire et se termine par la dissociation spontanée de SeqA de l'*oriC* puis la méthylation complète par Dam (Campbell and Kleckner, 1990). Le rôle de SeqA est double car elle permet également la séquestration du promoteur de *dnaA* inhibant ainsi sa transcription. Dans les cellules dépourvues de SeqA, les ré-initiations sont fréquentes et conduisent à une sur-initiation (Campbell and Kleckner, 1990).

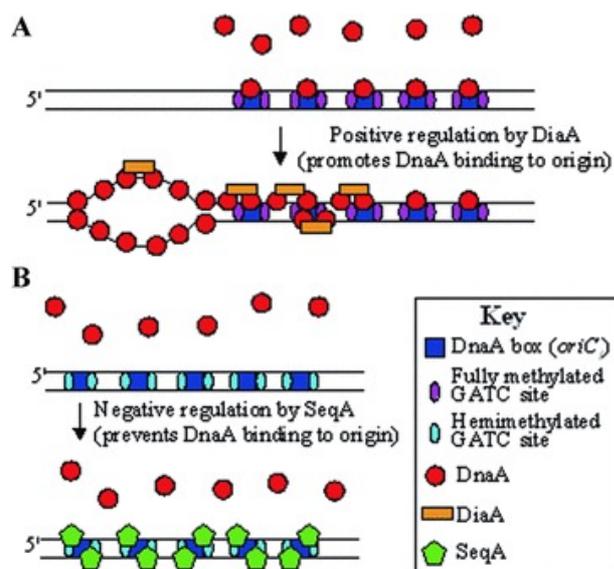


Figure 9. Séquestration de l'*oriC* par SeqA. En se fixant à *oriC* sur les sites GATC hémiméthylés SeqA empêche les protéines DnaA de se fixer et ainsi inhibe la re-initiation de la réplication de l'ADN.

On estime que *E. coli* est capable d'initier environ 10 copies d'*oriC* (portées par des plasmides) par *oriC* chromosomique, cela indique qu'il y a à la fois une grande capacité d'initiation et qu'il n'y a pas de mécanisme de "comptage" pour s'assurer que le bon nombre d'origines a été initié. Au contraire, toutes les origines présentes dans la cellule sont initiées simultanément puis séquestrées (Skarstad and Lobner-Olesen, 2003). Le niveau intracellulaire de DnaA-ATP est le déterminant principal de la fenêtre d'initiation de la réplication et ce niveau de DnaA-ATP est corrélé au volume de la cellule. Chez *E. coli* trois systèmes majeurs de régulation influencent le niveau de DnaA-ATP (Katayama et al., 2017), (Skarstad and Katayama, 2013), (Riber et al., 2016).

7. Inactivation de DnaA : RIDA et DDAH

Une fois que la réplication a été initiée, différents mécanismes vont venir empêcher une ré-initiation précoce et ainsi limiter le nombre de cycles de réplication chevauchant. Parmi ces processus, certains permettent l'inactivation directe de l'initiateur en régulant la forme active de DnaA, celle liée à l'ATP.

Le premier système est nommé RIDA pour Regulatory Inactivation of DnaA. C'est un mécanisme couplé à la réplication et qui permet l'hydrolyse de l'ATP fixé à DnaA. Dans celui-ci la protéine AAA+ Hda lié à l'ADP interagit avec la sous-unité β -clamp (DnaN) de l'holoenzyme PolIII fixé à l'ADN lors de la réplication. Le complexe Hda-clamp-ADN interagit spécifiquement avec DnaA-ATP et catalyse la formation de DnaA-ADP (Katayama *et al.*,

2017). RIDA est considéré comme le principal inactivateur de DnaA après l'initiation de la réplication, et réprime fortement la sur-initiation de la réplication (Figure 10).

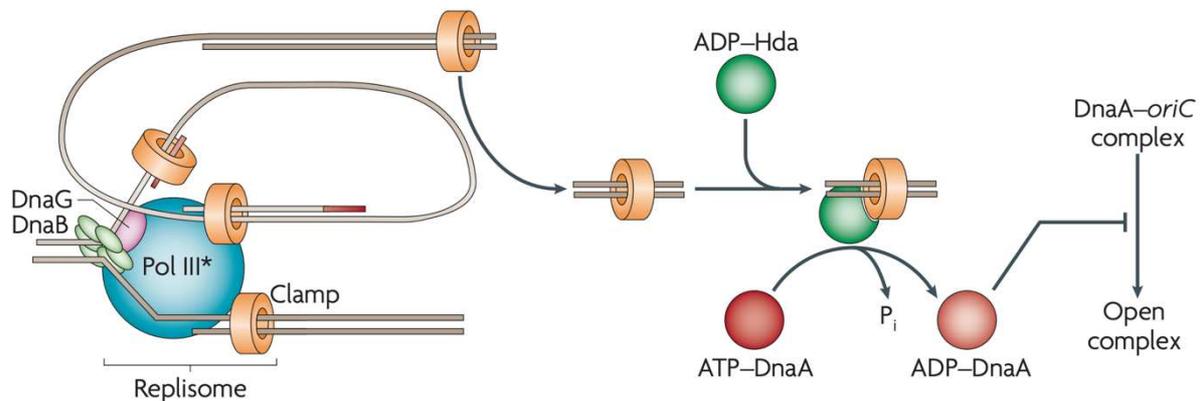


Figure 10. Inactivation de DnaA par RIDA chez *E. coli*. Après la formation du réplisome et l'achèvement d'un fragment d'Okazaki, le β -clamp reste sur l'ADN et se lie à la protéine Hda. Le complexe ADP-Hda-clamp-ADN interagit avec DnaA-ATP et favorise l'hydrolyse de l'ATP (Katayama *et al.*, 2010).

Le second système d'inactivation directe de DnaA implique la liaison de l'initiateur à une région chromosomique spécifique nommée *data*. En effet, la protéine DnaA se fixe en dehors d'*oriC* sur environ 300 boîtes DnaA (9pb) le long du chromosome de *E.coli* (Smith and Grossman, 2015). Certains de ces sites n'ont pas de rôle particulier, d'autres régulent la transcription de gènes en aval, et enfin certains régulent directement l'activité de DnaA (Olliver *et al.*, 2010). Ce système de régulation de DnaA est nommé DDAH pour *data*-dependent *DnaA*-*ATP* *hydrolysis*. Tout comme RIDA, il permet de réguler négativement DnaA en favorisant l'hydrolyse de sa molécule d'ATP. Il requiert la fixation de plusieurs DnaA à une région d'ADN proche d'*oriC* nommée *data*. La région *data* est un locus de 262 pb contenant plusieurs boîtes DnaA ainsi qu'un site de liaison à IHF. Le complexe *data*-IHF se forme après l'initiation de la réplication, et stimule le recrutement de DnaA-ATP puis l'hydrolyse de l'ATP lié à DnaA. Ce système DDAH joue un rôle d'inactivation complémentaire à RIDA (Kasho and Katayama, 2013).

Ensemble les deux systèmes RIDA et DDAH permettent d'inhiber la sur-initiation de l'*oriC* en convertissant rapidement le lot de DnaA-ATP libre en DnaA-ADP. Lorsque la réplication a progressé et que le volume cellulaire augmente, le lot de DnaA-ADP est progressivement reconvertie en DnaA-ATP, pour permettre une nouvelle initiation.

8. Régénération de DnaA-ATP

L'initiation de la réplication se produit uniquement lorsque le niveau cellulaire DnaA-ATP atteint un pic suffisamment élevé pour permettre l'assemblage sur *oriC*. Après, l'initiation le lot cellulaire de DnaA-ATP est inactivé après l'initiation par les mécanismes RIDA et DDAH en le convertissant en DnaA-ADP. Lorsque la réplication a suffisamment progressé le lot cellulaire de DnaA-ATP est progressivement reformé via différents mécanismes (Figure 11).

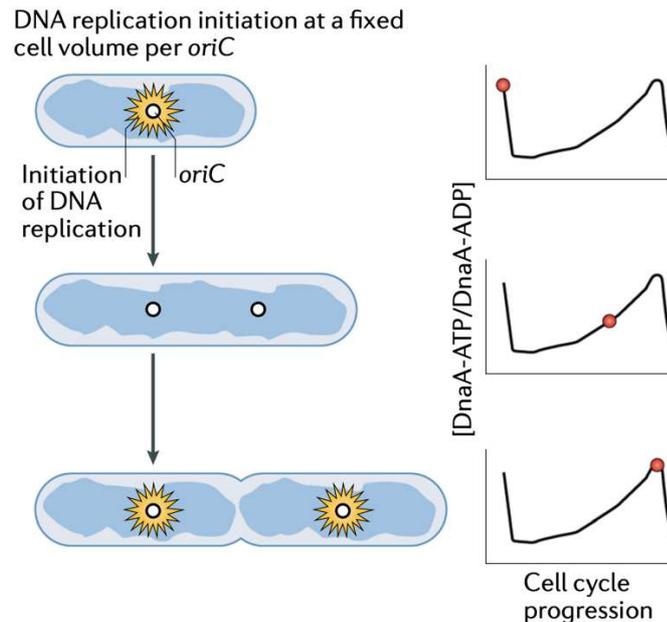


Figure 11. L'oscillation du niveau cellulaire de DnaA-ATP contrôle l'initiation d'*oriC*. L'initiation de la réplication de l'ADN se produit à un volume cellulaire fixe par origine (*oriC*), avec de multiples origines de réplication initiant la réplication de manière synchrone. Le rapport DnaA-ATP/DnaA-ADP fluctue au cours du cycle cellulaire. Les points rouges dans les graphiques représentent l'état de la DnaA dans le schéma cellulaire illustré à gauche : près de l'initiation, la concentration de DnaA-ATP est élevée et la concentration chute peu après l'initiation et augmente avec le temps avant l'événement d'initiation (Reyes-Lamothe and Sherratt, 2019).

DnaA-ATP est produit uniquement par la liaison passive des protéines DnaA libres à l'ATP qui est très abondant dans le cytosol. Trois mécanismes de production de protéines DnaA libres ont été identifiés : la traduction *de novo* de DnaA ; la dissociation de ADP-DnaA stimulée par l'interaction avec les phospholipides de la membrane ; et la fixation à des sites chromosomiques nommés DARS pour DnaA-Reactivating Sequences. Ce dernier est le principal mécanisme de réactivation.

Les séquences DARS ont d'abord été identifiées sur le plasmide ColE1 où l'on retrouve une séquence de 70pb contenant trois boîtes DnaA qui permet de convertir les DnaA-ADP en DnaA-ATP (Fujimitsu and Katayama, 2004). Puis deux sites DARS ont été identifiés sur le chromosome d'*E.coli* (Fujimitsu et al., 2009). DARS1 est situé dans la région promotrice de

uvrB et DARS2 dans le promoteur de *mutH*. DARS1 et DARS2 favorisent spécifiquement la dissociation de la molécule d'ADP de la protéine DnaA-ADP, ce qui stimule l'échange de nucléotides dans le cytoplasme et augmente les niveaux DnaA-ATP (Kasho et al., 2014).

In vitro, un ADN contenant DARS1 peut promouvoir la dissociation ADP de DnaA sans facteurs supplémentaires. A l'inverse, le site DARS2 a peu d'activité seul, il nécessite la fixation des protéines IHF et Fis, qui se lie à leur site respectif et plie l'ADN d'au moins 60°. Dans la cellule, DARS1 et DARS2 ont tous deux un rôle dans la production d'ATP-DnaA et favorisent l'initiation de la réplication. La délétion de l'un ou l'autre ou des deux retardent l'initiation (Kasho et al., 2014).

DARS1 et DARS2 sont de tailles différentes : 101 et 455 pb respectivement. Les deux sites contiennent une région centrale contenant trois boîtes DnaA, l'espacement entre ces boîtes est le même pour DARS1 et DARS2, mais pas leur orientation. En plus des régions centrales, DARS1 et DARS2 possèdent des régions régulatrices de longueurs différentes. La région régulatrice de ~50 pb de DARS1 renforce l'activité de dissociation DnaA-ADP (Fujimitsu et al., 2009). La région régulatrice ~400 pb de DARS2 contient les sites de fixation de IHF et de Fis, qui stimulent à la fois le recrutement de DnaA-ADP et la dissociation de l'ADP de DnaA (Kasho et al., 2014). La fixation de IHF et de FIS à DARS2 oscille pendant le cycle cellulaire permettant la production de DnaA-ATP au moment opportun (Miyoshi et al., 2021).

Cependant, la délétion de DARS2 provoque des initiations plus asynchrones que la délétion de DARS1. De manière similaire, la présence de DARS2 sur un plasmide multicopie entraîne une forte stimulation d'initiations supplémentaires comparée à celle de DARS1 (Kasho et al., 2014). Dans le modèle de régénération de DnaA-ATP, DARS1 permet d'en produire un niveau basal alors que DARS2 permet réellement de déterminer le moment de l'initiation (Figure 12).

Lorsqu'ils sont répliqués, ces sites doublent en nombre de copies et le rythme de formation de DnaA-ATP s'accélère. Ainsi leur positionnement sur le génome est important, et lorsque ces sites sont déplacés à d'autres loci cela perturbe la régulation de l'initiation de l'*oriC* (Frimodt-Moller et al., 2016).

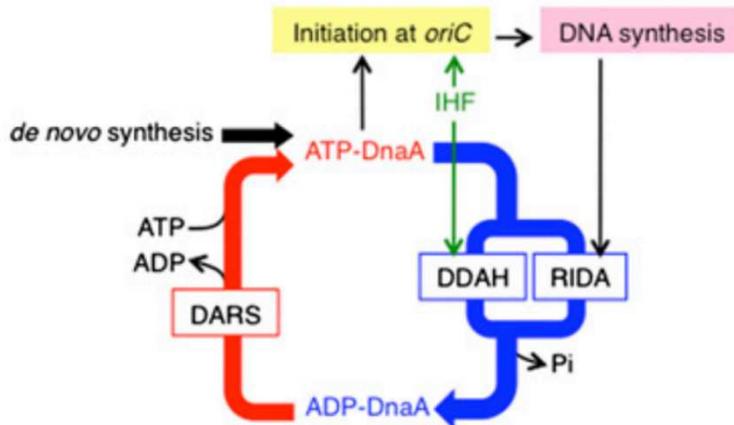


Figure 12. Cycle d'activité de DnaA. L'ATP-DnaA est hydrolysée par deux voies indépendantes, RIDA et DDAH. Les sites DARS réactivent l'ADP-DnaA par un échange ADP-ATP (Kasho and Katayama, 2013).

9. Chargement du réplisome lors de l'initiation de la réplication à *oriC*

La polymérisation de DnaA-ATP le long de l'*oriC*, et son interaction avec la DUE permettent l'ouverture des deux brins d'ADN. Après la formation de cette « bulle » d'ADN, deux réplisomes vont alors être recrutés pour commencer la réplication. Chez *E. coli*, le réplisome est constitué de douze protéines distinctes qui s'assemblent sur l'ADN et combinent leurs actions pour répliquer fidèlement le chromosome. L'hélicase répliquative DnaB joue un rôle clé au sein du réplisome. C'est une ATPase homo-hexamérique qui entoure l'ADN simple brin (ss). Une fois chargée sur l'ADNss, DnaB hydrolyse l'ATP pour permettre sa translocation dans la direction 5'-3'. DnaB sépare l'ADN double brin (ds) parental et fournit ainsi les matrices d'ADN simple brin pour l'ADN polymérase (Lewis et al., 2016).

Pendant l'initiation de la réplication, le chargement de DnaB à l'origine de réplication *oriC* est considéré comme l'étape clé dans l'assemblage du réplisome. Afin d'être chargée, DnaB doit former un complexe avec la protéine de chargement DnaC, puis être recrutée à *oriC* par la protéine initiateur DnaA. Le rôle de DnaC est de maintenir ouvert l'hexamère de DnaB dans le but de le placer sur l'ADN simple brin via l'interaction directe de DnaA (Arias-Palomo et al., 2013). Chez *B.subtilis*, ce processus est assuré par une protéine homologue nommée DnaI (Smits et al., 2010). Cependant dans la plupart des phylum bactériens dont *Vibrio cholerae*, il n'existe pas d'homologue de DnaC ni de DnaI. Au lieu de cela, c'est le chargeur bactérien ancestral DciA qui partage des similitudes avec DnaA et DnaX (Brezellec et al., 2016).

Une fois chargé sur l'ADN simple brin, l'hexamère de DnaB se referme et recrute la primase DnaG stimulant la dissociation de DnaC. DnaG est une ARN polymérase ADN-dépendante simple brin, elle permet la synthèse de courts fragments d'ARN complémentaires à l'ADN simple brin nommés amorces. DnaG synthétise l'amorce d'ARN une seule fois pour initier la synthèse de l'ADN du leading strand et à plusieurs reprises sur le lagging strand afin d'initier la synthèse des fragments d'Okazaki. DnaG reconnaît des sites d'amorçage spécifiques contenant préférentiellement le trinucéotide CTG (Mitkova et al., 2003). L'hétéroduplex ARN/ADN est le point de départ pour l'holoenzyme Pol III. L'holoenzyme ADN polymérase III (ou HE Pol III) est l'enzyme principale qui catalyse l'élongation de la chaîne d'ADN pendant la réplication (Figure 13).

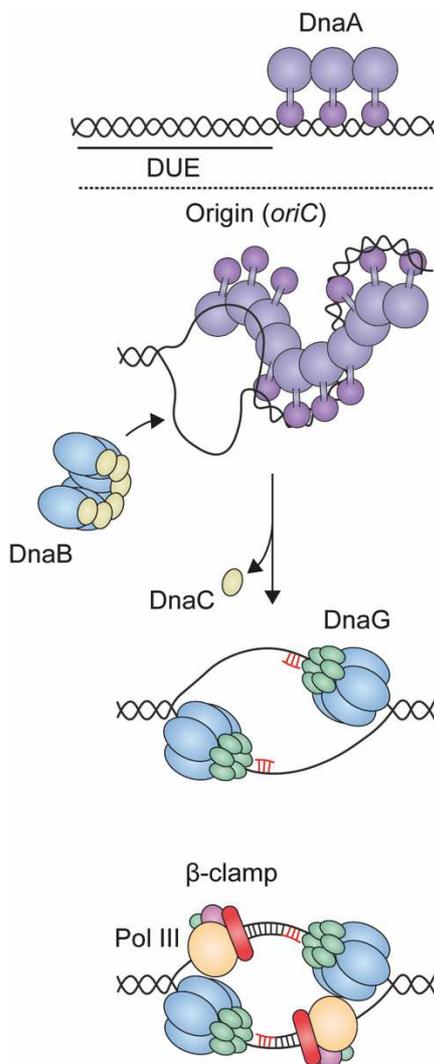


Figure 13. Modèle de recrutement des réplisomes lors de l'ouverture de d'*oriC* par DnaA. DnaA se fixe coopérativement à l'*oriC*, sa polymérisation entraîne l'ouverture du duplex d'ADN au niveau de la région DUE. DnaA recrute ensuite le complexe DnaBC, maintenu ouvert par l'action de DnaC. Une fois l'héxamère de DnaB chargée sur l'ADN simple brin, DnaC se retire du complexe aidé par le recrutement de la primase DnaG. DnaG synthétise de court fragment d'ARN nommé amorce qui permettent le recrutement de l'holoenzyme PolIII et du clamp. La réplication peut alors commencer.

10. Étape d'élongation lors de la réplication de l'ADN

Pendant l'initiation de la réplication au niveau de l'origine, deux machines moléculaires dédiées à la réplication, nommées réplisomes, sont recrutées. Les deux réplisomes progressent ensuite le long du génome en répliquant les deux bras chromosomiques, les réplichores, formant une structure nommée fourche de réplication. Les deux fourches de réplication se rencontrent et fusionnent au niveau de la région terminale du chromosome achevant ainsi la réplication complète du génome.

Lors de la progression d'une fourche de réplication, les deux brins d'ADN sont répliqués simultanément par un seul et même réplisome. La réplication se fait uniquement du sens 5' vers 3', ce qui impose deux modes de réplication différents. Le leading strand est répliqué de manière continue, tandis que la réplication du lagging strand est discontinue et se fait sous forme de courts fragments d'Okazaki.

Chez *E. coli*, la principale polymérase répliquative est l'holoenzyme Pol III. Elle est composée d'une dizaine de protéines différentes organisées en trois complexes distincts : le core $\alpha\epsilon\theta$, la pince coulissante β (ou β -clamp) et le complexe chargeur de pince (ou clamp loader). Cette holoenzyme permet le couplage de la réplication du brin leading et celle du brin lagging (Figure 14).

Pol III a une double activité : ADN polymérase 5'→3' et exonucléase de relecture 3'→5' assurant une grande fidélité dans la réplication. Après la synthèse d'une amorce d'ARN complémentaire par DnaG, le β -clamp est chargée sur l'extrémité de l'amorce par le clamp loader. Le core se fixe au β -clamp le stabilisant ainsi sur l'ADN. Enfin le complexe est capable de synthétiser le brin d'ADN complémentaire à grande vitesse (~1000 Nt/s) et avec une grande fidélité (fréquence d'erreur 10^{-7}) (Xu and Dixon, 2018).

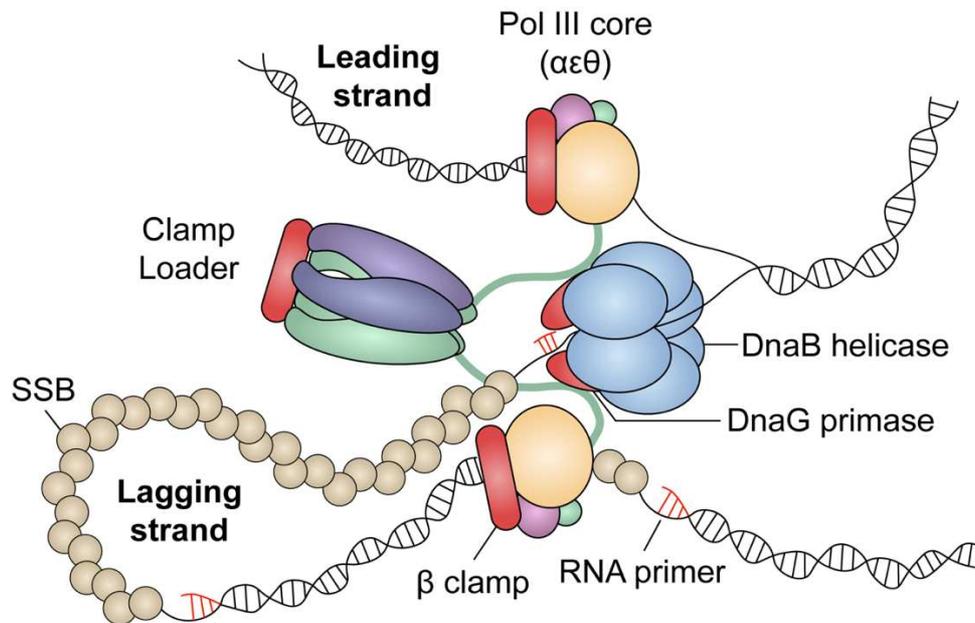


Figure 14. Représentation schématique du réplisome d'*E. coli*. Modèle classique du réplisome d'*E. coli* avec synthèse couplée du brin principal et du brin secondaire. La Pol III* est reliée à DnaB par la sous-unité τ du complexe clamp-loader et deux ou trois noyaux de polymérase de la même Pol III* répliquent à la fois l'ADN du brin principal et du brin secondaire. Le ssDNA dans la boucle du brin retardé est lié par la protéine de liaison ssDNA, SSB (Xu and Dixon, 2018).

Après le passage du réplisome, la ribonucléase H reconnaît spécifiquement les hétéroduplexes ARN/ADN et élimine l'amorce d'ARN générée par la primase. Puis l'ADN polymérase I complète les désoxyribonucléotides manquants entre les fragments d'Okazaki dans le sens 5' → 3'. Enfin, l'ADN ligase permet ensuite la suture entre les fragments d'Okazaki pour produire un brin d'ADN continu.

11. Terminaison de la réplication : le système Ter/Tus d'*E.coli*

La fin de la réplication du chromosome bactérien implique la rencontre des deux fourches de réplication dans la région terminale du chromosome, située à l'opposé de l'origine. La terminaison de la réplication peut se faire de manière active comme chez *E. coli* (système *ter/Tus*) et *Bacillus subtilis* (système RTP-*ter*) ou alors de manière passive par simple fusion des fourches convergentes comme chez *Vibrio cholerae*. La terminaison de la réplication au niveau de régions de terminaison dédiées, garantit que les réplichores sont de taille similaire et que les dernières régions répliquées sont accessibles à d'autres machineries cellulaires pour les étapes finales du traitement du chromosome, telles que la décaténation et la résolution des dimères de chromosomes (Lesterlin et al., 2004).

Chez *E. coli*, la région terminale contient dix sites ADN de 23pb orientés nommés *ter*, ces sites sont reconnus par la protéine Tus (Coskun-Ari and Hill, 1997). Le rôle du complexe nucléoprotéique *ter*/Tus est de bloquer la progression des fourches de réplication de manière polaire.

La protéine Tus se lie au site *ter*, Tus est asymétrique avec une face qui ne permet pas le passage de l'hélicase DnaB (à gauche), et une face qui permet le passage (à droite). Lorsqu'un réplisome s'approche de la face non permissive du complexe Tus-*ter*, une cytosine en position 6 du site *ter* se retourne hors du duplex et se lie fortement à la protéine Tus. Puis lorsque le réplisome atteint ce complexe verrouillé il se décroche de l'ADN. Enfin, un deuxième réplisome s'approche de la face permissive du complexe Tus-*ter* et décroche la protéine Tus de l'ADN. La réplication peut alors se terminer (Figure 15) (Kaplan, 2006).

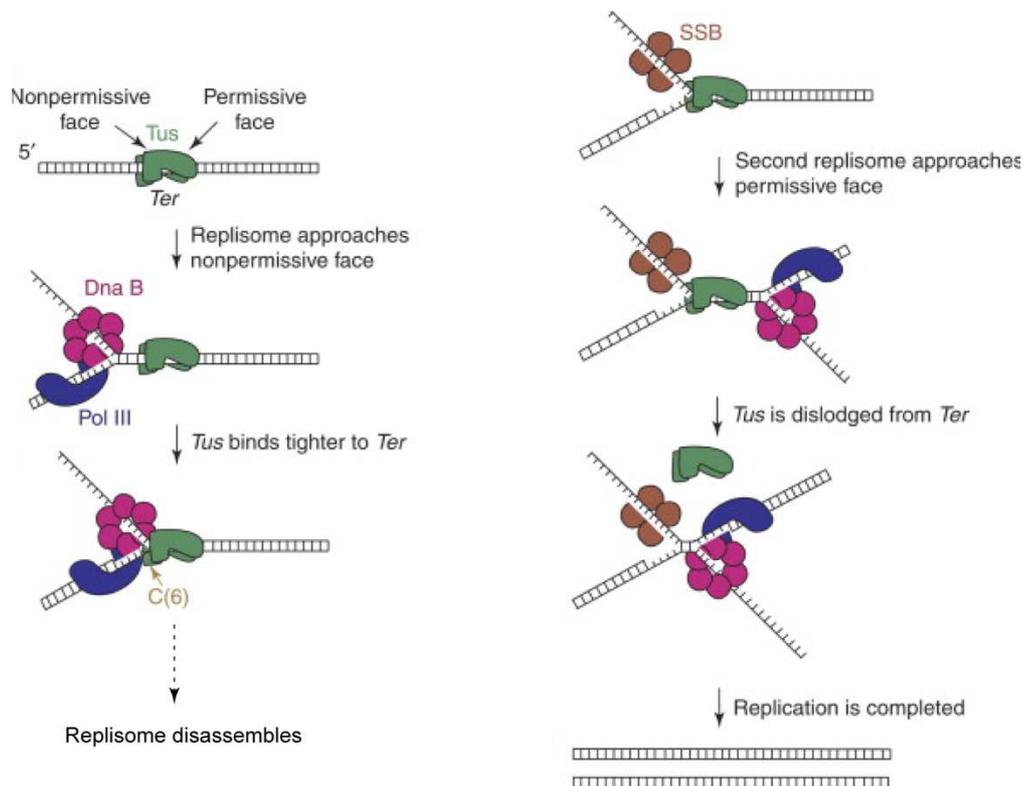


Figure 15. Modèle de terminaison de la réplication chez *E. coli*. Blocage polaire d'un réplisome par le complexe *Ter*/Tus permettant la terminaison de la réplication. Adapté de (Kaplan, 2006).

III. Réplication des chromosomes secondaires bactériens

Les chromosomes secondaires bactériens, aussi appelés chromides, sont des réplicons qui ont évolué à partir de mégaplasmides. Au cours d'un processus de domestication, ils ont acquis un ou plusieurs gènes de ménage ainsi qu'une signature génomique similaire à celle du chromosome principal (taux de GC, utilisation de codon) (Figure 16).

Un des traits attestant de leur passé plasmidique est la présence chez tous les chromides d'une origine de réplication de type plasmidique. Ces origines sont principalement de deux types et permettent de séparer les chromides en deux catégories. La première est basée sur la présence de l'opéron *repABC*, un système à trois gènes où RepC est l'initiateur et RepA/RepB sont impliqués dans la ségrégation. La seconde est basée sur des courtes séquences répétées nommées itérons reconnues par un initiateur Rep encodé proche de l'origine.

Tous les chromides présents chez les Alpha-protéobactéries, dont ceux de *Rhodobacter sphaeroides* ou d'*Agrobacterium tumefaciens* portent un système de réplication de type *repABC* (diCenzo et Finan, 2017). A l'inverse, toutes les espèces appartenant à la famille des *Vibrionaceae* possèdent un chromosome secondaire dont le système de réplication est de type plasmide à itérons (Okada et al., 2005).

L'initiation de la réplication est étroitement contrôlée dans les chromosomes afin de garantir qu'elle ne se produise qu'une fois par cycle cellulaire. A l'inverse, celle des plasmides peut avoir lieu de façon aléatoire tout au long du cycle cellulaire (Nordstrom, 2006). Dans le cas des chromides, la machinerie d'initiation de type plasmidique a évolué vers un mode de régulation chromosomique pour permettre une réplication unique par cycle cellulaire.

La majorité des connaissances sur la biologie des chromides provient de l'étude de *Vibrio cholerae*, cet organisme est considéré comme le modèle d'étude des génomes bactériens multipartites. Son génome est divisé en deux réplicons avec un chromosome principal et un chromosome secondaire contenant une origine de type plasmide à itérons.

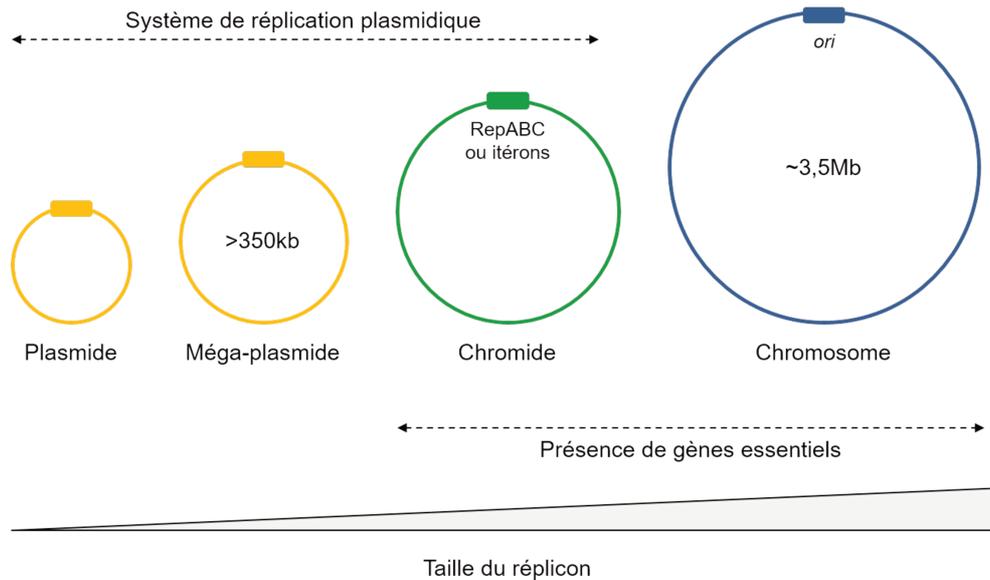


Figure 16. Les différentes classes de réplicons bactériens. Représentation schématique des 4 classes de réplicons que peuvent contenir les génomes bactériens. Les plasmides (<350kb) et les méga-plasmides (>350kb) (en jaunes) portent exclusivement des gènes adaptatifs. Les chromosomes et les chromides portent des gènes de ménages. Les plasmides, mégaplasmides et les chromides ont des origines de répllication de type plasmidique tandis que le chromosome est toujours initié par DnaA adapté de (Fournes et al., 2018).

1. Organisation des origines de type RepABC

Les génomes des Alpha-protéobactéries contiennent fréquemment de larges réplicons secondaires (diCenzo and Finan, 2017). Quelle que soit leur nature, mégaplasmides ou chromides, la répllication et la ségrégation de ces réplicons impliquent toujours trois gènes organisés en opéron : *repABC* (Figure 17) (Petersen et al., 2009).

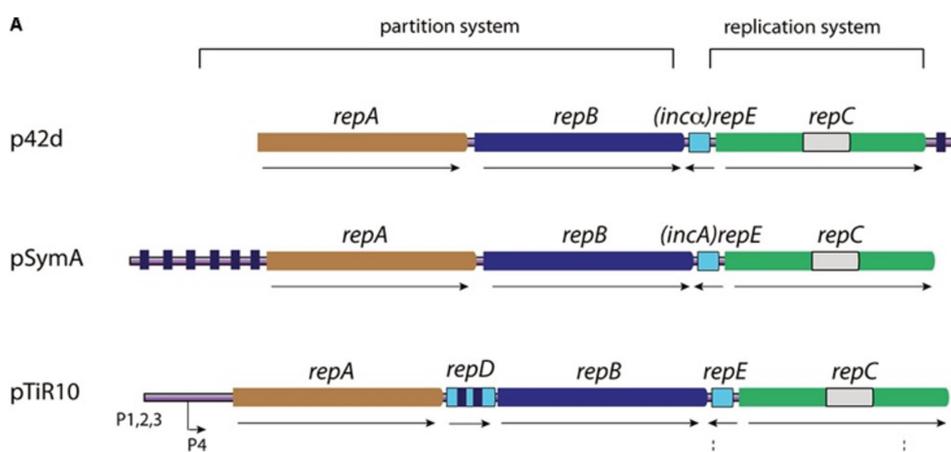


Figure 17. Organisation de l'origine de réplicon de type RepABC. Différents exemples de réplicon de type RepABC montrant la conservation de l'organisation fonctionnelle avec les gènes RepAB pour la partition et RepC pour l'initiation (Fournes et al., 2018).

Les protéines codées par l'opéron *repABC* sont impliquées dans deux mécanismes distincts ; RepC est l'initiateur de la réplication, RepA et RepB sont des protéines de partition. Ces trois gènes *repA*, *repB* et *repC* sont généralement exprimés à partir d'un même promoteur situé en amont de *repA*.

La protéine RepC est la protéine initiatrice des réplicons RepABC, et est présente uniquement chez les Alpha-protéobactéries (Petersen *et al.*, 2009). L'origine de réplication *oriV* est située dans la région codante du gène *repC*, si bien que le gène seul est suffisant pour initier la réplication d'un mini-réplicon (Pinto *et al.*, 2012). Dans le plasmide pTiR10 d'*Agrobacterium tumefaciens*, RepC se lie à une région de 150pb contenu dans sa propre séquence codante *repC*. Celle-ci contient un site en dyade avec deux régions répétées inversées imparfaites près d'une région riche en AT (Pinto *et al.*, 2012).

RepC présente certaines homologues avec d'autres initiateurs de réplication. Sa structure est divisée en deux domaines : un domaine N-terminal (NTD) des résidus 1 à 265 et un domaine C-terminal (CTD) des résidus 298 à 439. Les deux domaines NTD et CTD sont reliés par une région linker (Pinto *et al.*, 2012). Le domaine NTD de pTiR10 RepC contient un motif hélice-tour-hélice qui permet la fixation à l'ADN et qui est structurellement similaire aux facteurs de transcription de la famille MarR.

Le domaine CTD est incapable de lier l'ADN seul mais assure la spécificité de la fixation (Pinto *et al.*, 2012). Ce domaine CTD permettrait également la dimérisation de RepC et jouerait un rôle dans l'incompatibilité entre différents réplicons *repABC* (Cervantes-Rivera *et al.*, 2011).

De manière surprenante, l'initiateur RepC fonctionne exclusivement en *cis*. En effet, la surexpression de RepC dans *A. tumefaciens* induit une augmentation du nombre de copies de plasmides en *cis*, mais ne change pas le nombre de copies d'un second plasmide contenant également *oriV* en *trans* (Pinto *et al.*, 2012). Alors que les chromides à itérons aient été étudiés de manière relativement approfondie, notre connaissance des réplicons RepABC reste encore limitée, en particulier en ce qui concerne l'initiation de la réplication. Cette dernière a été relativement peu étudiée, ce qui rend difficile la compréhension de sa régulation.

2. La réplication des origines à itérons

L'autre grand type de chromide qu'on retrouve chez tous les *Vibrionaceae*, contient une origine proche de celle des plasmides à itérons. Les plasmides à itérons sont des éléments génétiques présents chez de nombreuses bactéries Gram-négatives. Ces plasmides portent pour la plupart des gènes de résistance aux antibiotiques, ils peuvent également apporter d'autres gènes accessoires comme des gènes impliqués dans la dégradation d'un composé spécifique ou encore des gènes de virulences.

Ces plasmides sont dits « à itérons » car ils possèdent dans leur origine de réplication des séquences répétées et dirigées caractéristiques nommés itérons. On peut pour exemple citer les plasmides à itérons les plus communément étudiés : R6K, RK2, F, P1, pSC101, pPS10.

Tous ces plasmides encodent à proximité de leur origine de réplication une protéine initiatrice nommée Rep. La protéine Rep joue un rôle multiple dans la régulation de la réplication du plasmide. Tout d'abord, Rep se fixe aux itérons et permet l'initiation de la réplication en favorisant l'ouverture de l'ADN ainsi que le recrutement du réplisome.

3. Structure de l'origine d'initiation de la réplication des plasmides à itérons

Les plasmides à itérons portent une origine de réplication qui diffère des origines chromosomiques, ils encodent leur propre protéine initiatrice nommé Rep. Ces deux éléments permettent une réplication autonome du plasmide, indépendamment de la réplication du chromosome et du cycle cellulaire. Les itérons sont des séquences répétées dont la longueur varie de 17 pb dans le cas du plasmide RK2 à 22 pb dans le plasmide R6K. Le nombre d'itérons et l'espacement entre les itérons peuvent également varier d'un plasmide à l'autre (Figure 18).

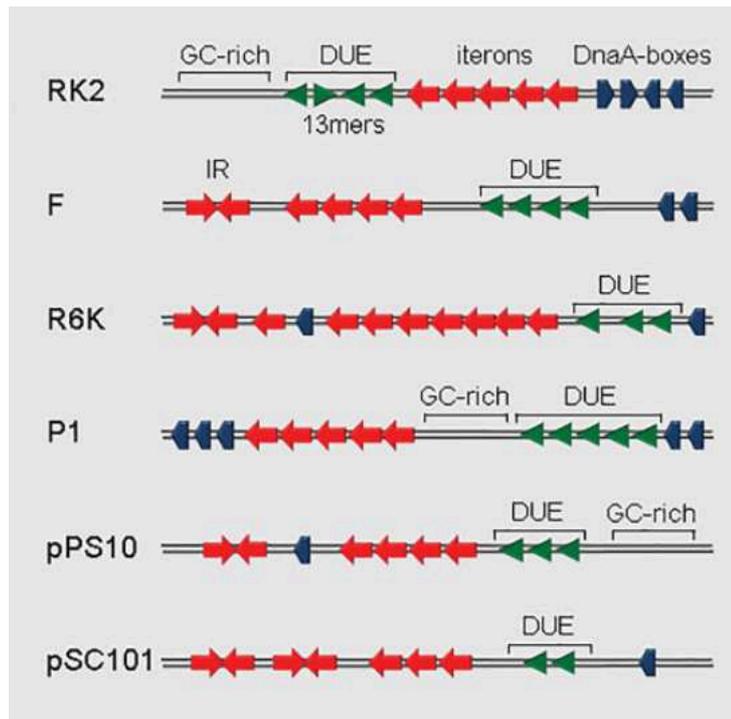


Figure 18. Schéma de la structure d'origine d'un plasmide contenant des itérons. Les répétitions directes - itérons - et les répétitions inversées (RI) sont représentées par des flèches rouges. La région DUE de chaque origine est indiquée, et les séquences répétées dans cette région sont représentées par en vert. Les DnaA-box sont en bleu. La région riche en GC est indiquée (Konieczny et al., 2014).

La liaison de Rep à l'ADN double brin (ADNdb) sur les itérons de l'origine entraîne une déstabilisation locale du duplex d'ADN et l'ouverture des deux brins de l'ADN au niveau de la région riche en AT. La protéine Rep est accompagnée dans son action initiatrice par la protéine DnaA de l'hôte. En effet, l'origine des plasmides à itérons contient également une ou plusieurs DnaA-boxes qui peuvent être localisés à différents endroits sur celle-ci.

Tout comme les origines chromosomiques, on retrouve dans les origines plasmidiques une région riche en AT. C'est au niveau de cette séquence située proche des itérons que l'ouverture locale de la double hélice d'ADN se produit lors de l'initiation de la réplication.

Dans la région riche en AT, on retrouve de courtes séquences répétées, généralement orientées directement (par exemple, l'origine des plasmides RK2 et pSC101). Ces séquences sont reconnues par Rep et permettent la polymérisation de Rep le long de l'ADN simple brin.

On a détecté également dans certaines origines de plasmides comme c'est le cas pour *oriC*, le site de liaison de la protéine IHF. La liaison de la protéine IHF entraîne la courbure de la molécule d'ADN ; celle-ci est essentielle pour l'activité de l'origine.

De manière générale, on ne retrouve pas de sites de méthylation GATC dans l'origine des plasmides à itérons comme c'est le cas pour *oriC*. Seul le plasmide P1 présente des sites GATC dans sa région riche en AT (Brendler et al., 1991). Ces sites sont essentiels à l'initiation car ils facilitent l'ouverture du duplex d'ADN, ils permettent par la même occasion la séquestration de l'origine par SeqA.

Certains plasmides à itérons (RK2, P1, pPS10, IncQ) portent également dans leur origine de réplication une région riche en GC qu'on ne retrouve pas dans les origines chromosomiques. Son rôle exact est encore mal compris mais elle semble servir de région *spacer* entre les itérons et la région riche en AT.

4. Comparaison de l'ouverture de l'ADN entre chromosome et plasmides à itérons

La première étape de l'initiation de la réplication des plasmides à itérons est la formation d'un complexe initial via l'interaction spécifique et coopérative des protéines Rep avec les itérons. De manière analogue au système DnaA-oriC, l'interaction de Rep avec les itérons génère également une déstabilisation localisée du duplex d'ADN au niveau de la région riche en AT.

Ce processus d'ouverture nécessite généralement différentes protéines de l'hôte telles que DnaA, SSB, IHF ou encore HU (Konieczny *et al.*, 2014). DnaA est principalement nécessaire pour renforcer ou stabiliser la formation du complexe ouvert. SSB protège l'ADN simple brin et participe également à la stabilisation de l'ouverture (Figure 19). L'interaction des protéines de type histone (HU et/ou IHF) aident au repliement de l'ADN.

Par exemple : dans le cas du plasmide F, la liaison de l'initiateur RepE aux itérons induit une ouverture localisée dans la région d'origine seulement en présence de HU. L'ajout de DnaA stimule le processus de l'ouverture de l'origine du plasmide F (Konieczny *et al.*, 2014).

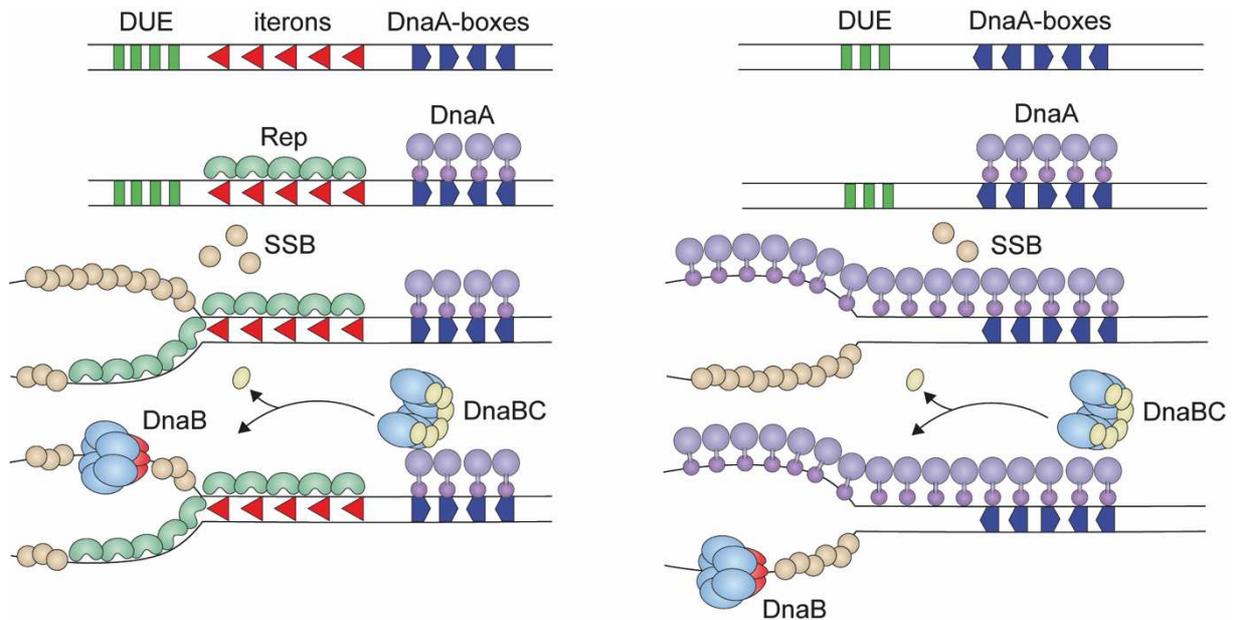


Figure 19. Comparaison des modèles de l'initiation de l'origine des plasmides à itérons avec celle des origines chromosomiques. La protéine initiateur codée par le plasmide (Rep) reconnaît et se lie aux itérons. L'interaction de Rep avec les itérons conduit à la formation d'un complexe ouvert. DnaA aide au recrutement du complexe DnaBC sur l'ADN simple brin. Par exemple : il a été montré que les interactions entre la protéine hôte DnaB et les initiateurs π_{R6K} , RepE_F et RepA_{pSC101} sont importantes pour le recrutement de l'hélicase (Konieczny *et al.*, 2014).

5. Structure des initiateurs de la famille Rep

Les protéines Rep sont composées de deux domaines nommés WH1 (N-terminal) et WH2 (C-terminal), chacun de ces domaines contient un motif *winged-helix* qui permet l'interaction avec l'ADN (Giraldo *et al.*, 1998). Contrairement à DnaA, on ne retrouve pas chez les protéines Rep de domaine AAA+ ATPase.

La structure de plusieurs initiateurs de type Rep a été élucidée (RepA_{pPS10}, RepE_F, π_{R6K}), cependant la majorité des informations sur les différentes conformations des Rep provient de l'étude de la protéine RepA du plasmide pPS10. Les protéines Rep peuvent se fixer à deux types de sites : itéron et opérateur. La fixation aux itérons permet l'initiation de la réplication du plasmide alors que la fixation à l'opérateur permet l'auto-répression du promoteur de Rep. La fixation sur ces deux types de sites nécessite des changements conformationnels dans la structure de la protéine.

Rep se fixe sur les itérons sous la forme d'un monomère, l'interaction avec l'ADN se fait avec les deux domaines WH1 et WH2. Le domaine WH2 se lie à la moitié 3' de l'itéron, tandis que le domaine WH1 change de conformation et entre en contact avec l'extrémité 5' de l'itéron (Giraldo *et al.*, 1998).

Rep se fixe à son opérateur sous la forme d'un dimère, l'interaction avec l'ADN se fait via les deux domaines WH2. Les deux domaines WH1 forme une interface de dimérisation (Diaz-Lopez et al., 2006). Le domaine WH2 permet donc la liaison à la fois aux itérons et aux répétitions inversées dans le promoteur de Rep (opérateur).

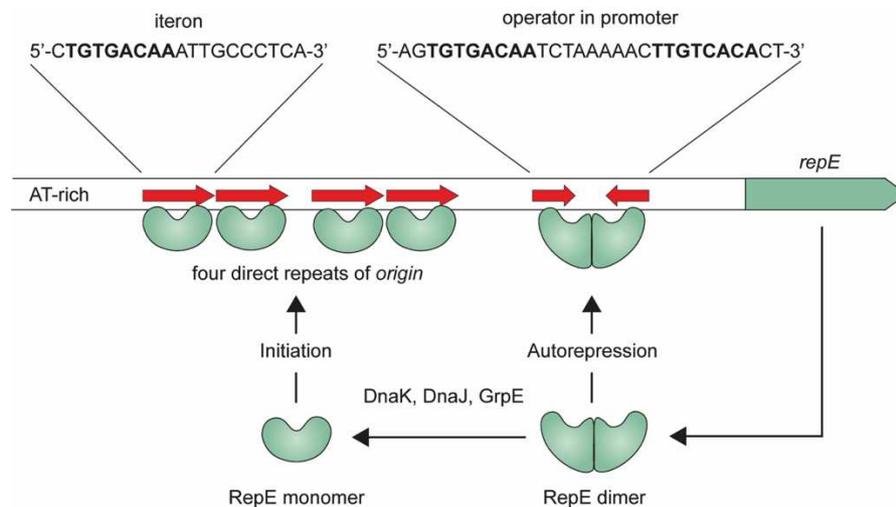


Figure 20. Reconnaissance d'une origine de plasmide à itérons par RepE. La protéine RepE est exprimée proche de l'origine de réplication du plasmide. Dans le cytoplasme RepE est sous forme dimérique et se fixe à son propre promoteur. RepE est principalement monomérisé par des protéines chaperonnes, sous sa forme monomérique il peut se fixe aux itérons (adapté de (Komori et al., 1999)).

Rep est principalement sous forme de dimères dans la cellule, cette dimérisation se fait via les domaines WH1 (Toukdarian et al., 1996). Dans cette forme dimérique, Rep est considéré comme « inactif » car il ne peut pas se fixer fortement aux itérons. Ces dimères peuvent être dissociés de deux manières via des changements de conformations.

Le mécanisme principal de dissociation se fait par l'action des protéines chaperonnes ClpB, DnaK, DnaJ et GrpE (Konieczny and Liberek, 2002). Le second est l'interaction directe avec un itéron via le domaine WH2 qui permet de remodeler le domaine WH1 depuis une interface de dimérisation vers un domaine de liaison à l'ADN (Diaz-Lopez et al., 2003). A noter que lors de la fixation sur les séquences inversement répétées de de son opérateur, le dimère de Rep ne subit pas de dissociation ni de changement de conformation (Nakamura et al., 2007) (Figure 20).

6. *Handcuffing* : mécanisme principal de la régulation négative de la réplication

Le mécanisme principal permettant de réguler négativement le nombre de copies des plasmides à itérons est nommé *handcuffing* (menottage). Ce dernier consiste en l'appariement des origines de deux plasmides via des interactions entre les protéines Rep fixées à l'ADN et inhibant l'initiation de la réplication par encombrement stérique (Figure 21).

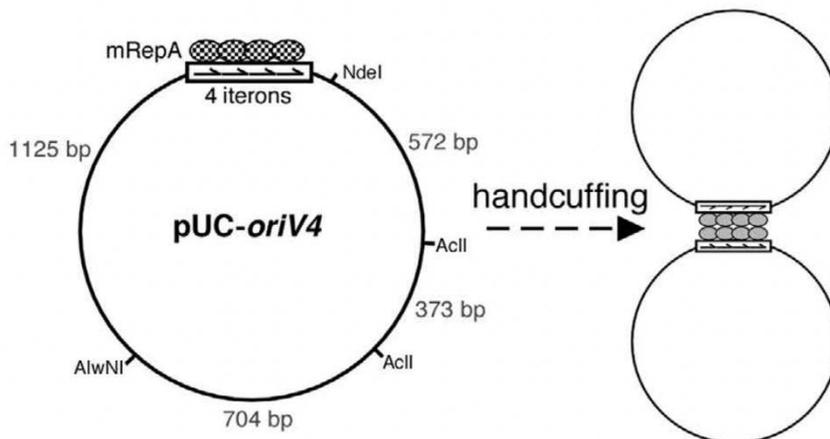


Figure 21. Appariement des origines *oriV* du plasmide pPS10 par *handcuffing*. Après la réplication, les origines du plasmide pPS10 sont couplées par l'initiateur RepA inhibant l'initiation de la réplication.

Ce mécanisme de régulation implique un complexe nucléoprotéique spécifique formé sur des répétitions directes (itérons) situées dans l'origine de réplication du plasmide. Il contrôle l'initiation de la réplication de l'ADN et donc le nombre de copies et le maintien stable du plasmide. Le *Handcuffing* empêche un nouveau cycle d'initiation de la réplication en inhibant la fusion de l'origine, ce qui est important pour le contrôle du nombre de copies du plasmide. Les détails exacts des structures des complexes de *handcuffing* sont encore débattus et semblent ne pas être les mêmes pour tous les plasmides à itérons.

Dans le cas du plasmide pPS10, son initiateur RepA contient deux domaines WH1 et WH2. Sous sa forme *libre* RepA est un dimère qui interagit via son domaine N-terminal WH1. Lors de la liaison aux itérons de l'origine (*oriV*) via le domaine WH2 ainsi que l'aide de DnaK/J et GrpE, le domaine WH1 est remodelé pour se fixer à l'ADN également. Une fois répliquées, les deux molécules d'ADN plasmidique sont couplées via un complexe d'*handcuffing* entre les molécules RepA fixées aux deux *oriV*. Ce complexe est médié par des monomères de RepA interagissant via leur domaine WH1 en impliquant une interface distincte de celle trouvée dans les dimères de RepA libres. On parle même de structure « amyloïde » pour évoquer ces oligomères d'initiateurs (Figure 22).

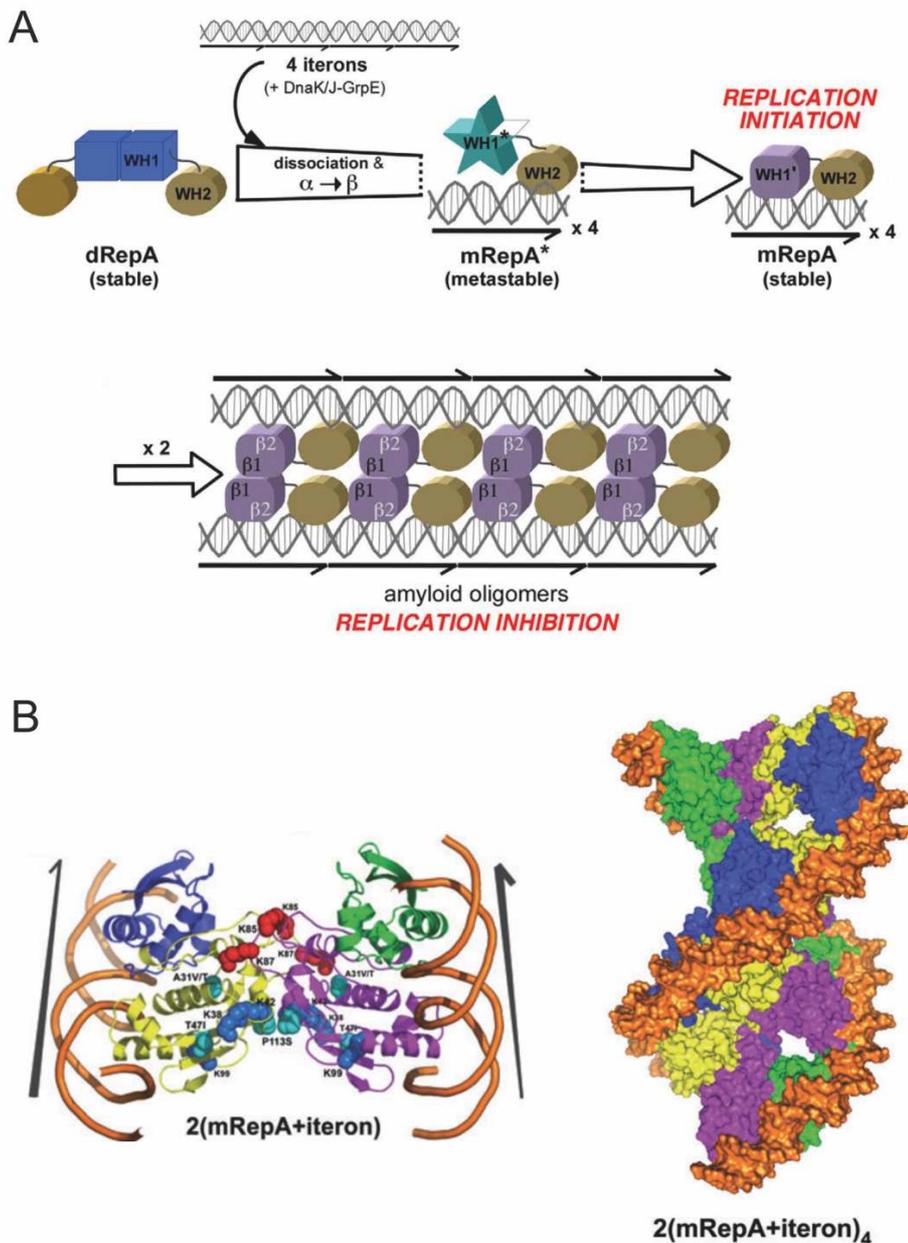


Figure 22. Modèle d'handcuffing de l'origine de pPS10 médié par l'initiateur RepA. A. La forme libre de RepA est un dimère interagissant via le domaine WH1. Sa dissociation est facilitée par la fixation aux itérons et par l'action des chaperonnes DnaK/J et GrpE. Le domaine WH2 permet l'attachement à l'ADN et le remodelage de WH1. Une fois remodelé WH1 interagit également avec l'ADN et permet une fixation stable. Après réplication une seconde interface de dimérisation de WH1 permet de coupler les deux origines. B. Détails du modèle d'interaction de RepA par le domaine WH1 dans un complexe de handcuffing (Gasset-Rosa et al., 2008); (Molina-Garcia et al., 2016).

Dans le cas des plasmides P1 (Das and Chattoraj, 2004), R6K (Kunnimalaiyaan et al., 2005) et RK2 (Bury et al., 2017) le complexe d'handcuffing adopterait une organisation différente de celle décrite précédemment. Ce complexe se formerait uniquement à partir de

dimères de Rep interagissant via leur domaine WH1 mais sans la participation de WH1 dans la fixation à l'ADN (Figure 23).

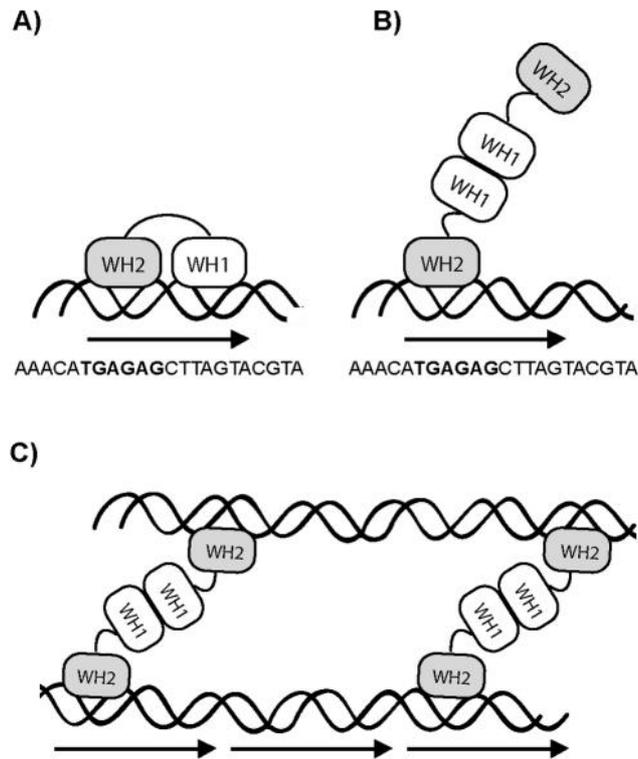


Figure 23. Modèle du complexe d'handcuffing pour le plasmide R6K. Pour le plasmide R6K, l'initiateur Pi médie le handcuffing entre deux origines par une auto-interaction via le domaine WH1, sans que celui-ci ne participe à la fixation de l'ADN (Kunnimalaiyaan *et al.*, 2005).

Enfin, pour certains plasmides, le complexe d'*handcuffing* se forme entre l'origine et un autre site ADN spécifique sur la même molécule, on parle alors d'*handcuffing* intramoléculaire. C'est le cas pour le plasmide F, qui est un plasmide en moyenne à une seule copie par cellule et qui porte un système actif de partition (ParB). Dès que le plasmide est répliqué, les deux molécules d'ADN sont ségréguées de part de d'autre de la cellule via l'action de ParB, ainsi les origines des deux plasmides néoformés ne sont quasiment jamais en contact. Pour ce plasmide F, la régulation négative de l'initiation se fait via un *Handcuffing* intramoléculaire entre les régions *oriF* (origine minimale) et *incC* (région d'incompatibilité). En effet, l'initiateur RepE se fixe aux itérons de l'*oriF* mais aussi à des itérons de *incC*, ces deux complexes interagissent ensemble et forme une boucle d'ADN inhibant l'initiation. Dans le modèle établi, la forme monomérique de RepE permet la liaison à l'ADN et la forme dimérique permet le lien entre les deux régions d'ADN (Zzaman and Bastia, 2005) (Figure 24).

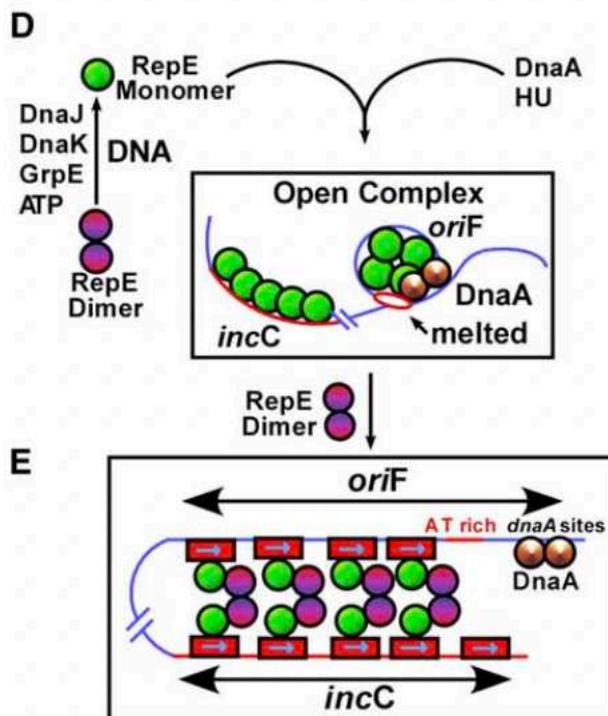


Figure 24. Handcuffing intramoléculaire dans le plasmide F. L'initiateur RepE permet l'initiation de la réplication lorsqu'il se fixe aux itérons de l'*oriF* sous forme monomérique. RepE inhibe également l'initiation en forme une boucle d'ADN en reliant les itérons de l'*oriF* et ceux de *incC*, ce complexe de handcuffing fait intervenir à la fois la forme monomérique et dimérique (Zzaman and Bastia, 2005).

Nous allons voir par la suite que les chromides à itérons, notamment le Chr2 de *Vibrio cholerae*, possèdent une origine ressemblant étonnamment à celle de ce plasmide F.

IV- *Vibrio cholerae* : un modèle d'étude des bactéries à génome multipartite

Vibrio cholerae est une bactérie gram-négative qui appartient à la famille des *Vibrionaceae* et au groupe des gamma-protéobactéries. Elle est l'agent pathogène responsable du choléra et également le modèle principal d'étude des bactéries à génome multipartite. Le génome de *V.cholerae* est constitué de deux réplicons de taille différente, dénommés Chromosome 1 (Chr1, 3 Mb) et Chromosome 2 (Chr2, 1 Mb) (Trucksis et al., 1998).

La désignation de « chromosome » est basée sur leur caractère essentiel (ils sont tous deux indispensables à la croissance de *V.cholerae*) et est couramment utilisée dans la littérature et les bases de données. Cependant, un examen plus approfondi des caractéristiques de ces deux réplicons indique qu'ils ne doivent pas être considérés comme des membres du même groupe. En effet, le Chr1 porte la grande majorité des gènes de ménages, dont la synténie est conservée entre les espèces et les niveaux supérieurs de taxonomie (Heidelberg et al., 2000). Le Chr2 ne contient que quelques gènes de ménages et il présente également une densité beaucoup plus importante d'informations génétiques spécifiques à l'espèce ou à la souche, dont une partie est accumulée dans un superintégron (Heidelberg et al., 2000), (Mazel et al., 1998).

Chaque réplicon porte sa propre machinerie d'initiation de la réplication (Duigou et al., 2006). La réplication de Chr1, qui est initiée à *ori1*, dépend de l'initiateur universel DnaA (Duigou et al., 2006). L'origine du Chr1 est un homologue fonctionnel d'*oriC* situé sur le chromosome de *E. coli* (Demarre and Chatteraj, 2010), et est apparentée aux origines chromosomiques trouvées dans presque toutes les bactéries (Mackiewicz et al., 2004). La réplication de Chr2, qui commence à *ori2*, dépend de l'initiateur RctB encodé proche de l'origine, qui partage des similarités structurales avec les initiateurs des plasmides à itérons (Orlova et al., 2017). Par conséquent, l'ancêtre du Chr2 doit se trouver parmi les plasmides. L'ancêtre plasmidique de Chr2 doit avoir été domestiqué pour devenir un réplicon stable et essentiel du génome de *Vibrio* en incorporant sa réplication dans le cycle de la cellule hôte (discuté en détail ci-dessous) ainsi qu'en acquérant quelques gènes de ménage du Chr1. À la lumière de cette histoire évolutive et selon les classifications actuelles (Harrison et al., 2010), (diCenzo and Finan, 2017), le Chr2 peut être considéré comme un chromide (en partie chromosome, en partie plasmide), tandis que le Chr1 est le véritable chromosome bactérien.

L'origine de réplication du Chr2 est un locus de 3kpb qui se divise en trois unités fonctionnelles : la région minimale d'initiation *ori2_{min}* (0.4kb), une région de régulation négative nommée *inclI* (0.6kb) et le gène codant l'initiateur *rctB* (2kb). L'initiateur RctB se fixe sur deux types de sites présents sur l'origine les itérons (11mer ou 12mer) et les 39m (39mer ou 29mer) (Figure 25).

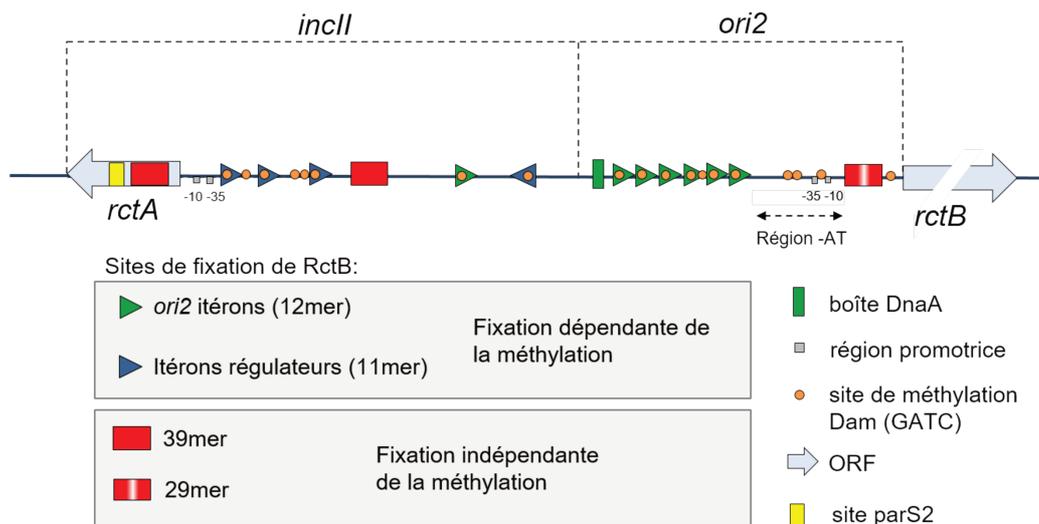


Figure 25. Organisation fonctionnelle de l'origine du chromosome 2 de *V.cholerae*. L'origine du Chr2 comporte trois unités fonctionnelles *inclI*, *ori2* et le gène *rctB*. RctB se fixe sur 4 types différents de sites 11mer, 12mer, 29mer, 39mer. La région *ori2* contrôle positivement la réplication du Chr2. La fixation de RctB sur les 6x12mer permet l'ouverture des 2 brins d'ADN au niveau de la région riche en AT et le démarrage de la fourche de réplication. La région *inclI* contrôle négativement la réplication du Chr2 via handcuffing et titration.

Les itérons sont des sites non palindromiques activateurs de l'initiation du Chr2. La région minimale d'initiation (*ori2_{min}*) porte six sites 12mer (6x12mer), tous ces sites sont dans la même orientation en phase hélicoïdale : c'est-à-dire que le même pas d'hélice sépare deux sites l'un de l'autre de telle sorte qu'ils soient tous alignés à la surface de la molécule d'ADN. Ces sites 6x12mer, lorsqu'ils sont fixés par RctB, permettent l'ouverture des deux brins d'ADN au niveau de la région riche en AT (Demarre and Chatteraj, 2010).

1. Un génome bipartite conservé chez les *Vibrionaceae*

L'organisation bipartite du génome est conservée dans toute la famille *Vibrionaceae*, avec le Chr1 dont la taille varie entre 1,4 Mpb et 4,1 Mpb et le Chr2 entre 0,78 Mpb et 2,4 Mpb (Figure 26). Le niveau de conservation de l'organisation multipartite est très élevé par rapport

à d'autres espèces bactériennes à génome multipartite, où la présence de chromides est généralement limitée à un seul genre et où d'autres espèces mono-chromosomiques sont présentes dans la même famille (Harrison *et al.*, 2010), (diCenzo *et al.*, 2019). On rapporte quelques exemples connus de *Vibrionaceae* mono-chromosomique, il s'agit cependant toujours d'évènement de fusion entre le Chr1 et le Chr2 (Xie *et al.*, 2017), (Bruhn *et al.*, 2018), (Yamamoto *et al.*, 2018). Des exemples similaires de co-intégrations ont également été observés chez d'autres bactéries à génome multipartite : *Sinorhizobium meliloti* (Guo *et al.*, 2003) et *Cupriavidus necator* (Mori and Kanaly, 2022). La conservation du génome bi-partite chez les *Vibrionaceae*, ainsi que les expériences d'évolution en laboratoire sur des fusions artificielles Chr1+Chr2 (qui ont tendance à s'inverser spontanément après plusieurs générations) (Val *et al.*, 2016), indiquent que cette architecture génomique est sélectionnée positivement et est importante pour la fitness des *Vibrio*.

Au-delà du « core genome » bipartite, les souches de *Vibrionaceae* portent un ensemble très variable de plasmides et de mégaplasmides, qui ne sont pas essentiels à la croissance de la cellule hôte, mais peuvent assurer certaines fonctions d'adaptation (Harrison *et al.*, 2010), (Hall *et al.*, 2022). Le 11/11/2022, la PLSDB (The Plasmid Database ; <https://ccb-microbe.cs.uni-saarland.de/plsdb>) contenait 331 plasmides de *Vibrionaceae* d'une taille comprise entre 2 kb et 1 Mb, dont 26 plasmides de souches de *V. cholerae* d'une taille comprise entre 2 et 354 kb (Figure 26). Il est intéressant de noter que les plasmides sont absents dans les souches de *V. cholerae* El Tor, responsables de la dernière pandémie de choléra. Il a été démontré que *V. cholerae* El Tor est porteur de deux systèmes de défense localisés dans deux îlots génomiques (l'îlot de pathogénicité 2 de *Vibrio*, VPI-2, et l'îlot II de *Vibrio* septième pandémie, VSP-II), qui rendent les plasmides instables et les éliminent de la population (Jaskolska *et al.*, 2022), ce qui constitue un exemple intéressant de barrières pour le transfert horizontal de gènes augmentant la stabilité du génome.

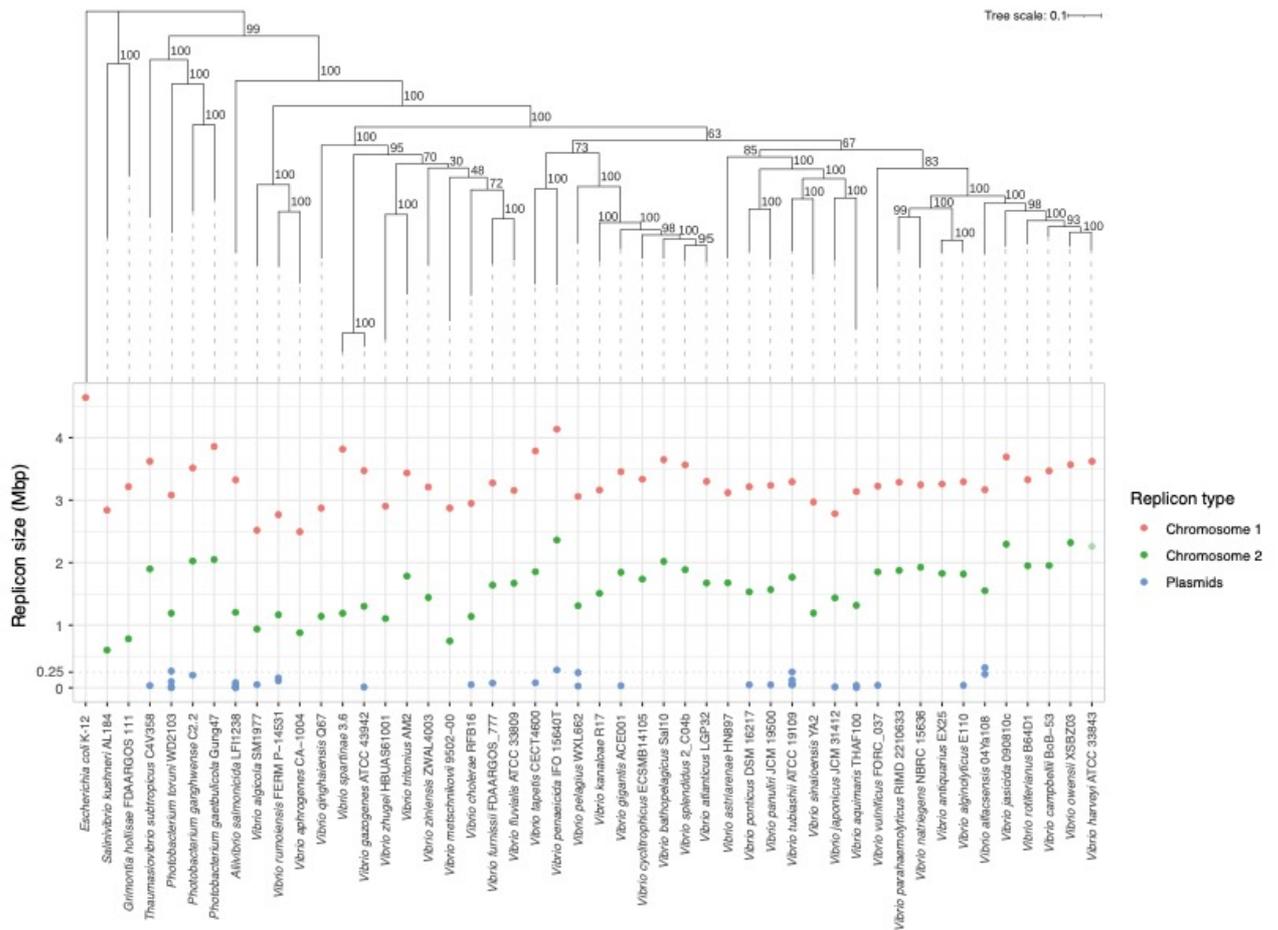


Figure 26. L'organisation du génome bipartite est conservée chez les *Vibrionaceae*. En haut du graphique est représenté l'arbre phylogénétique des *Vibrionaceae*. En bas sont représentés les différents réplicons Chr1 en rouge, Chr2 en vert et les plasmides en bleu.

2. Initiation du Chr1 par un mécanisme similaire à *oriC*-DnaA

Les deux chromosomes de *V. cholerae* initient la réplication par des mécanismes distincts. Alors que le Chr1 suit un processus d'initiation chromosomique standard, le Chr2 a un mécanisme d'initiation ressemblant à celui des plasmides à itérons (Duigou *et al.*, 2006), (Egan and Waldor, 2003). Le Chr1 est répliqué par l'initiateur bactérien universel DnaA, de façon similaire à la réplication du chromosome chez *E. coli*. Les origines de réplication (*ori*) du Chr1 de *V. cholerae* (*ori1*) et le chromosome de *E. coli* (*oriC*) partagent 58% d'identité de séquence en nucléotides, la présence de cinq boîtes DnaA fortes, une région riche en AT, et une boîte IHF (Figure 27). En outre, les protéines DnaA de *V. cholerae* (DnaA_{VC}) et de *E. coli* (DnaA_{EC}) présentent 79 % d'identité en acides aminés. Il a été démontré qu'un plasmide p*ORI1* peut se répliquer dans *E. coli* (Egan and Waldor, 2003), et qu'*ori1* peut remplacer fonctionnellement l'*oriC* d'*E. coli* (Demarre and Chattoraj, 2010), (Koch et al., 2010). Ces résultats suggèrent que les systèmes DnaA_{EC}-*oriC* et DnaA_{VC}-*ori1* sont pleinement compatibles.

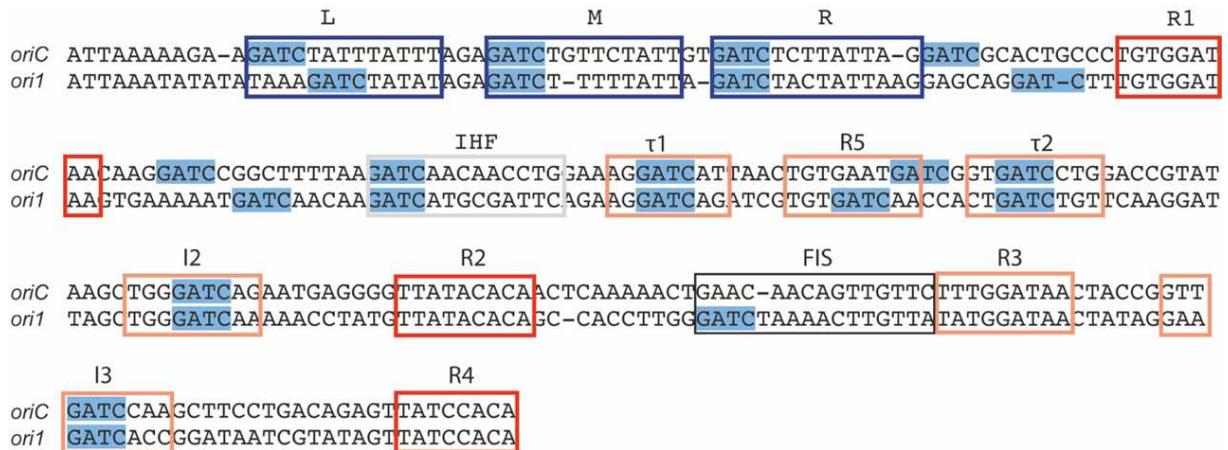


Figure 27. Alignement des régions *oriC* d'*E.coli* et *ori1* de *V.cholerae*. L'organisation de l'*oriC* et de l'*ori1* est très conservée chez les deux organismes.

Chez *E.coli*, divers mécanismes de régulation de l'activité de DnaA fonctionnent ensemble pour assurer une initiation robuste de la réplication à tous les taux de croissance (Katayama *et al.*, 2017), (Berger and Wolde, 2022). Bien que ces mécanismes n'aient pas été étudiés directement, on pense que la plupart sont également présents chez *V.cholerae*. L'origine de réplication du Chr1 chez *V.cholerae* est caractérisée par une surreprésentation de sites GATC, qui sont globalement conservés entre *V.cholerae* et *E.coli* (Koch *et al.*, 2010), (Demarre and Chatteraj, 2010) (Figure 27). Ces sites sont reconnus et méthylés par l'adénosine méthylase de Dam (Marinus and Lobner-Olesen, 2014). Lors de la réplication, les sites GATC deviennent transitoirement hémiméthylés et sont reconnus par la protéine SeqA, qui inhibe la ré-initiation de la réplication en empêchant le recrutement de DnaA pendant un tiers du cycle cellulaire (Nielsen and Lobner-Olesen, 2008). En plus de ce mécanisme, le Chr1 porte deux séquences de réactivation de DnaA DARS (Frimodt-Moller *et al.*, 2017), (Miyoshi *et al.*, 2021), qui sont présentes aux mêmes loci que chez *E.coli* dans les régions promotrices de *uvrB* et de *mutH* (Frimodt-Moller *et al.*, 2017).

Chez *E.coli*, le locus de titrage *datA* s'étend sur une région de 1 kb qui comprend à la fois la région intergénique *orn-queG* et la CDS du gène *queG* (Kasho and Katayama, 2013). La région intergénique entre *VC0341* (*orn*) et *VC0342* (*queG*) chez *V.cholerae* est également enrichie en séquences DnaA-box et IHF-box prédites. Ce locus semble donc bien être le site *datA* de *V.cholerae*.

Enfin, l'inactivation régulatrice de DnaA (RIDA) chez *E. coli* impliquant la protéine Hda semble fonctionner différemment chez *V. cholerae* car on ne retrouve aucun homologue de

Hda. RIDA étant le principal mécanisme de conversion de DnaA-ATP en DnaA-ADP, nous avons initié un projet au laboratoire pour comprendre comment fonctionne l'inactivation de la DnaA chez *V. cholerae*.

D'autres mécanismes secondaires qui régulent l'activité de DnaA seraient partagés par *E. coli* et *V. cholerae*. La protéine DiaA pourrait être VC0579, ce gène est un homologue de la famille DiaA/GmhA et il contient dans la partie C-terminale les résidus Leu190 et Phe191 spécifiques pour l'interaction avec DnaA (Keyamura et al., 2007).

Lors d'une limitation en acides aminés, l'augmentation des niveaux intracellulaires de ppGpp active la protéase Lon. Lon dégrade ensuite spécifiquement DnaA-ADP et inhibe ainsi l'initiation de la réplication (Gross and Konieczny, 2020). Ce mécanisme est conservé chez *E. coli* et chez *V. cholerae* (Kemter et al., 2018). Pendant l'activation de la réponse stringente, les fourches de réplication déjà assemblées peuvent poursuivre et achever la réplication du génome. Le laboratoire de Torsten Waldminghaus a montré que, comme pour *E. coli*, l'induction de la réponse stringente avec un analogue de la sérine (SHX) peut être utilisée pour synchroniser une population de cellules de *V. cholerae* (Ferullo et al., 2009). Il s'agirait d'ailleurs de la première méthode de synchronisation de la réplication pour Vibrio. Ils ont également observé que *E. coli* et *V. cholerae* ne réagissent pas exactement de la même manière à l'induction de la réponse stringente. Chez *E. coli*, son induction bloque complètement l'initiation de la réplication, les cellules terminent leur réplication mais ne se divisent pas. Chez *V. cholerae*, l'initiation est d'abord bloquée, puis un second cycle de réplication reprend et lorsque les chromosomes ont terminé leur réplication, les cellules se divisent. Ceci suggère l'existence d'autres mécanismes de régulation chez Vibrio (Kemter et al., 2019).

3. Le Chr1 et le Chr2 suivent une chorégraphie d'initiation

Les deux chromosomes ne sont présents qu'en une seule copie par cellule mais des cycles de réplication chevauchant apparaissent en milieu riche (Stokke et al., 2011). La terminaison de la réplication ne se produit qu'une fois par cycle cellulaire pour les deux chromosomes et est assurée par un réseau de régulation complexe. En effet, les deux chromosomes suivent une chorégraphie d'initiation. Tout d'abord, le chromosome principal est initié à *ori1* par la protéine DnaA en réponse aux paramètres de croissance cellulaire (Egan et al., 2004). Ensuite, lorsque les 2/3 du Chr1 ont été répliqués, le Chr2 commence à se répliquer

de telle sorte que les deux chromosomes terminent leur réplication de manière synchrone (Rasmussen et al., 2007) (Figure 28).

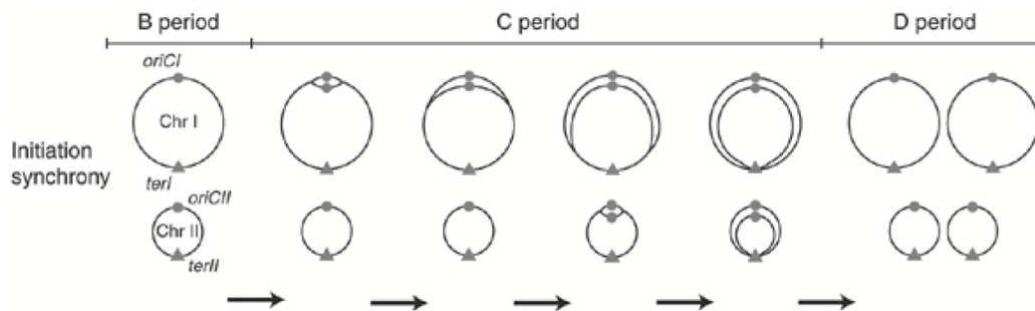


Figure 28. La terminaison de la réplication des deux chromosomes de *V.cholerae* est synchrone. (Rasmussen et al., 2007)

La terminaison synchrone de la réplication des deux chromosomes est une caractéristique sélectionnée par l'évolution chez les Vibrionaceae car elle est observée chez de nombreuses espèces : *Vibrio cholerae*, *Vibrio anguillarum*, *Vibrio coralliilyticus*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio nigripulchritudo*, *Vibrio parahaemolyticus*, *Vibrio tasmaniensis*, *Vibrio vulnificus*, *Photobacterium profundum*, *Aliivibrio fischeri* (Kemter et al., 2018).

Plusieurs hypothèses ont été proposées pour expliquer l'avantage sélectif d'une terminaison synchrone de la réplication. Ce processus peut réduire le besoin de chevauchement des cycles de réplication lors d'une prolifération rapide (Rasmussen et al., 2007). De plus, il pourrait réduire la complexité de la ségrégation et permettre le couplage avec la division cellulaire.

4. Activation de l'initiation du Chr2

L'origine minimale du Chromosome 2 a été étudiée et disséquée avec précision. Il s'agit d'une région d'ADN très complexe et compacte qui engage de nombreux régulateurs. Cette *ori2_{min}* est composée d'une boîte DnaA, d'un réseau de six itérons régulièrement espacés, d'une boîte de liaison pour IHF et d'une région riche en AT. La région riche en AT est la région où l'ADN est ouvert après la liaison de RctB aux itérons. De manière général, *ori2* tolère très peu de changements, car ni la position ni l'orientation des sites ne peuvent être modifiées. Les séquences DnaA-box ainsi que IHF-box sont essentielles. Ni la séquence ni l'espacement entre les itérons ne peuvent être modifiés. Un seul des six itérons est inactivable à la fois, certains de ces itérons ne peuvent pas être inactivés du tout (Gerding et al., 2015).

Lorsque l'initiation de l'origine est déclenchée, RctB se lie en coopération sur la série de six itérons (Gerding *et al.*, 2015). Puis, les protéines RctB attachées aux 6x12mer se lient simultanément au brin inférieur de la région riche en AT formant un complexe ternaire (ADNdb-protéines-ADNsb). La région riche en AT d'*ori2* contient six répétitions directes régulièrement espacées de tétra-nucléotides 5'-ATCA reconnues par RctB (Chatterjee *et al.*, 2020). La formation de ce complexe ternaire nécessite la fixation de la protéine IHF en aval de série de 6x12mer, ce qui plie l'ADN à 180°. Cette structure conduit à l'ouverture de l'ADN au niveau des deux sites ATCA les plus proches de la boîte IHF (Figure 29) (Chatterjee *et al.*, 2020).

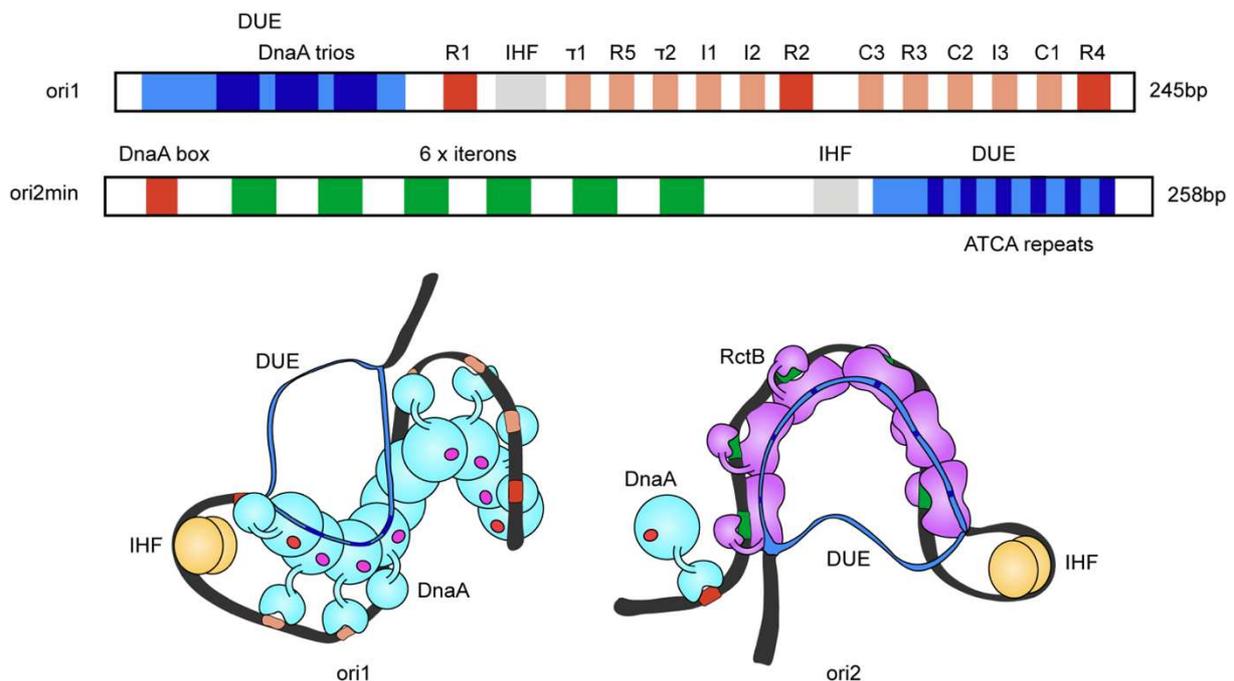


Figure 29. Modèle d'ouverture de l'origine de réplication dans Chr1 et Chr2. L'origine de réplication de Chr1 est une séquence longue de 245 pb contenant de nombreux sites de liaison DnaA. Il existe trois types de sites de liaison DnaA : -les boîtes DnaA fortes (R1, R2, R4) -les boîtes DnaA faibles (T1, R5, T2, I1, I2, C3, R3, C2, I3, C1) et les trios DnaA (13mer) à l'intérieur de l'élément de déroulement de l'ADN (DUE). L'origine de réplication du Chr2 (*ori2*) est longue de 258 pb et contient une boîte DnaA forte, six répétitions régulièrement espacées de 12 pb (itérons), une boîte IHF et six répétitions ATCA tétranucléotidiques à l'intérieur du DUE. La DnaA est représentée en bleu clair, l'association de la DnaA avec l'ATP est obligatoire pour les sites de liaison faible (cercles roses) ; pour les sites de liaison forte, la DnaA est associée à l'ATP ou à l'ADP (cercles rouges). RctB est en violet, IHF en jaune, DUE en bleu. Les domaines d'oligomérisation de DnaA (domaine I) et de RctB (domaine IV) ne sont pas représentés.

Bien que la liaison de DnaA soit nécessaire pour l'initiation de Chr2 (de Lemos Martins *et al.*, 2018), elle n'est pas impliquée dans le moment de l'initiation (Duigou *et al.*, 2006). On retrouve des boîtes DnaA dans la plupart des origines de plasmides à itérons (Wegrzyn *et al.*, 2016). Les deux principales hypothèses du rôle de DnaA sont : -la stabilisation de l'ouverture

de la région riche en AT de l'origine -l'aide au chargement de l'hélicase DnaB. Les mêmes hypothèses s'appliquent pour *ori2*.

Globalement, il apparaît que pour l'ouverture de la DUE, RctB suit un mécanisme analogue à l'ouverture de l'*oriC* par DnaA chez *E.coli* (Ozaki and Katayama, 2012), (Sakiyama et al., 2017). Dans ce cas, les protéines DnaA_{Ec} se lient à plusieurs boîtes DnaA et interagissent ensuite avec l'ADN simple brin via un motif H/B (résidus V211 et R245) contenu dans le domaine III (Duderstadt et al., 2011), (Ozaki et al., 2008). Il pourrait être intéressant de savoir quelle région de la protéine RctB permet l'interaction avec l'ADN simple brin ; des études complémentaires sont nécessaires.

En conclusion, l'ouverture de la région riche en AT d'*ori2* permet le recrutement de deux réplisomes et conduit au démarrage bidirectionnel de la réplication.

5. Contrôle par la méthylation de la réplication de Chr2

La plupart des plasmides à itérons ne contiennent pas d'enrichissement particulier en séquence GATC dans leur origine, à l'exception du plasmide P1. Pour le plasmide P1, la méthylation par Dam est essentielle car elle facilite le déroulement de l'ADN dans la DUE (Abeles and Austin, 1987), (Brendler *et al.*, 1991). Dans le plasmide P1, les sites GATC à l'extérieur de l'itéron, près de la boîte DnaA, sont séquestrés par SeqA lorsqu'ils sont hémiméthylés afin d'empêcher de nouveaux cycles de réplication (Brendler et al., 1995).

Par conséquent, Chr2 diffère des plasmides à itérons car il contient une surreprésentation de sites de méthylation GATC. Ceux-ci restent hémiméthylés pendant une période de temps significative (Demarre and Chatteraj, 2010). La perte de SeqA entraîne une dérégulation du *timing* de réplication pour le Chr2 suggérant que les deux origines sont sujettes à une séquestration post-reproduction par SeqA similaire à ce qui a été décrit chez *E. coli* (Lobner-Olesen et al., 2005) (Figure 30).

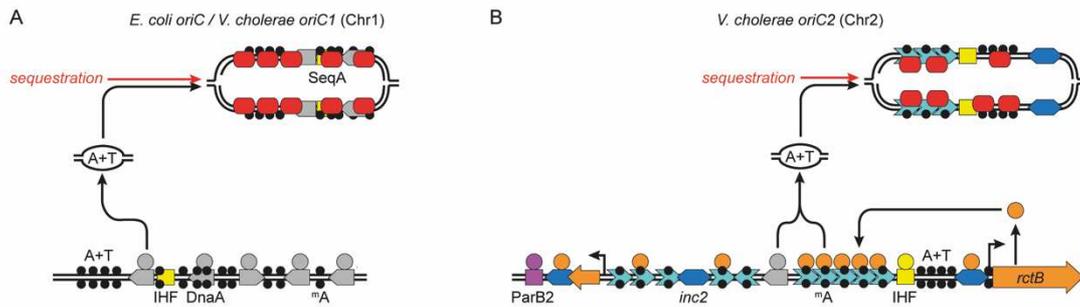


Figure 30. Séquestration des *ori1* et *ori2* par SeqA. Les deux origines *ori1* et *ori2* contiennent des sites GATC qui seront reconnus et séquestrés par SeqA post réplication. Adapté de (Espinosa et al., 2017)

De plus, chaque itéron à l'intérieur d'*ori2* contient un motif de méthylation GATC (Egan and Waldor, 2003). Dam joue une fonction essentielle supplémentaire dans le contrôle de la réplication de chr2 en ce sens que RctB se lie uniquement aux itérons entièrement méthylés par Dam (Demarre and Chatteraj, 2010). En l'absence de méthylation, RctB ne se lie pas aux itérons et l'initiation de chr2 est bloquée, entraînant la mort cellulaire (Val et al., 2012). La protéine Dam est donc essentielle chez *V.cholerae* alors qu'elle ne l'est pas chez d'autres organismes comme *E.coli*. A noter que Dam n'est plus essentielle dans une souche où les deux chromosomes sont fusionnés car la réplication du Chr2 dépend alors de l'origine du Chr1 (Val et al., 2014).

Juste après le recrutement des deux réplisomes, les origines sont répliquées. Dans certains plasmides à itérons, les origines interagissent ensemble via les protéines Rep lorsque le nombre de plasmides augmente. Ce mécanisme inhibe la réplication excessive (Chatteraj, 2000). Cependant, ce n'est pas le cas pour Chr2. En effet, RctB se lie uniquement aux itérons complètement méthylés (Venkova-Canova and Chatteraj, 2011). Ainsi, le menottage intermoléculaire entre deux array d'itérons (6x12mer) n'est pas possible *in vivo* même s'il a été observé *in vitro* (Venkova-Canova and Chatteraj, 2011). De plus, les deux *ori2* répliquées sont ségréguées très rapidement du fait d'un système de partition ParAB2 et ne sont donc plus à proximité (Val et al., 2016).

Pour conclure, la méthylation a un double rôle pour *ori2* : la séquestration par SeqA et le recrutement de l'initiateur RctB (Demarre and Chatteraj, 2010).

6. Régulation négative de l'origine du chromosome 2

Le Chr2 de *V.cholerae* a vraisemblablement évolué à partir d'un plasmide itéron dont la régulation de la réplication n'était pas aussi complexe. Pour mieux comprendre ce qui se passe au niveau d'*ori2*, il est intéressant de dresser un parallèle avec la régulation négative de l'origine des plasmides à itérons.

Tout d'abord, dans des plasmides tels que pPS10, F, pSC101, R6K et RK2, la protéine Rep peut se lier à un opérateur au sein de son propre promoteur. Dans ce cas, Rep se lie en tant que dimère à des itérons répétés inversement, inhibant la transcription (Komori *et al.*, 1999), (Ingmer *et al.*, 1995), (Filutowicz *et al.*, 1985).

De plus, les itérons peuvent être situés à différents endroits sur l'origine, leur localisation change radicalement leur fonction. Lorsqu'ils se trouvent près de la région riche en AT, ils ont un effet positif sur l'initiation en permettant l'ouverture de l'ADN. Cependant, les itérons peuvent également se retrouver dans la région d'incompatibilité (*inc*) comme dans le cas du plasmide F (Zzaman and Bastia, 2005). Les protéines RepE attachées à ces itérons vont interagir de manière intramoléculaire via le repliement de l'ADN avec la RepE attachée au réseau d'itérons près de la région riche en AT. Cela conduit à l'inhibition de l'initiation de la réplication et ce mécanisme est appelé "handcuffing".

Globalement, l'interaction de Rep avec un itéron peut avoir trois rôles distincts : ouverture de l'ADN, auto-répression de la transcription, *handcuffing* de l'origine inter/intra.

Dans le cas de l'origine du chromosome 2, l'initiateur RctB se lie à deux types de sites : aux itérons (12pb) ainsi qu'à un second type de site différent à la fois par sa séquence et sa longueur : les 39m. Il y a trois sites 39m sur *ori2*, deux 39mer dans la région d'incompatibilité et un 29mer dans le promoteur de RctB.

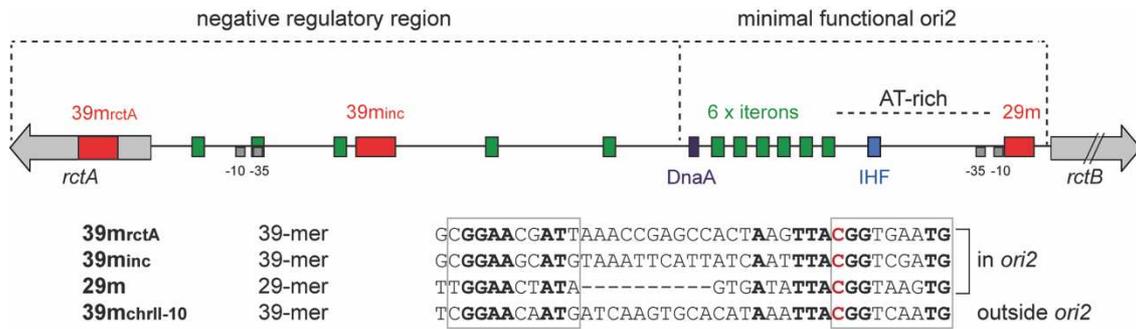


Figure 31. Organisation de l'origine de réplication du chromosome 2 de *Vibrio cholerae*.

Les 39m contiennent des répétitions directes riches en GC à leurs deux extrémités avec une séquence intermédiaire riche en AT (Figure 31). Les répétitions GC sont conservées dans tous les membres séquencés de la famille *Vibrio*. Ces répétitions imparfaites riches en GC sont des sites de liaison de RctB en phase du même côté de l'hélice d'ADN. Le 29mer contient une délétion de 10nt dans sa partie centrale riche en AT, cette délétion correspond à un tour complet de l'hélice d'ADN et ne perturbe pas la phase des deux parties riches en GC du site (Venkova-Canova and Chatteraj, 2011). RctB semble se lier au site 39m en tant que dimère formant une structure dyade, mais les preuves expérimentales directes manquent (pas de structure sur l'ADN).

Les 39m sont conservés chez tous les *Vibrionaceae*, on ne retrouve pas d'équivalent dans les plasmides à itérons. Ils permettent un contrôle étroit de l'initiation et sont considérés comme le régulateur négatif principal de l'*ori2*. Ils inhibent fortement l'initiation d'*ori2* par le biais d'un handcuffing intramoléculaire entre la région d'incompatibilité (*inc*) et *ori2*_{min} (Venkova-Canova and Chatteraj, 2011), (Venkova-Canova et al., 2012). Les détails structurels de ce complexe d'inhibition sont encore inconnus mais le modèle proposé est que les molécules de RctB liées aux itérons et celles liées aux 39m forment un pont qui empêche l'initiation de la réplication (Jha et al., 2012), (Venkova-Canova and Chatteraj, 2011), (Venkova-Canova et al., 2006).

Pour le 29mer, le rôle est double car il participe au handcuffing intramoléculaire mais il se trouve également dans le promoteur de *rctB* chevauchant le TSS de l'ARNm (position -4:+25). Le site 29mer permet une boucle de rétroaction négative de l'expression de *rctB*. D'un point de vue évolutif, le site 29mer peut être comparé aux sites opérateurs des plasmides à itérons qui sont liés par la forme dimérique de Rep.

On peut se demander pourquoi *ori2* a acquis un nouveau type de site (les 39m) alors que les autres plasmides à itérons ont déjà une réplication très stricte. Par exemple, le plasmide F est maintenu de façon très stable dans la cellule avec un nombre de copies de 1 alors qu'il ne contient qu'un seul type de site de liaison pour Rep. Le 39m constitue peut-être l'innovation fonctionnelle qui a permis l'apparition de coordination de la réplication entre le Chr1 et le Chr2.

Notons que la région d'incompatibilité de l'origine du Chr2 contient également cinq itérons régulateurs. Lorsque ces itérons sont inactivés par mutation, le nombre de copies d'un plasmide *pORI2* diminue, ces itérons ont donc un rôle positif dans l'initiation de Chr2 (Figure 32) (Venkova-Canova and Chattoraj, 2011). Ce résultat est à l'opposé de ce qui se passe pour le plasmide F, pour lequel les itérons de la région d'incompatibilité inhibent l'initiation de la réplication. Il semble que le rôle principal de ces itérons régulateurs serait de contrer l'action des 39m par un mécanisme qui n'est pas encore compris.

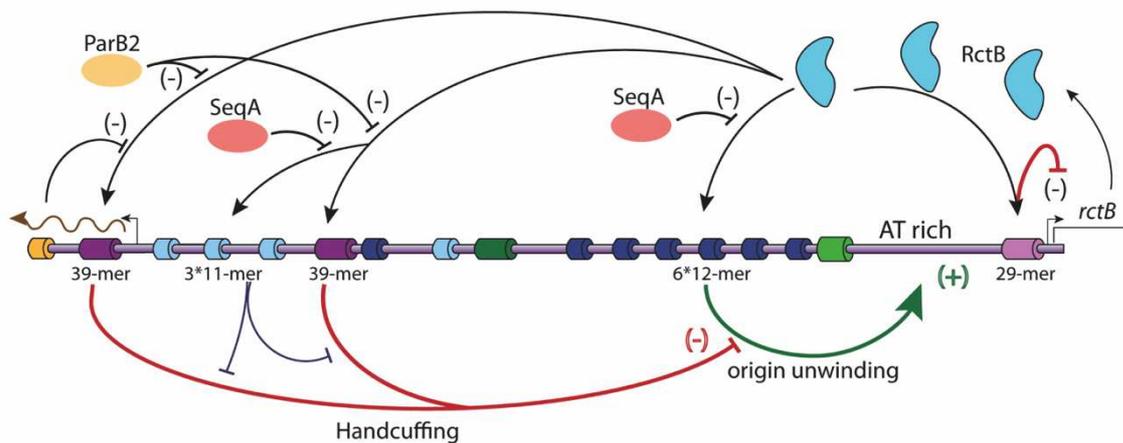


Figure 32. Schéma de la régulation de l'initiation de la réplication du Chr2 de *Vibrio cholerae*. Différents mécanismes régulent l'initiation de la réplication à l'*ori2*. L'initiateur RctB (en bleu clair) se fixe à deux types de sites : les sites inhibiteurs (39mer et 29mer) et les itérons (11mer et 12mer). Lorsque RctB se fixe à la série de 6x12mer (bleu foncé) permet l'ouverture de l'ADN sur la région riche en AT. RctB se fixe sur le site 29mer dans son propre promoteur inhibant sa propre expression. Les sites 39mer sont les principaux inhibiteurs de l'ouverture de l'ADN (violet), ils inhibent l'initiation par un mécanisme nommé handcuffing. Les sites 11mer de la région *inc* (bleu clair), contre-carre l'action des 39mer. SeqA et ParB2 empêchent la fixation de RctB après la réplication (Fournes et al., 2018).

7. Coordination de la terminaison de la réplication du Chr1 et du Chr2

En 2007, Rasmussen et al. ont proposé deux mécanismes qui pourraient permettre une terminaison synchrone de la réplication pour les deux chromosomes chez *V.cholerae* (Rasmussen *et al.*, 2007). La première est que l'initiation des deux chromosomes pourrait être liée à la masse cellulaire, mais avec une masse d'initiation plus importante pour le Chr2. La seconde pourrait être la présence d'un *timer* sur le Chr1 qui, lorsqu'il est répliqué, envoie un signal au Chr2.

En 2014, Baek et al ont découvert un site de fixation de RctB présent à environ 2/3 du réplichore droit du Chr1 qui régule positivement l'initiation du Chr2 (Baek and Chattoraj, 2014). Pour cette découverte, ils se sont inspirés de travaux alors récemment publiés sur DnaA, pour laquelle des sites régulateurs importants avaient été trouvés en dehors de l'origine de réplication *oriC* d'*E.coli*, notamment *datA* qui favorise l'hydrolyse de la DnaA-ATP (Kasho and Katayama, 2013) ainsi que DARS1 et DARS2 permettant la régénération de la DnaA-ADP en DnaA-ATP (Fujimitsu *et al.*, 2009). Suivant cette idée, Baek et al. ont réalisé une expérience de CHIP-on-chip avec l'initiateur du Chr2, RctB, afin de trouver d'autres sites de liaison en dehors d'*ori2*. Ils ont alors identifié un site de liaison de RctB dans une région intergénique située sur le bras droit du Chr1 à 690kb en aval d'*ori1*. Cette région ne contient aucun site connu ni itéron et ni 39m. Ce site a un effet positif sur la réplication d'*ori2*, en effet lorsqu'il est mis en *trans* sur un plasmide il augmente la réplication d'un *pORI2* chez *E.coli* (Baek and Chattoraj, 2014).

L'équipe de D.Mazel a démontré que ce site agissait comme un *timer*, car sa réplication permet de déclencher l'initiation du Chr2 (Val *et al.*, 2016). Le site a donc été nommé *crtS* pour Chromosome two Replication Triggering Site. Par la suite, des expériences de délétion successives ont prouvé que la taille minimale de ce site est de 62pb (de Lemos Martins *et al.*, 2018). Bien qu'il ne soit pas essentiel sa délétion conduit à un phénotype de filamentation pour la cellule, sans ce site la réplication de Chr2 est fortement perturbée (Figure 33). Les souches Δ *crtS* gagnent rapidement des mutations compensatoires : soit une fusion des chromosomes, soit une mutation dans le gène/promoteur de RctB (Val *et al.*, 2016).

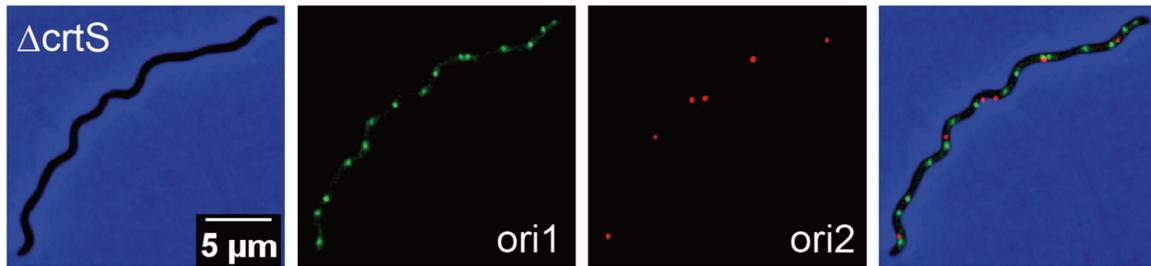


Figure 33. La délétion de *crtS* entraîne une sous-initiation du Chr2. Une souche $\Delta crtS$ a un phénotype de filamentation caractéristique d'une inhibition de la réplication de l'ADN. Dans cette souche le Chr1 (vert) continue à répliquer et le Chr2 est sous répliqué (rouge). A gauche, observation en contraste de phase. L'*ori1* est tagguée par GFP et *ori2* tagguée par RFP. A droite, les trois images sont superposées. (Val *et al.*, 2016)

Enfin, une approche par génomique comparative chez les *Vibrionaceae* prouve que la distance entre *crtS* et Ter1 est en corrélation avec la longueur d'un réplichone du Chr2. Cela signifie que chez les espèces de *Vibrio* ayant un Chr2 plus grand, le locus *crtS* est plus proche de *ori1* (Kemter *et al.*, 2018).

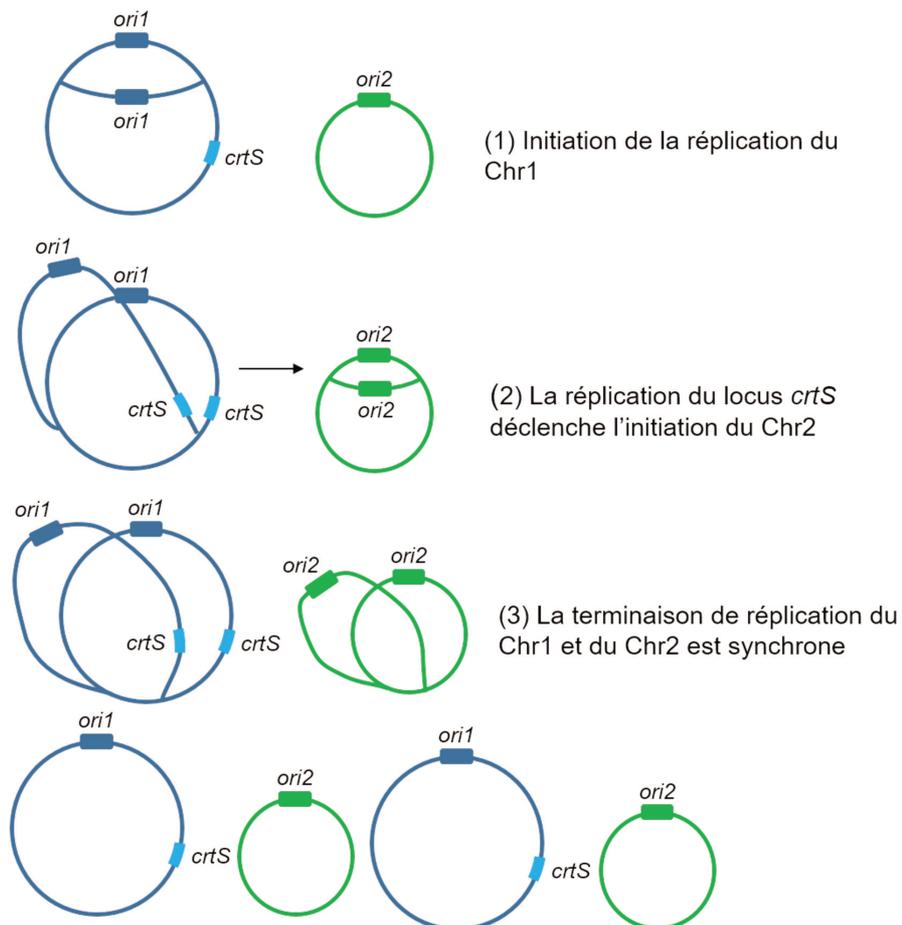


Figure 34. Le site *crtS* permet la synchronisation de la réplication chez *V.cholerae*. Le chromosome 1 (bleu foncé) et le chromosome 2 (en vert) sont répliqué de manière coordonnée. L'*ori1* est initiée en premier et la fourche de réplication se déplace bidirectionnellement le long du Chr1. Lorsque la fourche passe au locus *crtS* (bleu clair), un signal (flèche noire) est délivré et déclenche l'initiation de la réplication du Chr2. (Adapté de Fournes et al, 2018)

Le site *crtS* contrôle non seulement le moment de l'initiation mais aussi le nombre de copies du Chr2, lorsqu'on ajoute des sites *crtS* au génome de Vibrio, le nombre de copies du Chr2 augmente progressivement (de Lemos Martins *et al.*, 2018) (Figure 34).

8. *crtS* et *ori2* : à proximité dans la cellule

Une information cruciale a été obtenue en étudiant une souche avec deux copies chromosomiques de *crtS* et en particulier le mutant *crtSWT/VC23* qui a une copie ectopique de *crtS* près de l'*ori1* ainsi que le *crtS* au locus natif. Les cellules avec une *ori2* tagguée ont été observées par microscopie à fluorescence. Dans cette population, les cellules filles naissent avec deux *ori2*. Elles passent ensuite par un stade trois *ori2* puis quatre *ori2* avant la division cellulaire (Figure 35) (Val *et al.*, 2016). Par conséquent, dans cette souche, il y a une activation séquentielle des origines suggérant que la duplication d'un site *crtS* active une seule *ori2*. Ces observations semblent indiquer une action locale de *crtS*.

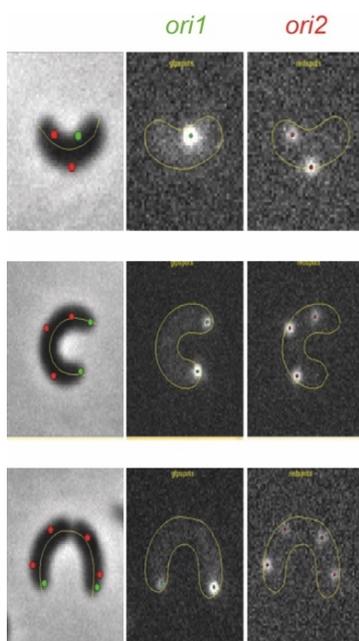


Figure 35. Activation séquentielle de l'*ori2* dans une souche avec deux *crtS*. Microscopie à fluorescence d'une souche de *V.cholerae* contenant deux copies chromosomiques de *crtS* (une proche de l'*ori1* et une en position native). L'*ori1* est tagguée en vert l'*ori2* en rouge

Ainsi, l'activation des initiateurs dans les systèmes RctB-*ori2* et DnaA-oriC semble de nature complètement différente. En effet, Le système DnaA-oriC est basé sur une activation globale pendant une fenêtre de temps réduite du lot de protéines DnaA libres dans la cellule. Lorsqu'un plasmide multicopie mini-oriC est inséré dans une souche d'*E.coli*, tous les oriC sont activés simultanément (Skarstad and Lobner-Olesen, 2003).

Un autre argument en faveur de cette localité de l'action de *crtS* a été obtenu à partir d'une expérience 3C (chromosomal conformation capture) réalisée sur une population de cellules en croissance exponentielle. Avec cette expérience on observe que l'*ori2* est en contact préférentiel avec le réplicore droit du Chr1 autour du locus *crtS* (Figure 36).

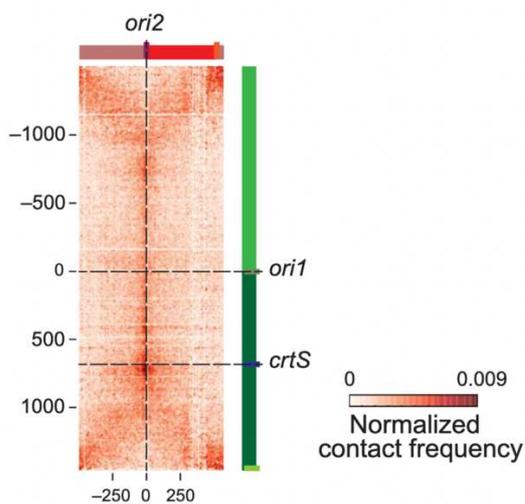


Figure 36. Les régions *ori2* et *crtS* sont en contact dans la cellule. Données de contact entre les deux chromosomes de *V.cholerae* obtenues par 3C.

Enfin, la colocalisation de *crtS* et *ori2* dans la cellule a également été confirmée par microscopie à fluorescence (Val *et al.*, 2016). L'ensemble de ces résultats suggère une interaction physique entre le site *crtS* et *ori2* dans la cellule.

9. Implication du régulateur global Lrp dans l'activité de *crtS*

En utilisant un screen génétique, le régulateur global Lrp (Leucine Responsive regulatory Protein) a été identifié comme se liant au site de *crtS* (Ciaccia *et al.*, 2018). En effet, *crtS* contient dans sa partie 5' la séquence TAGATTTTTATTCTT très proche (12/15 nucléotides) du consensus du site de liaison de Lrp (le consensus étant YAGHAWATTWTDCTR, (Cui *et al.*, 1995)) (Figure 37). Dans la souche Δ lrp, le site *crtS* ne fonctionne pas et Chr2 est sous-répliqué. In vitro, Lrp facilite la liaison de RctB à *crtS*. Lrp est

connu pour enrouler l'ADN (Beloin et al., 2003), ce changement de conformation local de l'ADN permet probablement à RctB de reconnaître le site *crtS*.

```
crtS TTGTTTTTAGATTTTATTCTTTTAAAAATTGAAATGCAGAATATGTAACCTTATGCTTTCGG 62bp  
                Lrp binding site                Lrp binding site
```

Figure 37. *crtS* 62bp minimal sequence contains two predicted Lrp binding sites.

Sous cet aspect, *crtS* est analogue au fonctionnement du site DARS-2 de *E.coli* (DnaA Reactivation Site), qui nécessite la liaison de deux NAPS pour régénérer la DnaA-ATP : les masters régulateurs Fis et IHF (Kasho *et al.*, 2014). De plus, il a été récemment démontré que la liaison de Fis à DARS2 fluctue au cours du cycle cellulaire, ce qui est essentiel pour coordonner l'initiation de la réplication (Miyoshi *et al.*, 2021). Peut-être que tout comme Fis est la clé de la fonction du cycle cellulaire de DARS-2, Lrp peut être la clé de la détection du croisement de la fourche de réplication à *crtS*. Il a déjà été montré que Lrp permet également dans certains cas de détecter le croisement des fourches, la variation de phase de l'opéron *pap* mais ce mécanisme est dépendant de la méthylation (récemment revu dans : (Zamora et al., 2020)). D'autres études sur la régulation de *crtS* par Lrp sont nécessaires.

10. RctB : une protéine Rep avec deux domaines supplémentaires

L'initiateur RctB est une protéine de 658 acides aminés unique au genre *Vibrionaceae*, sa séquence est très conservée. Sa structure complète n'est pas élucidée, nous savons cependant qu'elle comporte quatre domaines structuraux (Figure 38). La structure des domaines centraux II et III est homologue à celle d'initiateurs de plasmides à itérons de type Rep. Les domaines I, II et III possèdent chacun un motif Hélice-Tour-Hélice (HTH) permettant la fixation à l'ADN (Orlova *et al.*, 2017).

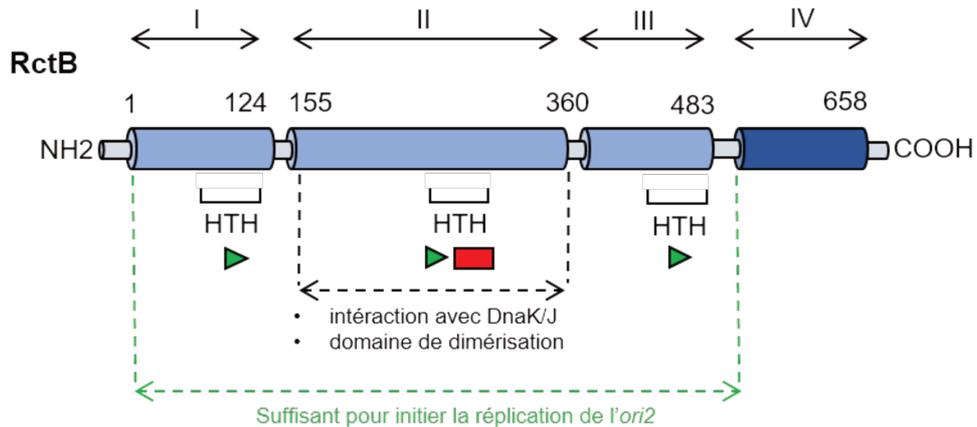


Figure 38. Domaines structuraux de la protéine RctB : initiateur de la réplication du chromosome 2. RctB comporte quatre domaines structuraux. Les domaines I, II et III portent un motif HTH, les trois premiers sont nécessaires pour se fixer aux itérons (triangle vert), HTH-2 est le seul motif HTH à être nécessaire pour la fixation aux 29/39mer (rectangle rouge). Les 493 premiers acides aminés sont suffisants pour l'initiation de l'ori2 (en bleu clair). Le domaine IV serait dédié à la régulation (en bleu foncé). (Adapté de (Fournes *et al.*, 2018)).

RctB est deux fois plus grande que les initiateurs Rep et contient deux domaines supplémentaires (domaines I et IV). Premièrement, le domaine I contient un motif de liaison à l'ADN wHTH impliqué dans la liaison des itérons. Le domaine I est essentiel pour l'initiation de la réplication. Dans un modèle proposé par Orlova et al. le domaine 1 et le domaine 2-3 se lient aux faces opposées de l'ADN de l'itéron. Avec le domaine 1 et le domaine 2 se lient directement au 12mer, et le domaine 3 se lient aux séquences adjacentes de manière non spécifique (Orlova *et al.*, 2017) (Figure 39).

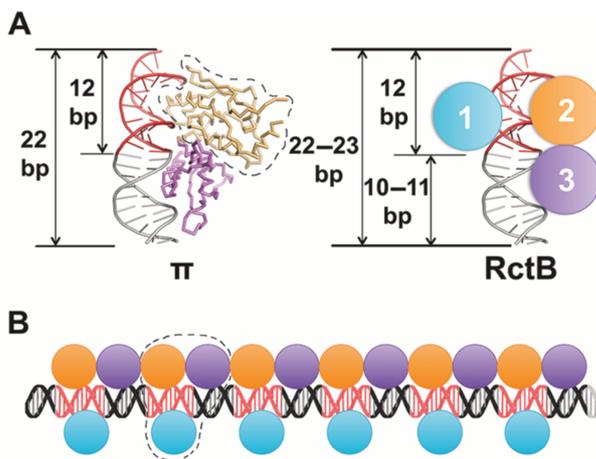


Figure 39. Modèle de l'interaction de RctB avec les itérons de l'ori2.

Le gain d'un domaine N-terminal de liaison à l'ADN a déjà été observé dans les initiateurs de plasmides. Pour le plasmide RK2, l'initiateur de réplication trfA possède un domaine de liaison à l'ADN (DBD) N-terminal supplémentaire. Comme le domaine 1 de RctB, ce domaine DBD

est essentiel pour l'interaction avec les itérons. La modélisation suggère d'ailleurs que chaque domaine interagit avec une face différente de l'ADN (Wegrzyn et al., 2021).

Les 493 premiers acides aminés de RctB (domaine I, II, III + 10 acides aminés), sont suffisants pour initier la réplication du Chr2 (Orlova *et al.*, 2017). Contrairement au domaine additionnel I, le domaine IV n'est pas essentiel à l'initiation mais est impliqué dans la régulation négative du nombre de copies (Jha *et al.*, 2012), (Jha et al., 2014), (Fournes et al., 2021). A ce jour, il n'y a pas de structure publiée du domaine IV. Cependant, d'après les modélisations structurales, il apparaît qu'il ne ressemble à aucune structure connue. Le domaine IV est au moins impliqué dans deux mécanismes : le premier est le *handcuffing*, c'est un mécanisme régulant négativement la réplication via des interactions protéine-protéine entre deux RctB fixées sur l'origine (Venkova-Canova and Chattoraj, 2011). Le second est l'oligomérisation, lorsque RctB se fixe sur l'ADN à un site spécifique, il entraîne la fixation aspécifique de plusieurs autres monomères de RctB via une interaction protéine-protéine (de Lemos Martins *et al.*, 2018). Enfin, il a été montré que le domaine IV serait impliqué dans la fixation aux sites inhibiteurs 29/39mer (Jha *et al.*, 2014).

11. Similarité structurale entre Rep et RctB

Les initiateurs de type Rep sont de petites protéines d'environ 300 acides aminés contenant généralement deux domaines (WH1 et WH2). Chaque domaine Rep contient un motif wHTH qui sert de médiateur pour la liaison à l'ADN. Au contraire, RctB est deux fois plus grande que les initiateurs Rep avec 658 résidus et est divisée en quatre domaines structurels nommés domaine I, II, III et IV (Orlova *et al.*, 2017).

Le domaine II (aa. 182:360) et le domaine III (aa. 361-472) de RctB ont été cristallisés ensemble formant un segment stable de ~38 kDa contenant les résidus 155-483 (Orlova *et al.*, 2017). Ces deux domaines centraux sont structurellement homologues aux domaines WH1 et WH2 des initiateurs de type Rep (pi, RepE, RepA, 1Rep).

En solution, Rep s'organise en dimère, l'interaction se produit entre chaque domaine WH1 dans un arrangement tête-à-tête. De manière similaire les domaine II-III de RctB cristallisent sous forme de dimère en tête-à-tête, l'interface de dimérisation est contenue principalement dans le domaine II (Figure 40). Les domaines II et III contiennent chacun un

motif de liaison à l'ADN de type HTH (Helix-Turn-Helix), essentiel pour la liaison à l'ADN sur les sites itérons.

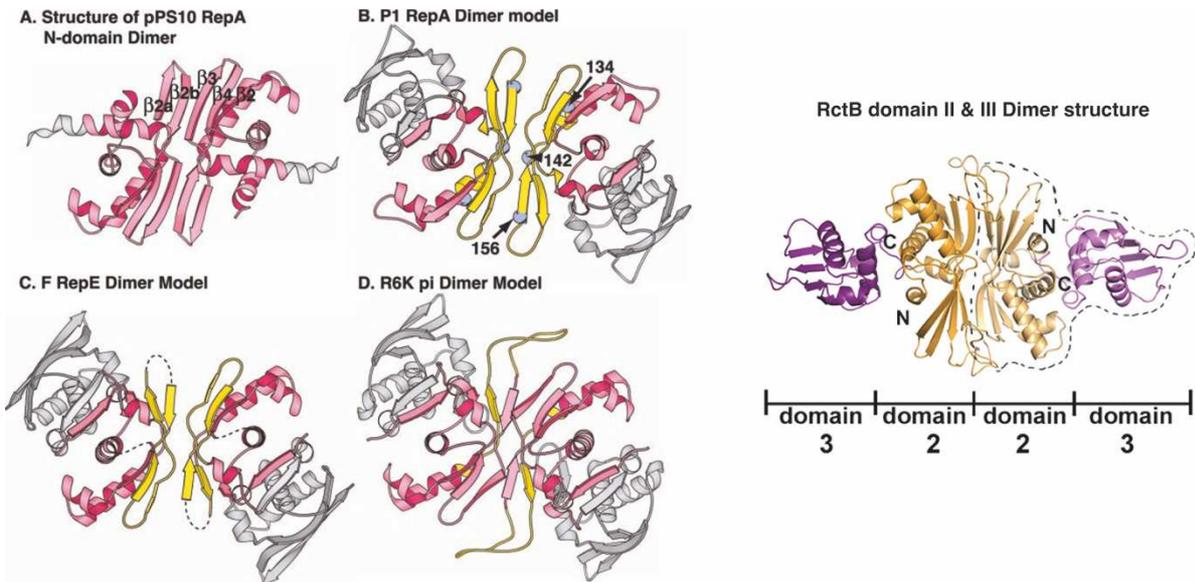


Figure 40. RctB adopte une structure dimérique via son domaine II ressemblant aux dimères de Rep. Les protéines de type Rep (RepA, RepA_{P1}, RepE, pi) forme un dimère en interagissant avec leur domaine WH1 (Sharma et al., 2004)

L'organisation dimérique tête à tête de RctB n'est pas compatible avec la liaison de l'ADN sur les itérons. Cela signifie que RctB a besoin d'un remodelage par monomérisation pour se lier correctement aux itérons. De manière similaire aux initiateurs de type Rep, le dimère subit une activation par monomérisation assistée par les protéines chaperonnes DnaK/J. RctB possède un site d'interaction avec DnaK dans le domaine II, qui s'étend des résidus 150 à 163. De façon surprenante, ce site d'interaction, appelé site KI, ne ressemble pas au KI de Rep (Jha et al., 2017).

12. Comparaison entre la régulation de la réplication du Chr2 et des plasmides à itérons

La réplication du chromosome bactérien commence lorsqu'une certaine masse cellulaire est atteinte, ce n'est pas le cas pour la réplication des plasmides. La réplication des plasmides peut se produire à tout moment du cycle cellulaire (Ramachandran et al., 2017). Bien que le Chr2 partage beaucoup de points communs avec les plasmides itérons, il initie toujours la réplication à un point fixe du cycle cellulaire. Ainsi, il y a eu un changement au cours de l'histoire de l'évolution d'*ori2* d'un mode de réplication aléatoire à un mode régulé par le cycle cellulaire.

- Ouverture de l'origine

L'organisation de l'origine de réplication de chr2 présente de nombreuses similitudes avec celle de la famille des plasmides itéron : RK2, F, R6K, P1, pPS10 ou pSC101 (Kim et al., 2020). Dans cette famille, l'initiation se produit lorsque la protéine Rep codée par le plasmide se lie à des séquences d'ADN répétées et orientées, appelées itérons. Cet événement de liaison introduit des forces de torsion facilitant la fusion de l'hélice d'ADN double brin (Konieczny *et al.*, 2014). Ces sites itérons sont au moins au nombre de trois dans le cas d'oriPSC101 et leur taille varie de 17bp pour oriRK2 à 22bp pour oriR6K. Le Chr2 code sa propre protéine initiateur nommée RctB à côté de son origine de réplication : *ori2*. RctB se lie à un réseau de six itérons régulièrement espacés. Les itérons d'*ori2* sont plus courts que ceux des plasmides (12 pb). De plus, il existe d'autres éléments communs comme la présence d'une région riche en AT qui sera ouverte lors du processus d'initiation. Comme le plasmide F, *ori2* contient également une région d'incompatibilité qui régule l'initiation. Il existe également des sites de liaison pour d'autres protéines essentielles à la réplication : la boîte DnaA et la boîte IHF (Konieczny *et al.*, 2014).

- Régulation négative

Les plasmides itérons sont présents en un faible nombre de copies par cellule, leur maintien nécessite donc une régulation très stricte de la réplication. Cette régulation se fait principalement par deux mécanismes qui sont l'auto-répression de l'initiateur et le handcuffing (Chattoraj, 2000). En effet, la protéine Rep peut se lier à son propre promoteur pour en limiter l'expression. Le handcuffing est un mécanisme de pontage de l'ADN médié par Rep à l'origine des plasmides et qui empêche l'initiation par encombrement stérique. Il existe deux types de handcuffing : intermoléculaire et intramoléculaire. Le premier se produit entre différentes molécules d'ADN lorsque le nombre de copies du plasmide augmente, provoquant un appariement des origines. Le second se produit entre la région d'incompatibilité et la région minimale de l'origine. Dans le cas de l'origine du Chr2, on retrouve également ces deux mécanismes. L'initiateur est capable de se lier à son promoteur et d'inhiber sa transcription. Le handcuffing de l'*ori2* est intramoléculaire, se faisant entre la région *inc* et les itérons de la région minimale comme c'est le cas pour le plasmide F (Uga et al., 1999), (Zzaman and Bastia, 2005).

13. Chargement du réplisome, élongation et terminaison de la réplication des deux chromosomes

Pour *ori1*, il a été récemment démontré que l'hexamère libre de DnaB de *Vibrio cholerae* contrairement à celui d'*E.coli* adopte une conformation partiellement ouverte et est capable de se charger sur de l'ADN simple brin. DciA, bien que non essentiel à l'initiation, interagit avec DnaB dans une stoechiométrie 3:6, facilitant son chargement sur l'ADN ainsi que son activité hélicase (Marsin et al., 2021), (Cargemel et al., 2022).

Pour l'instant, le recrutement de l'hélicase au niveau d'*ori2* n'a pas été étudié. Dans les plasmides itéron, plusieurs scénarios ont été observés (Wegrzyn *et al.*, 2016). Dans le cas de l'origine pSC101, RepA, DnaA et DnaC participent tous au chargement de DnaB (Datta et al., 1999). Alors que dans le cas du plasmide RK2 de *Pseudomonas putida*, l'initiateur *trfA-44* peut directement recruter DnaB sans l'aide de DnaA ni de DnaC (Jiang et al., 2003), (Caspi et al., 2001). Des études supplémentaires sont nécessaires pour déterminer comment le recrutement des hélicases se produit après l'ouverture d'*ori2* par RctB.

Le réplisome bactérien est hautement conservé, il coordonne le déroulement de l'ADN et la réplication des deux brins leading et lagging. Il est composé de nombreuses protéines : Le complexe polymérase PolIII, la primase DnaG, le chargeur de clamp, la protéine SSB et l'hélicase DnaB (Lewis *et al.*, 2016), (Xu and Dixon, 2018).

Après l'initiation et le chargement des réplisomes, les chromosomes sont répliqués de manière bidirectionnelle. Les deux fourches de réplication progressent depuis l'origine de réplication jusqu'à la région Ter pour les deux chromosomes. La vitesse de réplication est constant le long du génome et identique pour Chr1 et Chr2 (Val *et al.*, 2016). Cela est attendu car sur le Chr2 il n'y a pas de gènes supplémentaires impliqués dans le réplisome, ni polymérase ni hélicase, la réplication dépend donc entièrement de la machinerie encodée sur le chromosome principal. Pendant la croissance exponentielle, un contact en forme de croix est observé dans les données de HiC entre le tiers inférieur des réplichores Chr1 et les réplichores Chr2 (Val *et al.*, 2016). Ces contacts sont perdus dans la phase non-réplivative, ce qui suggère que la réplication est responsable de ces contacts le long de leurs bras chromosomiques.

Espinosa et al. ont développé une nouvelle méthode pour capturer la cohésion derrière la fourche de réplication nommée HiSC2 pour High-Resolution Whole-Genome Analysis of

Sister Chromatid Contacts. Après le passage de la fourche de réplication, les deux chromatides sœurs restent en contact préférentiel dans certaines zones. Il existe cinq régions majeures de cohésion le long du génome : *ori1*, *ori2*, *dif1*, *dif2* et autour de l'île de pathogénicité VPI-1. La cohésion n'est pas liée aux différences de vitesse de la fourche de réplication, ni aux processus d'initiation et de terminaison. On peut supposer que la cohésion est causée par des facteurs inconnus spécifiques à la séquence, sauf pour VPI1. Dans ce cas, la cohésion est due à la protéine H-NS qui étend la cohésion à l'intérieur de l'îlot de pathogénicité VPI-1 et, dans une moindre mesure, sur le locus O-antigène. Ensuite, à la fin du cycle cellulaire, la plupart de la cohésion entre les chromatides sœurs sont libérées avant la division cellulaire, à l'exception de *dif1* (Espinosa et al., 2020).

Lorsque les deux fourches de réplication ont complété chaque réplichore, elles finissent par se rencontrer dans la région Ter du chromosome. *Vibrio cholerae*, bien qu'étroitement apparenté à *E. coli*, n'héberge aucun homologue de Tus ou RTP. Il a été montré que chez *Vibrio*, si l'*ori* est déplacé, ou si un *ori* ectopique est inséré dans le chromosome, le point de convergence de la fourche est déplacé. Ce point de convergence n'est donc pas dépendant d'un système de piège actif. La réplication se termine spontanément lorsque les deux fourches de réplication se rencontrent et fusionnent (Galli et al., 2019).

QUESTIONS ET HYPOTHESES

Questions et hypothèses du projet de thèse

Concernant le chromosome d'*E.coli*, nous commençons à avoir une vision complète de comment fonctionne l'initiation de sa réplication. Le mécanisme central est que l'initiateur DnaA fonctionne grâce à son domaine AAA+ comme un interrupteur, le lot cellulaire de DnaA peut osciller entre un état inactif DnaA-ADP et un état actif DnaA-ATP permettant des cycles d'initiation successifs particulièrement robustes.

Cependant, dans le cas de RctB l'initiateur du Chr2 il n'y a pas de domaine AAA+, son fonctionnement apparait donc complètement différent de DnaA-oriC. RctB ressemble à un initiateur de plasmide et pourtant contrairement aux plasmides l'initiation de la réplication du Chr2 est coordonnée avec le cycle cellulaire.

L'étude de la réplication du second chromosome de *V.cholerae* commence en 2003 dans l'équipe de M.K. Waldor. Dans cette étude, il est démontré que le Chr1 suit une initiation classique ressemblant à DnaA-oriC chez *E.coli* alors que l'origine du Chr2 ressemble aux origines des plasmides à itérons et encode son propre initiateur. Au cours des dernières années, de nombreuses avancées ont été réalisées dans la compréhension de la régulation de l'*ori2* grâce aux efforts de cinq équipes de recherche à travers le monde (les laboratoires de M.K. Waldor, D.K. Chattoraj, A. Lobner-Olesen, T. Waldminghaus, et le nôtre).

En 2014, dans le laboratoire de D.K. Chattoraj, est découvert un site de fixation de RctB trans-activateur de l'*ori2* présent sur le Chr1. En 2016, nous avons démontré dans notre laboratoire que ce site avait un rôle de *timer* permettant la terminaison synchrone deux chromosomes. Le mécanisme d'activation de l'initiation de la réplication d'un chromosome par un site présent sur le chromosome principal apparait comme complètement inédit et pourrait nous en apprendre beaucoup sur le processus de domestication des mégaplasmides.

Comment la réplication du site *crtS* permet de déclencher l'initiation de la réplication du Chr2 chez *Vibrio cholerae* ?

Au cours de ma thèse, j'ai étudié l'action de la réplication *crtS* sur l'*ori2* et comment elle contribue à l'initiation de la réplication du Chromosome 2. Lorsque j'ai commencé cette étude, nous possédions alors déjà un certain nombre d'informations. Nous savions que l'*ori2* est activée après la réplication du locus *crtS* sur le Chr1, site de fixation de RctB. Nous savions également que les deux loci *ori2* et *crtS* semblent être en contact dans la cellule. Nous

pensions alors déjà que l'action de *crtS* sur l'*ori2* est probablement directe et localisée, et qu'il n'y a pas de forme active de RctB libre dans la cellule qui oscille.

Il restait de nombreuses questions sans réponse concernant la régulation de l'origine de la réplication du Chr2. Nous ignorions notamment la structure du domaine IV de RctB et son rôle précis dans la régulation négative de l'origine. Nous ne comprenions pas non plus quels changements se produisaient au niveau de l'*ori2* après la réplication de *crtS* et comment *crtS* percevait le passage de la fourche de réplication. Enfin, le régulateur global Lrp semblait important pour la régulation de l'*ori2* par *crtS*, mais son rôle précis *in vivo* restait encore flou.

RESULTATS

Résultats

I. Premier article: “The coordinated replication of *Vibrio cholerae*’s two chromosomes required the acquisition of a unique domain by the RctB initiator”

Florian Fournes*, [Theophile Niault*](#), Jakub Czarnecki, Alvisé Tissier-Visconti, Didier Mazel, Marie-Eve Val. *: co-premier auteur

Cet article a été publié dans le journal *Nucleic Acids Research* le 8 novembre 2021.

Contribution personnelle : partie *in vivo*, constructions des souches et plasmides, dPCR.

Le mécanisme par lequel la réplication du site *crtS* déclenche l’initiation du Chromosome 2 de *Vibrio cholerae* reste encore mal compris. Dans cette étude, nous fournissons de nouvelles informations sur la façon dont l’initiation de la réplication de Chr2 est régulée par *crtS* via RctB.

Le point de départ de cette étude est l’isolement au laboratoire d’un mutant compensatoire d’une souche Δ *crtS*, cette mutation se trouve dans le site 29mer présent dans le promoteur de *rctB* à la position 20 (29mC₂₀>A). Cette cytidine est conservée dans tous les sites 39m et serait par conséquent importante pour la reconnaissance du site par RctB. Nous avons décidé d’étudier l’impact de cette mutation sur la réplication du Chr2, nous savions que son effet pouvait être double car le 29mer agit à la fois sur la transcription du gène *rctB* et sur le *handcuffing* de l’origine.

Nous avons démontré que la mutation 29mC₂₀>A empêche la fixation de RctB *in vitro* et *in vivo*. Cette mutation cause une augmentation de l’expression de RctB et augmente le nombre de copies du Chr2. Lorsqu’on mute cette même cytidine dans les autres 39m (C₃₀>A) le nombre de copies du Chr2 augmente également sans pour autant changer l’expression de *rctB*.

Lorsqu’un seul des sites 39m est muté (39mer_{rctA} ou 39mer_{inc} ou 29mer), le nombre de copies d’un *pORI2* devient insensible à la présence de *crtS*. Suggérant que *crtS* agit sur les 39mer. Nous proposons un modèle dans lequel *crtS* (sur Chr1) agit comme un site anti-inhibiteur en empêchant les sites 39m (sur Chr2) de réprimer l’initiation.

Pour confirmer ce rôle d'anti-inhibiteur, nous avons également étudié un mutant du domaine IV RctB qui ne se fixe plus aux 39m. Avec cette mutation le nombre de copies du p*ORI2* devient insensible à *crtS*.

Nous montrons également que la queue C-terminale de RctB, essentielle pour l'auto-interaction de RctB, est cruciale pour le contrôle exercé par *crtS*. Cette sous-région de RctB est conservée chez tous les *Vibrio*, mais absente chez les autres initiateurs de type Rep. Par conséquent, la réplication coordonnée des deux chromosomes résulte probablement de l'acquisition de ce domaine unique par RctB.

The coordinated replication of *Vibrio cholerae*'s two chromosomes required the acquisition of a unique domain by the RctB initiator

Florian Fournes^{1,2,†}, Theophile Niaux^{1,2,3,†}, Jakub Czarnecki^{1,2,4}, Alvisé Tissier-Visconti^{1,2}, Didier Mazel^{1,2,*} and Marie-Eve Val^{1,2,*}

¹Institut Pasteur, Université de Paris, Unité Plasticité du Génome Bactérien, Département Génomes et Génétique, Paris 75015, France, ²Centre National de la Recherche Scientifique, UMR3525, Paris 75015, France, ³Sorbonne Université, Collège Doctoral, Paris 75005, France and ⁴University of Warsaw, Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, Warsaw 02-096, Poland

Received April 21, 2021; Revised August 30, 2021; Editorial Decision September 16, 2021; Accepted September 22, 2021

ABSTRACT

Vibrio cholerae, the pathogenic bacterium that causes cholera, has two chromosomes (Chr1, Chr2) that replicate in a well-orchestrated sequence. Chr2 initiation is triggered only after the replication of the *crtS* site on Chr1. The initiator of Chr2 replication, RctB, displays activities corresponding with its different binding sites: initiator at the iteron sites, repressor at the 39m sites, and trigger at the *crtS* site. The mechanism by which RctB relays the signal to initiate Chr2 replication from *crtS* is not well-understood. In this study, we provide new insights into how Chr2 replication initiation is regulated by *crtS* via RctB. We show that *crtS* (on Chr1) acts as an anti-inhibitory site by preventing 39m sites (on Chr2) from repressing initiation. The competition between these two sites for RctB binding is explained by the fact that RctB interacts with *crtS* and 39m via the same DNA-binding surface. We further show that the extreme C-terminal tail of RctB, essential for RctB self-interaction, is crucial for the control exerted by *crtS*. This subregion of RctB is conserved in all *Vibrio*, but absent in other Rep-like initiators. Hence, the coordinated replication of both chromosomes likely results from the acquisition of this unique domain by RctB.

INTRODUCTION

Genome architecture is defined in most bacterial lineages, conserved through speciation, and therefore is a fundamental question of microbial evolution. 5–10% of bacte-

ria have a multipartite organization of their genome with at least one secondary chromosome containing essential core genes (1). Secondary chromosomes originate from plasmids that have been domesticated as *bona fide* components of the genome, as evidenced by several of their plasmid features (2). Although plasmids confer adaptive functions, which augment fitness under certain conditions, they are also frequently poorly adapted to a new genetic background and can be quickly lost in the absence of selective pressure. Therefore, their domestication as secondary chromosomes requires co-evolutionary processes involving their adaptation to host cellular mechanisms as well as accommodation of the host (3). In this study, we made an in-depth structure-function study of RctB, the replication initiator of *Vibrio cholerae* secondary chromosome, to understand the mechanism underlying the orchestrated replication of both chromosomes, which is controlled by this protein.

Like all *Vibrionaceae*, *V. cholerae* has its genome split on two circular replicons, Chr1, a main 3 Mbp-chromosome and Chr2, a secondary 1 Mbp-chromosome. The Chr1 replication origin (*ori1*) is similar to the single chromosome origin of *Escherichia coli* (*oriC*) and the ubiquitous bacterial replication initiator DnaA governs Chr1 initiation (4–6). The initiation of Chr2 replication, however, is different from that of a typical bacterial chromosome and its origin (*ori2*) is closely related to that of iteron-containing plasmids (6). Like iteron-plasmid origins, *ori2* replication initiation requires an accurate arrangement of DNA motifs comprising iterons, an IHF binding site, an AT-rich DNA unwinding element and a DnaA-box (7) (Figure 1A). Iterons are directly repeated sequences recognized by a plasmid-encoded initiator (Rep) (8). The binding of Rep to iterons near the AT-rich element promotes local DNA unwinding where replication starts (8). The specific equivalent of Rep for *Vibrio*

*To whom correspondence should be addressed. Tel: +33 186467361; Email: marie-eve.kennedy-val@pasteur.fr
Correspondence may also be addressed to Didier Mazel. Email: didier.mazel@pasteur.fr

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

© The Author(s) 2021. 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-NonCommercial 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

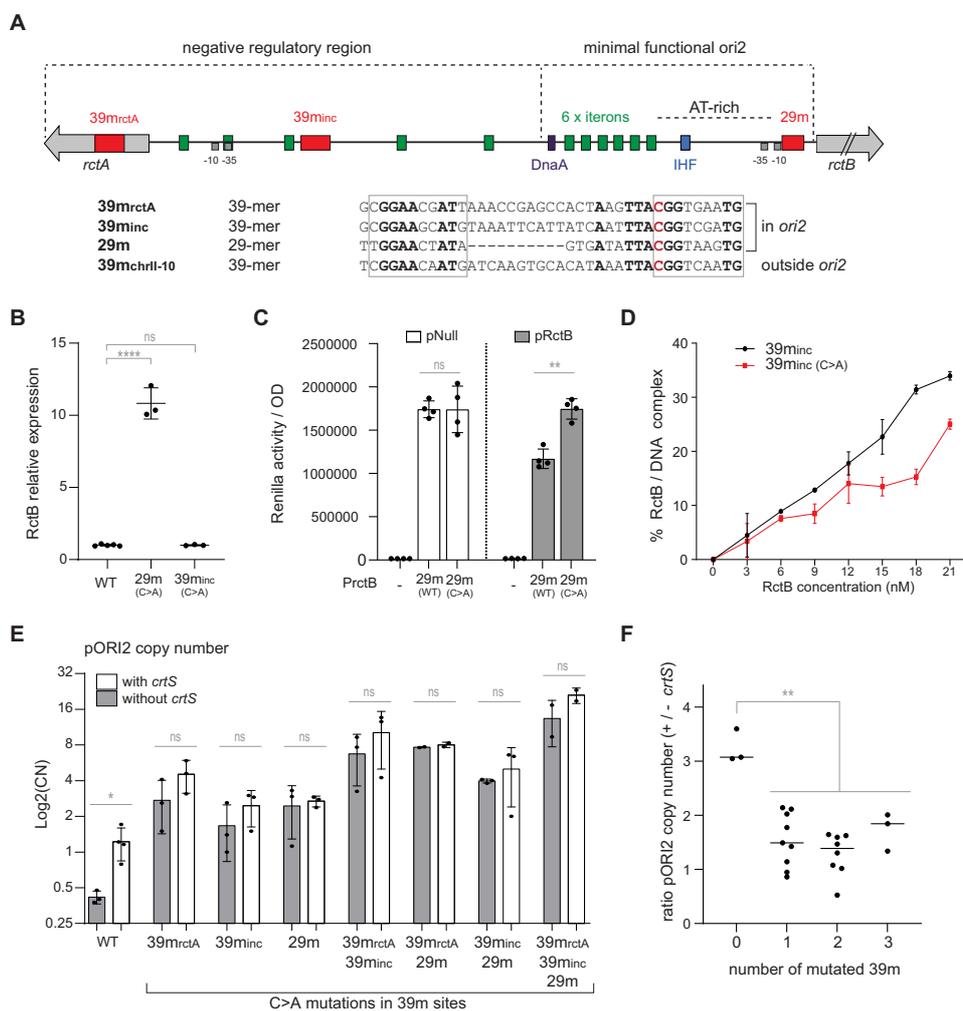


Figure 1. *crtS* counteracts 39m inhibitory activity. (A) Representation of Chr2 origin of replication (*ori2*) and sequence alignment of all 39m sites. RctB binding sites are indicated in green (iterons) and red (39m). All 39m contain two GC-rich repeats at both ends with an intervening 19-bp (or 9-bp) non-conserved AT-rich sequence. The imperfect GC-rich repeats are RctB binding sites (gray box), which are phased to be on the same side of the DNA helix (13). Deletion of 10 bp in the 29m intervening sequence (deletion of a full helix turn) maintains the phasing between the RctB binding sites and does not abrogate the function of the site. Conserved bases are shown in bold. The (C > A) mutation is indicated in red (18). (B) Expression of *rctB* relative to the housekeeping gene *gyrA* in the N16961 wild-type strain (WT), the 29m (C > A) and the 39m_{inc} (C > A) *V. cholerae* mutants. (C) Measure of repression by RctB of its promoter with and without the (C > A) mutation: 29m(WT), 29m(C > A). Promoter activity is assessed by measuring Renilla luciferase activity (Experimental set-up in Supplementary Figure S1). RctB expression is induced with 0.2% arabinose from a low-copy plasmid (pTVC11, Supplementary Table S2). The same empty plasmid is used as a control of absence of repression (pNull). Data are normalized to the optical density of each culture (OD at λ 600 nm). (D) Quantification of complex formation between a 40 bp DNA probe carrying a 39m_{inc} (with or without a C > A mutation) and increasing concentrations of RctB (nM : nanomolar), without DnaK/J_{Vc} (Supplementary Figure S2). Percent of bound DNA is equal to (intensity of the shifted band/combed intensities of shifted and unbound DNA bands) x 100. (E) pORI2 copy number (CN) relative to the chromosome [Log₂(pORI2/*oriC*)] in *E. coli* strains with and without a chromosomal *crtS* site. pORI2 contains either 1, 2 or 3 mutated 39m sites (C > A). (F) Graph representing the effect of *crtS* (pORI2 *E. coli* / pORI2 *E. coli::crtS*) as a function of the number of mutated 39m sites (experimental values of (E)). In all graphs, data represent averages from at least three independent biological replicates \pm standard deviations. Each dot represents an independent experiment. Statistics were applied on pairwise comparisons using Student's t-test and on multiple comparisons using one-way ANOVA. Statistical significances are indicated (ns : non significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$).

Chr2 is called RctB (6). Similar to Rep, RctB proteins form homo-dimers in solution with a head-to-head configuration (9). As Rep, RctB dimers are monomerized by DnaJ and DnaK chaperones, which favours their binding to DNA in a cooperative manner (10–12). Iteron-type plasmids limit their copy number via a mechanism called origin-pairing or handcuffing. In this process, the coupling of origins via iteron-bound Rep protein causes steric hindrance to initiation (8). The handcuffing of *ori2* is also one of the main regulatory mechanisms that limit Chr2 copy number; however, it requires the participation of additional regulatory sequences other than iterons, which are discussed below (13).

Despite many features common to both Chr2 and iteron plasmids, the regulation of Chr2 replication involves additional functions that are not found in plasmids and which lead to a more elaborate control of Chr2 replication (13). First, Chr2 replication initiation is tightly regulated to occur only once per cell cycle, while plasmids generally initiate replication several times over the bacterial cell cycle (14,15). Secondly, Chr2 replication is initiated at a specific time of the cell cycle so that Chr1 and Chr2 complete their replication at the same time, an evolutionary selected characteristic in *Vibrio* (16–18). It thus seems that during their domestication, plasmid-type machineries, which are suitable for maintaining high-copy-number plasmids, have evolved towards systems capable of sustaining the replication of low-copy-number replicons in harmony with the host cell cycle (1,14). Indeed, the structure of RctB is more complex than those of Rep. RctB is a large protein (658 amino acids (AA)), twice as large as the Rep initiators (e.g. π is 305 AA, TrfA is 382 AA). RctB harbors four domains (I, II, III, IV) with three DNA binding surfaces (in domains I, II and III) (9) whereas most Rep initiators contain only two domains (N-terminal and C-terminal) with two DNA binding surfaces that bind to the two halves of the iterons (9). RctB central domains (domains II and III) are structurally related to Rep N-terminal and C-terminal domains and are essential for Chr2 replication. RctB additional domains (domains I and IV) are unique and may have emerged during the evolution of *Vibrionaceae* to help meet the stricter regulatory requirements necessary for proper chromosome maintenance (9). Domain I is essential for binding to iterons and for replication (9). Domain IV, which is not essential for replication, is however critical for limiting the number of Chr2 copies (11).

Unlike Rep, RctB specifically recognizes DNA motifs other than iterons. RctB binds to three classes of DNA sites of different sequences and sizes: the iterons (11-mer or 12-mer), the 39m sites (29-mer, 39-mer) and the *crtS* site (62-bp) (10,18) (Figure 1A). Eleven iterons are present in *ori2*, six of which are essential for replication, while the others have a regulatory function (6). Four 39m (three in *ori2* and one located 40 kb away from *ori2*) act as strong negative regulators of *ori2* initiation (13,19,20) (Figure 1A). The unique *crtS* site located on Chr1 exerts a positive regulation of Chr2 replication (20). At these three types of sites, RctB monomers specifically bind cooperatively and then oligomerize non-specifically on adjacent DNA (10). The initial binding of RctB as a monomer requires domains I, II and III (10,11). However, various DNA binding surfaces are required for binding to iterons and 39m suggesting that the

binding conformation of RctB is different to these sites (9). Both the cooperative binding and oligomerization activities of RctB require its domain IV (10,11). As mentioned above, Chr2 copy number is limited by a handcuffing mechanism which is conveyed by the 39m sites that are contiguous to the iterons in *ori2* (13). In this process, iteron-bound RctB and 39m-bound RctB form a bridge that hinders the initiation of replication (11,13,21). This negative regulation of *ori2* initiation mediated by the 39m requires RctB domain IV. A large deletion of RctB domain IV ($\Delta 157$) causes loss of handcuffing activity and a copy-up phenotype (11). More specifically, a 71 AA region in domain III-IV (AA 450 to 521) has been extensively characterized and found to be involved in Chr2 copy number limitation, DNA binding and RctB dimerization (22).

The above-mentioned knowledge of Chr2 replication regulation mechanisms is insufficient to explain how the initiation of Chr2 replication is synchronized with Chr1 replication. This coordination requires the third RctB binding site, *crtS* (18), located in trans, halfway between the origin and the terminus of Chr1 (20). Upon replication, *crtS* triggers the initiation of Chr2, therefore its location on Chr1 ensures termination synchrony of the two chromosomes (18). The deletion of *crtS* severely impairs growth and is associated with the loss of Chr2 (18). It was proposed that *crtS* could act as a DNA chaperone to remodel RctB to an active form with altered DNA binding activities to iteron and 39m sites (20,23). In this study, we dissected RctB to better understand how *crtS* steers Chr2 replication. We found that *crtS* triggers Chr2 initiation by acting as an anti-inhibitory site, preventing the 39m sites from negatively regulating replication. This *crtS* activity involves the same RctB DNA binding region that is required for the binding to the 39m. The action of *crtS* also involves a protein-protein interaction domain located at the extreme carboxy-terminal end of RctB. We propose that the cooperative binding of RctB to DNA, mediated by protein-protein interactions, is necessary for *crtS* activity as an anti-inhibitor of the 39m.

MATERIALS AND METHODS

Bacterial strains, plasmids and nucleotides

All *E. coli* and *V. cholerae* strains and plasmids used in this study are listed in Supplementary Tables S1 and S2. Further details regarding their engineering are provided in the Supplementary Methods. The oligonucleotide sequences used for dPCR, RT-dPCR, *in vitro* experiments and cloning purposes are listed in Supplementary Table S3.

Protein purification

The expression plasmids derived from pET32b were transformed into *E. coli* BL21-DE3 (Supplementary Table S2). Plasmids used to express the RctB HTH mutants were built as described in (9). Proteins were produced and purified as described in (10). More details are provided in Supplementary Methods.

In vitro experiments

Electrophoretic Mobility Shift Assays (EMSA) of double-stranded DNA probes containing the RctB *crtS*, 39m or

iteron binding sites were radiolabeled in 5'-end [γ 32P]. A 114-bp long probe containing *crtS* was generated by PCR as described in (10). The iteron-containing probes were methylated *in vitro* as described in (10). The resulting labeled probes were used in a binding reaction performed as described in (10) and analyzed with the Typhoon FLA 9500 laser scanner. All protein concentrations are indicated in the figure legends. More details are provided in Supplementary Methods.

Bacterial Two-hybrid

BTH101 cells were co-transformed with plasmids encoding a T18 N-fusion of RctB and a T25 N-fusion of RctB according to the instructions described in (24). The different plasmid combinations are described in Tables S2. Co-transformants were selected on LB agar plates containing carbenicillin (100 μ g/ml) and kanamycin (25 μ g/ml). Co-transformants were pooled, resuspended in 100 μ l of LB medium, and spotted (5 μ l drop of cell suspension) onto LB agar plates containing X-Gal (40 μ g/ml), IPTG (0.5 mM), carbenicillin (100 μ g/ml), and kanamycin (25 μ g/ml). Plates were incubated for 2 days at 30°C, followed by an additional 24 h at 4°C. Blue spots indicate protein-protein interactions.

Digital PCR quantification of DNA loci (dPCR)

Quantifications of (*ori1* and *ori2*) in *V. cholerae* and (*oriC* and *pORI2*) in *E. coli* were performed as described in (10) using multiplex digital PCR (Stilla Technologies) (25). Primers and probes are listed in Supplementary Table S3. Total genomic DNA was purified using the DNeasy Blood & Tissue Kit (QIAGEN). PCR reactions were performed with 0.1 ng of DNA using the PerfeCTa[®] MultiPlex qPCR ToughMix[®] (Quantabio) on a Sapphire chip (Stilla Technologies). Digital PCR was conducted on a Naica Geode and Image acquisition on the Naica Prism3 reader. Images were analyzed Crystal Miner software (Stilla Technologies). The dPCR run was performed using the following steps: droplet partition (40°C, atmospheric pressure AP to +950mbar, 12 min), initial denaturation (95°C, +950 mbar, for 2 min), followed by 45 cycles at (95°C for 10 s and 60°C for 30 s), droplet release (down 25°C, down to AP, 33 min). More details are provided in Supplementary Methods.

Digital PCR quantification of RNA (RT-dPCR)

Quantification of RctB mRNA was performed in *V. cholerae* from exponentially growing cultures (OD₆₀₀ 0.4) using multiplex digital RT-dPCR (Stilla Technologies) (25). Primers and probes are listed in Supplementary Table S3. *V. cholerae* RNA was prepared as described in (26). PCR reactions were performed with 1 ng of RNA using the qScript[™] XLT One-Step RT-qPCR ToughMix[®] (Quantabio). RT-dPCR and data analysis were conducted as described above for dPCR. The RT-dPCR run was performed in the following steps: droplet partition (40°C, AP to +950 mbar, 12 min), cDNA synthesis (50°C, +950 mbar, 10 min), initial denaturation (95°C, +950 mbar, for 1 min), followed by 45

cycles at (95°C for 10 s and 60°C for 15 s), droplet release (down 25°C, down to AP, 33 min). Expression values were normalized to the expression of the housekeeping gene *gyrA* as described in (26). More details are provided in Supplementary Methods.

Luciferase assay

E. coli reporter strains were engineered with a *lacZ* insertion containing the *renilla* luciferase gene (27) under the control of PrctA or PrctB promoters. More details for generating reporter strains are provided in the Supplementary Methods. Reporter strains were transformed with the plasmid pTVC11 containing a transcriptional fusion of *rctB-firefly* luciferase genes under the P_{BAD} promoter (28). Overnight cultures from transformed cells were diluted 1:100 in Mueller-Hinton medium supplemented with 50 μ g/ml spectinomycin and 0.2% arabinose and grown to an OD₆₀₀ ~0.5. To 50 μ l of reporter cell culture, we added 40 μ l of *E. coli* cell culture (J665, OD₆₀₀ 0.5) and 10 μ l of a KH₂HPO₄ (1M)-EDTA (20 mM) solution. The mixture was placed on dry ice for 5 min and left to equilibrate at room temperature for 15 min. The cells were then lysed with 300 μ l of a lysis buffer for 10 min. Both luciferase activities of Firefly and Renilla were sequentially measured using the dual-luciferase reporter assay (DLR, Promega) following the manufacturer's recommendations. 20 μ l of the lysed extract was added to 100 μ l of Firefly substrate (LARII) in a glass tube. The light emission was acquired for 10 s in a luminometer (Lumat LB9507, Berthold). Quenching of Firefly luciferase luminescence and concomitant activation of Renilla luciferase was achieved by adding 100 μ l of STOP&GLO Reagent to the tube. Light emission from Renilla activity was acquired for 10 s. Luminescence data were normalized to the OD₆₀₀. MG1655 wild-type cultures were used as non-luminescent controls.

RESULTS

crtS exerts a positive control on Chr2 replication by counteracting the inhibitory activity of 39m

We previously isolated mutations in *ori2* that compensate for the loss of *crtS* in *V. cholerae* (18). In particular, we identified a mutation in the 29-mer site (29m) located in the promoter region of the *rctB* gene. This mutation is an adenine substitution of a cytosine (C > A), which is conserved at all 39m sites (Figure 1A), indicating the importance of this nucleotide. In addition to its role in hindering *ori2* initiation via handcuffing, the binding of RctB to the 29m represses its own transcription (19). We tested the impact of the C > A mutation on RctB expression in *V. cholerae* by RT-dPCR. We observed that RctB is 10 times more expressed when the 29m sequence contains the C > A mutation than in the WT strain or in a mutant where the C > A mutation is inserted in a 39m site outside the RctB promoter region (39m_{inc}) (Figure 1B). Overexpression of RctB could thus be due to a lack of self-repression due to a decrease in RctB binding to the 29m in its promoter region. Since RctB binding to the 29m represses the activity of its own promoter (19), we used this repression as a

proxy to monitor RctB binding *in vivo*. We developed a luciferase reporter tool in *E. coli* by fusing the promoter region of RctB (PrctB) to a Renilla luciferase reporter gene. RctB fused to a Firefly luciferase gene is supplied on a low-copy plasmid (PrctB) under the control of an arabinose-inducible promoter (P_{BAD}) (Supplementary Figure S1A). Using a dual-luciferase assay system, we can both monitor RctB binding to 39m by measuring PrctB activity (via Renilla-generated luminescence measurement) while controlling the induction of RctB expression on the plasmid (via Firefly-generated luminescence measurement). As expected, upon RctB induction (Supplementary Figure S1B, pink bars), PrctB activity decreased significantly (Supplementary Figure S1B, blue bars). In absence of RctB (Figure 1C, pNull), the C > A mutation in the 29m site in the promoter region of RctB (PrctB-29m_(C>A)) had no impact on Renilla expression compared to the WT promoter (PrctB-29m_(WT)) meaning that the C > A mutation in the 29m does not interfere with the activity of the promoter. In presence of RctB (Figure 1C, pRctB), the level of expression of Renilla decreased significantly in PrctB-29m_(WT) but remained unchanged with PrctB-29m_(C>A). Therefore, the C > A mutation alters the repressive activity of RctB by altering its binding to the 29m. We further assessed whether the impact of the C > A mutation on RctB binding was also detrimental at other 39m sites. We chose to test the effect of this mutation on the 39m_{inc} site, because RctB binds to this site similarly as to the 29m site (19). *In vitro*, we observed a decrease in RctB/DNA complex formation when the 39m_{inc} site is mutated, which corroborates our *in vivo* observation on the 29m (Figure 1D, Supplementary Figure S2). EMSA showed that both DNA probes containing either wild-type or mutated 39m sites mainly formed C2 complexes with RctB (i.e. the 39m is bound by two molecules of RctB, (10)), suggesting that the C > A mutation does not affect the cooperative binding of RctB to 39m (Supplementary Figure S2). We further studied the impact of mutating all three 39m sites in *ori2* on the control of replication by *crtS*. We used an artificial replication system in *E. coli*, using a plasmid, pORI2, which replication depends solely on *ori2* (10). By quantitative digital PCR (dPCR), we monitored pORI2 copy number (CN) variation in *E. coli* MG1655 without and with a *crtS* chromosomal site. In *E. coli*, the presence of *crtS* up-regulates pORI2 copy number (10). We then introduced the C > A mutations at the three 39m sites of the pORI2, either individually, in pairs or all three together, and measured the relative copy number of pORI2 in the absence or presence of *crtS*. As expected, since the 39m are strong inhibitory sites, their incremental mutations resulted in an increase in the copy number of pORI2 in absence of *crtS* (Figure 1E – without *crtS*, Supplementary Figure S3A). Furthermore, we revealed that *crtS* no longer has any effect on the control of pORI2 copy number when the 39m sites carry the C > A mutation, regardless of which site has been mutated (Figure 1E – with *crtS*, Supplementary Figure S3B). Surprisingly, the control by *crtS* was lost once a single site was mutated (Figure 1F). Since each of the C > A mutations renders the 39m sequence insensitive to *crtS* activity, it is likely that *crtS* inhibits the handcuffing activity of the 39m and not only the transcriptional control role of RctB through its binding to the 29m. All these results strongly suggest that *crtS* up-

regulates the initiation of replication at *ori2*, primarily by neutralizing the inhibitory effect of 39m.

The binding of RctB to *crtS* and 39m involves the same DNA binding surface

To understand how *crtS* competes with the 39m sites, we investigated which RctB DNA binding surfaces are involved in its interaction with *crtS*. RctB harbours three winged-helix-turn-helix motifs (HTH1, HTH2, HTH3) in domains I, II and III respectively, that mediate binding to DNA (9). As no full length RctB structure is currently available, we used a recently released AlphaFold2-MMseqs2 webserver to generate a model of RctB complete structure (29,30) which we visualized with ChimeraX (31) (Figure 2A, Video S1). This predicted structure shows that the three HTH motifs are close together in a pocket that could accommodate a DNA molecule. This structure is in agreement with Orlova's model (9) which speculates that domain I and domain II–III bind opposite faces of the iteron DNA site. They showed that all three HTH are crucial for RctB binding to iterons, but only HTH2 is essential for the binding to 39m (9). To explore the importance of these HTHs in *crtS* binding, we purified the three RctB HTH mutants (RctB_{HTH1}, RctB_{HTH2}, RctB_{HTH3}) and confirmed that their affinities for 39m were as described in (9) (Supplementary Table S2, Supplementary Figure S4). We further compared the binding of RctB_{HTH1}, RctB_{HTH2}, RctB_{HTH3} to the binding of RctB_{WT} by EMSA with a radiolabelled DNA probe containing the *crtS* site (Figure 2B). With all four RctBs tested (WT and HTH mutants), we observe the formation of higher molecular weight complexes. This observation depends on the ability of several RctB monomers to bind cooperatively to DNA (10). Thus, all three RctB_{HTH} mutants retain their ability to bind cooperatively to DNA. Quantification of the formation of RctB-*crtS* complexes at increasing RctB concentrations reveals that RctB, RctB_{HTH1} and RctB_{HTH3} have the same affinity for *crtS* while RctB_{HTH2} hardly forms complex with *crtS* (Figure 2C). At the highest concentration of the RctB_{HTH2} mutant, we could observe the formation of a higher molecular weight nucleoprotein complex, meaning that RctB_{HTH2} is still able to bind cooperatively to *crtS* (10). Thus, the disruption of HTH2 strongly reduced the binding efficiency of RctB to the *crtS* site as it is the case for the 39m without abolishing its ability to bind cooperatively. This suggests that the binding conformation of RctB, involving domain II, could be similar for 39m and *crtS*. Consistent with these results, the addition of non-radioactive *crtS* DNA competitor reduces the formation of RctB/39m complexes *in vitro* while the interactions between iterons and RctB are not altered by the addition of *crtS* (Supplementary Figure S5). Since 39m and *crtS* compete for RctB binding, we wondered if *crtS* could partially be substituted by a 39m in *V. cholerae* on Chr1. To do so, we first inserted the three types of RctB binding sites (iteron, 39m and *crtS*) into the Tn7 insertion site (*attTn7*) located on the Chr1 of *V. cholerae*, and then excised the native *crtS* site (Δ *crtS*). Using digital PCR (dPCR), we measured the copy number variation between Chr1 and Chr2 of non-replicating *V. cholerae* bearing either of the RctB sites on Chr1 (Figure 2D). If the RctB binding site is fully

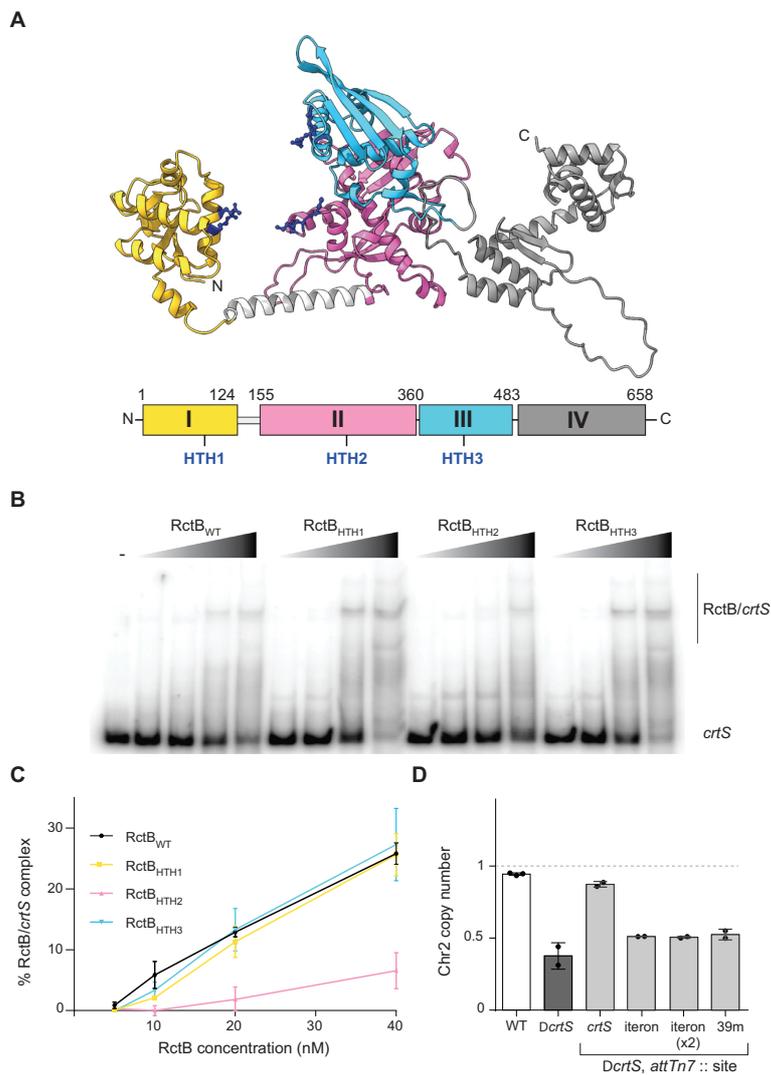


Figure 2. RctB share the same DNA binding surface to bind to *crtS* and 39m. (A) Model of RctB structure predicted with AlphaFold2 (29) and visualized with ChimeraX (31). The four domains of RctB, described by Orlova (9), are shown with the same color code in the secondary structure model and in the linear representation of RctB below. The three winged-helix-turn-helix motifs in domains I, II and III are indicated HTH1, HTH2 and HTH3 on the linear representation of RctB. The AAs that we mutated into these HTH motifs are shown in dark blue on the secondary structure model. (B) Electrophoretic mobility shift assays of wild-type RctB (RctB_{WT}) and HTH mutants: RctB_{HTH1} (Q83A-R84A-R86A), RctB_{HTH2} (K271A-K272A-S274A) and RctB_{HTH3} (R420A-R423A) with a 114bp-probe containing a *crtS* site. RctB concentration for each gel is incremented as follows: 5, 10, 20, 40 nM. (C) Mean percentage of RctB/DNA complex formation with *crtS* in three EMSA experiments (\pm standard deviation). Percent of bound DNA is calculated as in Figure 1D. (D) Chr2 copy number relative to Chr1 (*ori2/ori1*) in non-replicating *V. cholerae*. In all mutants, the native *crtS* site is deleted and RctB binding sites are inserted in the *attTn7* insertion site of Chr1 (mean \pm standard deviation).

active in triggering Chr2 replication, we expect a Chr2 copy number (*ori2/ori1* ratio) of about 1 (e.g. WT). If the site is inactive, Chr2 will be lost which will translate into a lower Chr2 copy number (e.g. $\Delta crtS$, *attTn7::O*). Our results show that only *crtS* can trigger Chr2 replication when located on Chr1 and cannot be substituted by any other RctB binding site (i.e. 39m, 1 or 2 iterons). Although the binding of RctB to *crtS* and 39m appears to be similar, only *crtS* can perform its function, suggesting that the triggering of Chr2 initiation by *crtS* may involve other activities of RctB.

RctB domain IV is necessary for both 39m and *crtS* regulation of *ori2* initiation

Although RctB domain IV is not essential for replication initiation, its role in Chr2 copy number control is fundamental (11). Domain IV encompasses RctB handcuffing activity and RctB oligomerization ability, both of which are mediated by protein-protein interactions (10,11,22). We investigated whether these activities could be involved in *crtS*-mediated control of Chr2 replication. To circumvent the toxic effect of Chr2 over-replication in *V. cholerae* caused by the deletion of RctB domain IV, we used our reporter plasmid pORI2 in *E. coli* (10). We built a set of pORI2 with truncations from RctB domain IV by inserting stop codons in the *rctB* gene (Figure 3A). All the truncated versions of pORI2 could replicate autonomously except for the complete deletion of domain IV (ΔIV) for which the few clones that we obtained integrated pORI2 into the chromosome. We measured the copy number of all pORI2 truncated RctB mutants by quantitative dPCR (Figure 3B). We observed that most of the truncations in domain IV increase pORI2 copy number. Only the deletion of 2 AA in RctB ($\Delta 2$) had no impact on pORI2 copy number compared to WT (Figure 3B, white bars). We noticed that, depending on the size of the truncation, there were two levels of increase in pORI2 copy number, suggesting that the negative control of *ori2* replication occurs at two levels in RctB domain IV. Larger deletions of domain IV ($\Delta 157$, $\Delta 147$, $\Delta 137$) considerably increased pORI2 copy number by a factor of about 50 (Figure 3B, dark grey bars). We found that these large deletions overlap with a 71 AA regulatory region previously identified for its role in limiting *ori2* initiation, DNA binding and RctB/RctB interactions (22). Shorter deletions ($\Delta 101$, $\Delta 82$, $\Delta 41$, $\Delta 27$, $\Delta 17$) moderately increased pORI2 copy number by a factor of about 4 (Figure 3B, light grey bars). This region closer to the carboxyl-terminal end of RctB has not yet been studied nor identified for its role in controlling Chr2 replication. To assess the role of RctB domain IV in *crtS*-mediated control of replication, we perform the same assay in *E. coli* carrying a *crtS* site inserted in the chromosome (Figure 3C, Supplementary Figure S6). We observed that most of the deletions in domain IV abolished the control of *crtS* on pORI2 copy number (Figure 3C, $+/- crtS \sim 1$). Only the shortest deletion of RctB ($\Delta 2$) did not abolish the control of *crtS* over pORI2 copy number (Figure 3C, $+/- crtS > 3$). We further assessed the *in vivo* binding of a selection of truncated variants (ΔIV , $\Delta 147$, $\Delta 82$, $\Delta 17$, $\Delta 2$) to the 39m sites by measuring the activity of the Renilla reporter transcribed from the RctB promoter (including a 29m) as in Figure 1C. We observed that in the pres-

ence of RctB, Renilla's activity increased significantly in all truncated mutants (ΔIV , $\Delta 147$, $\Delta 82$, $\Delta 17$) with the exception of RctB $\Delta 2$, to a level equivalent to that obtained in the absence of RctB (Figure 3D). This result means that mutants (ΔIV , $\Delta 147$, $\Delta 82$, $\Delta 17$) are no longer able to repress the 39m-containing RctB promoter. In parallel, we assessed the *in vivo* binding of RctB to iteron sites by measuring the activity of the Renilla reporter transcribed from the *rctA* promoter (PrctA contains two iterons and no 39m sites) as described in (11). Unlike PrctB, the *rctA* promoter remains repressed by all RctB mutants (ΔIV , $\Delta 147$, $\Delta 82$, $\Delta 17$, $\Delta 2$) as well as by RctB WT (Supplementary Figure S7), which suggests that domain IV mutations have no impact on the regulation of iteron-containing PrctA. Thus Chr2 replication regulation by 39m and *crtS* entails the domain IV of RctB. These results reinforce the hypothesis that the binding of RctB to 39m is crucial in the control of Chr2 replication by *crtS*.

The extreme C-terminal of RctB domain IV is crucial for replication coordination by *crtS*

Unlike the other RctB domains, the entire domain IV or its segments have not been crystallized. Our AlphaFold2 structure prediction shows that domain IV is poorly structured between AA 480 and 538 and that the C-terminal part contains a well-structured module (Figure 2A, Video S1). More precisely, the secondary structure prediction analysis of RctB revealed a region with high probability to form an α -helix within the last 16 AA C-terminal extremity (Figure 4A and Supplementary Figure S8). To assess its role in Chr2 replication control and *crtS* activity, we performed an alanine scanning in the 16 AA C-terminal tail of RctB. Substitutions with alanines are expected to alter the interactions in the α -helix without causing major disruption of the overall protein folding (32). We substituted, three residues at a time, 15 AA with alanines (Figure 4A, left panel). The pORI2 mutants were then analysed by quantitative dPCR to measure the effect of alanine substitutions on their copy number and on *crtS* activity. All pORI2 mutants have a copy number that is not significantly affected compared to the WT (Supplementary Figure S9), meaning that the negative regulatory control is still functional when the α -helix is mutated. However, whether mutations were inside or outside the α -helix, the influence of *crtS* on pORI2 copy number is different (Figure 4A, right panel). When RctB is mutated upstream or downstream of the α -helix, the $+/- crtS$ ratio is higher than 1 meaning the *crtS* still exerts a role on pORI2 replication control. When RctB is mutated in the α -helix, the $+/- crtS$ ratio is close 1, which means that *crtS* no longer influences pORI2 replication. We further substituted a conserved leucine located in the middle of the α -helix with a proline (L651P) in order to disrupt the α -helix (32). With the RctB_{L651P} mutation, *crtS* has no longer an effect on pORI2 copy number, whereas with a substitution outside the α -helix, RctB_{G657P}, *crtS* increases pORI2 copy number to a level similar to wild-type (Figure 4A). Thanks to the mild copy-up phenotype caused by the RctB_{L651P} mutation (Supplementary Figure S9), we were able to insert this mutation into *V. cholerae* without it being lethal which allowed us study its effect under endogenous conditions. We

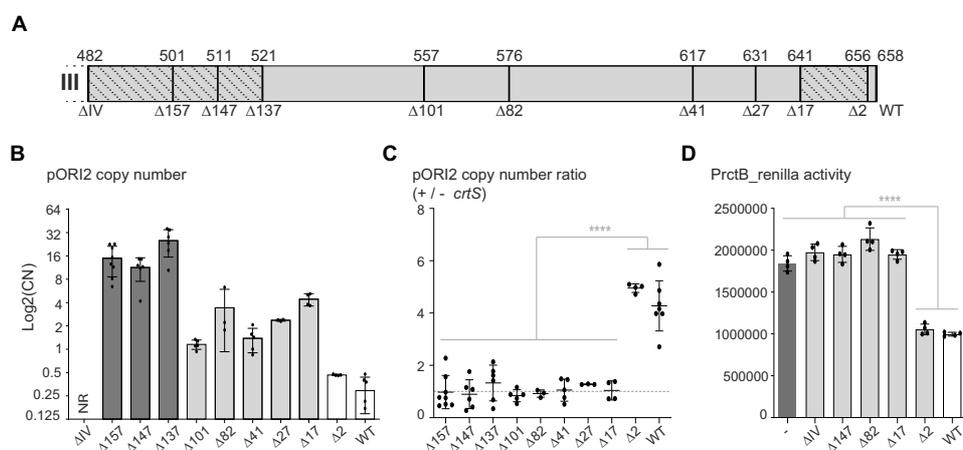


Figure 3. *crtS* regulation of Chr2 replication involves RctB domain IV. (A) Domain IV representation of RctB. Stop codons were inserted to generate truncated proteins as indicated in amino acids coordinates (above, AA) and amino acids deletion (bottom, ΔAA). (B) Copy-number (CN) of pORI2 relative to the chromosome in *E. coli* strains. The various pORI2 derivatives were built by inserting stop codons in *rctB* as indicated in panel A. The Y-axis represents [Log₂(CN)]. NR means not replicative. The different shades of grey represent the two different levels of up-regulation of *ori2* replication upon truncation of domain IV (dark grey: very high 32 > CN > 8; light grey: moderately high 4 > CN > 1). (C) Ratio of pORI2 copy number in *E. coli* with and without a chromosomal *crtS* site (+/- *crtS*). An absence of activation by *crtS* corresponds to a ratio of 1 (indicated by a dashed line). (D) Promoter repression by RctB (and truncated mutants) binding to the 29m in PrctB, as Figure 1C. RctB and mutants are induced with arabinose (0.2%) from a low-copy plasmid (Supplementary Table S2). In all graphs, data represent averages from at least three independent biological replicates ± standard deviations. Each dot represents an independent experiment. Statistics were applied on pairwise comparisons using Student's *t*-test and on multiple comparisons using one-way ANOVA. Statistical significances are indicated (ns: non significant; * *P* < 0.05; ** *P* < 0.01; *** *P* < 0.001; **** *P* < 0.0001).

measured the *ori2/ori1* ratios in presence or absence of the native *crtS* site both in non-replicating (stationary phase) and replicating (exponential phase) conditions to differentiate the effect of *crtS* on Chr2 copy number and Chr2 initiation timing respectively (Figure 4B). We observed that the RctB_{L651P} mutation causes a doubling of the Chr2 copy number compared to Chr1. As in *E. coli*, Chr2 copy number in the RctB_{L651P} mutant of *V. cholerae* is no longer under the influence of *crtS*, whereas it maintains a Chr1:Chr2 ratio of 1:1 in the WT strain (Figure 4B, stationary phase, *ori2/ori1* = 1 in WT with native *crtS*). We also observed, in exponential phase, that the timing of Chr2 replication in the RctB_{L651P} mutant is no longer controlled by *crtS* (Figure 4B, exponential phase, *ori2/ori1* < 1 in WT with native *crtS*). In the RctB_{L651P} mutant, *crtS* thus no longer has control over the copy number and the timing of initiation of Chr2. We further characterized RctB_{L651P} for its promoter repressing activity on PrctB (containing a 29m) and PrctA (containing two iterons) using the luciferase reporter. On PrctB, the Renilla activity, using RctB_{L651P}, is similar to results obtained with the empty plasmid (Figure 4C) while on PrctA, it is similar to results obtained with RctB_{WT}. This observation means that RctB_{L651P} is still capable of repressing PrctA but not PrctB. This de-repression of PrctB was confirmed in *V. cholerae* by RT-dPCR in a RctB_{L651P} mutant where the expression level of RctB increases to >20-fold that of WT (Supplementary Figure S10A). We compared the amount of RctB protein by Western blot in both the 29m(C > A) and RctB_{L651P} *V. cholerae* mutants where

Chr2 replication initiation is no longer under the control of *crtS* (Supplementary Figure S10B). In the 29m(C > A) mutant, we saw an increase in the amount of RctB protein. However, in the RctB_{L651P} mutant, the amount of RctB protein was unaltered compared to WT. This suggests that increasing the amount of RctB is not the only way to overcome the control by *crtS*. The defective binding of RctB_{L651P} to 39m sites could cause the independence of Chr2 replication from *crtS* control. The predicted α-helix in the C-terminal tail of RctB is conserved among 26 *Vibrio* species with the leucine at position 651 being fully conserved (Supplementary Figure S11). We tested the effect of mutating this leucine in two phylogenetically distant *Vibrio* species (*Vibrio nigripulchritudo* and *Grimontia hollisae*). We generated pORI2 derivatives containing the [*ori2*-rctB] regions of *V. nigripulchritudo* and *G. hollisae* that we introduced into *E. coli* containing a chromosomal copy of their corresponding *crtS* sites. Both *V. nigripulchritudo* and *G. hollisae* pORI2 copy number increases in the presence of their respective *crtS* sites in *E. coli*, as it does for *V. cholerae* pORI2 (Figure 4D). When the last leucine of RctB is mutated (*V. cholerae* RctB_{L651P}, *V. nigripulchritudo* RctB_{L651P} and *G. hollisae* RctB_{L653P}), the copy number of the respective pORI2 is no longer controlled by *crtS* (Figure 4D). Note that in *G. hollisae*, the L653P mutation is not copy-up as it is for *V. cholerae* and *V. nigripulchritudo*. These results show that the unique C-terminal tail of RctB is critical in the control of Chr2 replication by *crtS* and that it is conserved among *Vibrionaceae*.

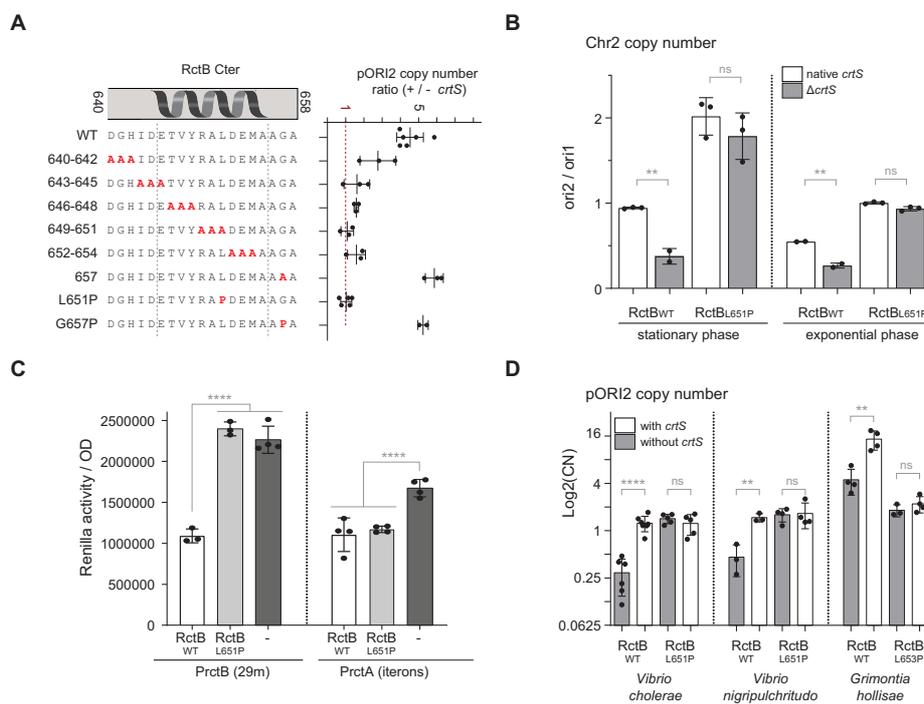


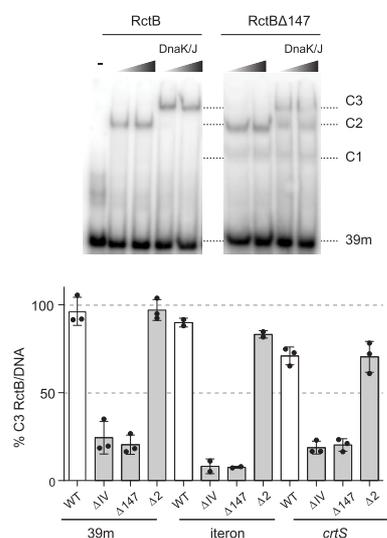
Figure 4. The C-terminal tail of RctB is essential for the control of Chr2 replication by *crtS*. (A) Mutagenesis of the last 19 AA in the carboxy-terminal end of RctB. Left panel: The coordinates of the putative alpha helix (645 to 655 AA), based on Supplementary Figure S8D, are indicated by a dashed line. Alanine and proline substitutions are indicated below the wild-type sequence (WT) in red. Right panel: Ratio of pOR12 copy number in *E. coli* with and without a chromosomal *crtS* site (+/- *crtS*) (data from Supplementary Figure S9). An absence of activation by *crtS* corresponds to a ratio of about 1 (indicated by the red dashed line). (B) *ori2* copy number relative to *ori1* (*ori2/ori1*) in *V. cholerae* wild-type (RctB_{WT}) or with a point mutation in RctB (RctB_{L651P}) in presence (native *crtS*) or absence (Δ *crtS*) of *crtS*. Total genomic DNA was purified either from replicating cells (exponential phase) or from non-replicating cells (stationary phase). (C) Promoter repression by RctB_{WT} and RctB_{L651P} mutants of PrctB (29m) or PrctA (itersons) as in Figure 1C. RctB is induced with arabinose (0.2%) from a low-copy plasmid (Supplementary Table S2). (D) Copy number of pOR12 derivatives relative to the *E. coli* chromosome (pOR12/*oriC*). The pOR12 derivatives carrying the *ori2* and *rctB* genes from either *V. cholerae*, *V. nigrripulchritudo* or *G. hollisae* are transformed into *E. coli* strains carrying their respective *crtS* sites. In all graphs, data represent averages from at least three independent biological replicates \pm standard deviations. Each dot represents an independent experiment. Statistics were applied on pairwise comparisons using Student's *t*-test and on multiple comparisons using one-way ANOVA. Statistical significances are indicated (ns : non significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$).

The extreme C-terminal of domain IV supports RctB self-interaction

In an earlier study, we showed that the domain IV of RctB is required for the cooperative binding and further non-specific oligomerization of RctB onto DNA after its initial specific binding (10). RctB Δ IV is no longer able to oligomerize on itersons even in the presence of DnaK/J chaperones which stimulate RctB cooperative binding (10). For this reason, we explored the *in vitro* binding activities of the RctB truncated mutants on all three binding sites. None of the mutants with deletions ranging from AA 101 to 17, as well as the RctB_{L651P} mutant, could be produced in soluble form. We were however able to purify RctB Δ 157, RctB Δ 147 and RctB Δ 2. We compared their binding activities to RctB_{WT} and RctB Δ IV by EMSA us-

ing probes carrying either a 39-mer, a methylated iterson or a *crtS* site in the presence or absence of DnaK/J chaperones (Supplementary Figure S12). RctB Δ IV, RctB Δ 157 and RctB Δ 147 could hardly form the complex C3 on a 39m site (i.e. three RctB molecules bound to the probe, (10)), even in presence of DnaK/J, suggesting an alteration of RctB oligomerization activity (Figure 5A, Supplementary Figure S12A). Conversely, RctB Δ 2 could still form C3 complexes similar to RctB_{WT}. Similar results were observed with probes containing *crtS* (Supplementary Figure S12B) and methylated itersons (Supplementary Figure S12C). RctB Δ IV, RctB Δ 157 and RctB Δ 147 formed mainly C1 complexes with itersons and *crtS*, whereas RctB Δ 2 bound similarly to RctB_{WT} to all the DNA probes. We quantified the formation of C3 complexes, reflecting the

A



B

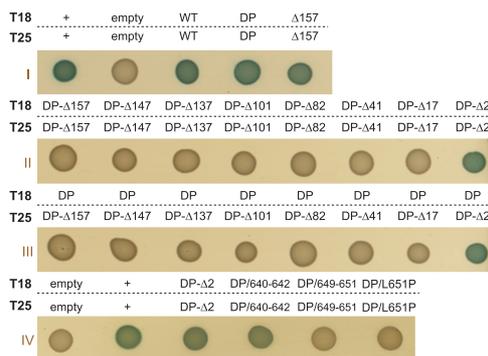


Figure 5. RctB domain IV is involved in RctB self-interaction and cooperative DNA binding. (A) Top panel : Example of cooperative binding of RctB and RctBΔ147 to 39m by EMSA. Free DNA probes (39m) and RctB nucleoprotein complexes (C1, C2 and C3 = 1, 2 and 3 bound RctB respectively) are indicated (as in (10)). The protein concentrations are 10 nM or 20 nM RctB with constant concentrations of 5 nM DnaK_{VC} and 5 nM DnaJ_{VC}. Bottom panel: percentage of RctB/DNA higher complex formation (C3) in EMSA shown in Supplementary Figure S12. EMSAs were performed with RctB (WT) or domain IV mutants (IV, Δ147, Δ2) with DNA probes containing either a 39m site, an iteron, or a *crtS* site. Quantifications were done under conditions using 10 nM RctB and in the presence of DnaK/J_{VC} chaperones. Percent of C3 RctB/DNA complexes is equal to (intensity of the C3 shifted band/combined intensities of C1, C2, C3 shifted bands) × 100. Mean of at least two independent experiments ± standard deviation. Each dot represents one replicate. (B) Bacterial two-hybrid interactions of RctB with itself in *E. coli* on X-Gal plates. Blue indicates interaction and white indicates no interaction. Panel (I) shows the positive leucine zipper (+/+) and negative (empty/empty) controls of interaction, the interactions of RctB_{WT} with itself (WT/WT) and the interactions of the two single mutants RctB_{D314P} (DP/DP) and deletion of 157 AA in domain IV (Δ157/Δ157) with themselves. Panels (II) and (III) show the interactions that were tested with RctB double mutants having an increasing deletion of domain IV (DP/ΔAA). Panel (II): interaction of the double mutants with themselves. Panel (III): interaction of the double mutants with the single mutant RctB_{D314P} (DP). Panel (IV) shows the interactions that were tested within the last C-terminal predicted α-helix of RctB. The vectors used for bacterial two-hybrid are described in Supplementary Table S2.

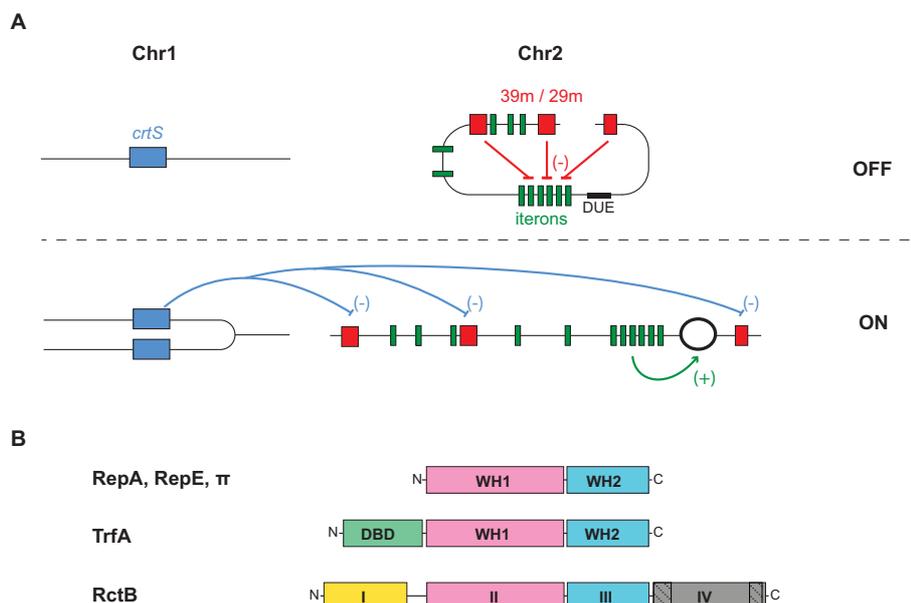


Figure 6. Model for coordination of replication between Chr1 and Chr2 by *crtS*. (A) RctB binding sites are shown in green (iterons), red (39m) and blue (*crtS*). A positive control is represented by an arrow associated with (+), and a negative control is represented by a flat-ended arrow associated with (-). **OFF** = Chr1: *crtS* has not been replicated. Chr2: RctB binding to the 39m sites inhibits replication initiation at *ori2* mainly by handcuffing (red arrows). **ON** = Chr1: *crtS* has been replicated. We hypothesize that the transient binding of RctB to the replicated *crtS* sites results in either a competition with the 39m sites or a remodeling of RctB that results in a decrease in its affinity for the 39m sites (blue arrows). Chr2: RctB binding to 39m is inhibited which releases *ori2*. Binding of RctB to methylated iterons causes opening of the DNA unwinding element (DUE) and oligomerization of RctB onto single-stranded DNA (green arrows) (7). (B) Acquisition of a unique domain IV by the initiator RctB to coordinate replication of Chr1 and Chr2. Comparison of protein domain organization of three types of iteron initiators. The different structural domains are indicated (WH = winged helix, DBD = DNA binding Domain). Additional N-ter domains (such as the domain I of RctB) have been found in TrfA-like plasmid initiators (DBD) (43). The role of domain I in RctB is still elusive. Additional C-terminal domains (such as the domain IV of RctB) have not found in iteron-like plasmid initiators. RctB domain IV is crucial for coordinating Chr1 and Chr2 replication. Representation adapted from (43).

whereas previous studies have been performed with DnaK_{Ec} and DnaJ_{Ec} from *E. coli* (20). We already observed such discrepancies in the binding of RctB to its DNA substrates between *E. coli* and *V. cholerae* (10). This suggests that species-specific interactions between DnaK_{Vc}/J_{Vc} and RctB are important to promote the binding of additional RctB monomers to DNA (10).

In conclusion, it appears that the positive control of *crtS* on *ori2* replication initiation occurs through the abolishment of 39m inhibitory activities. Competition for RctB binding between *crtS* and 39m could impede the handcuffing of *ori2* by 39m, allowing replication to be initiated. Importantly, a 39m does not functionally replace *crtS*, even in trans on *V. cholerae* Chr1 (Figure 2D), implying that RctB-*crtS* and RctB-39m interactions are different; only the RctB-*crtS* interaction can exert a role in initiating Chr2 replication. How the interactions between RctB-*crtS* and RctB-39m differ is not yet understood, and may be due either to the specific DNA sequences contained in *crtS* or to the binding of the transcription factor Lrp (34).'

The self-interaction of RctB molecules through their C-terminal tails is crucial for *crtS* control of Chr2 replication

The down-regulation of *ori2* replication initiation by RctB requires its domain IV (11). This domain, whose structure is still unknown, does not harbor a known DNA binding motif, although it has been shown to bind to DNA (9,22). However, domain IV is required for RctB handcuffing and oligomerization activities (10,11,22). We used the recently released AlphaFold2 algorithm (29), which uses machine learning to predict protein structures with high accuracy, in order to predict the full length RctB secondary structure (Figure 2A, Video S1). This gives us new information about how RctB may interact with DNA and with itself. In this model, it clearly seems that all three HTH are close together within a pocket that could accommodate the DNA molecule, in agreement with Orlova's model (9). We also see that RctB domain IV is separated from the other three domains. It is therefore conceivable that RctB C-terminal part is more accessible to interact with the C-terminal module of another RctB molecule. This would be consistent with our

in vivo observations of domain IV-domain IV interactions (Figure 5B).

In this study, we discovered a novel function of domain IV, which allows for the positive regulation of Chr2 replication by *crtS*. All domain IV RctB truncated mutants, except for RctB Δ 2, are no longer regulated by *crtS* (Figures 3C, 4A, B). This is in contrast to published results showing replication-stimulating activity by *crtS* in the RctB Δ 157 mutant (20). These discrepancies can be explained by our different setups. In our system, *crtS* is inserted into the *E. coli* chromosome and the entire [*ori2*-RctB] region is on the same plasmid, whereas in Baek's setup three plasmids contain the three components separately: RctB, *ori2* and *crtS*. In our study, once the integrity of domain IV is affected, the replication-enhancing activity of *crtS* is lost in *E. coli* (Figures 3C, 4A) and this observation was confirmed in *V. cholerae* (Figure 4B).

We observed that RctB domain IV includes at least two regions of replication regulation, which are reflected by two levels of Chr2 copy limitation (Figure 3B). The first region, between AA 501 and 521 (Figure 3A), overlaps a regulatory region previously characterized for its role in Chr2 copy number down-regulation, binding to DNA and RctB dimerization (22). In both plasmids and Chr2, handcuffing is the main safeguard mechanism that keeps copy number low. Increasing the concentration of initiator above the physiological level does not significantly increase plasmid copy number, whereas altering handcuffing can drastically increase plasmid copy number (8). Since all three RctB Δ 157, RctB Δ 147 and RctB Δ 137 mutants are deleted for the dimerization interface involved in handcuffing, this explains the high pOR12 copy number (Figure 3B). *In vivo*, RctB handcuffing mutants can be studied in *E. coli* using pOR12 (Figure 3B) but not in *V. cholerae* because these mutations are likely to be lethal due to the toxicity caused by excessive Chr2 copy number. These RctB domain IV mutants (RctB Δ 157, RctB Δ 147 and RctB Δ 137) are defective in cooperatively binding to DNA (Figure 5A, Supplementary Figure S12). These mutants preferentially form C1 complexes with DNA, corresponding to the binding of a monomer of RctB to DNA (10). This implies that the cooperative binding of RctB is not required for initiation of *ori2* replication. Plasmid initiators catalyze origin opening via a DNA distortion mechanism (35,36). DNA distortion is induced by binding of the plasmid initiator to an iteron array (37,38). It is likely that binding of RctB to the six-iteron array of *ori2* catalyzes the initiation of *ori2* replication by mediating a change in DNA conformation (7) without requiring RctB cooperative binding to DNA.

The second, newly identified region between AA 641 and 656 (Figure 3A, Supplementary Figure S9) is mildly involved in the control of Chr2 copy number (Figure 3B). All truncations from Δ 17 to Δ 101 have the same mild impact on pOR12 copy number. Between Δ 2 and Δ 17 truncations resides a predicted α -helix, conserved in other *Vibrionaceae* (Supplementary Figures S8, S11). We show that this putative α -helix is crucial for *crtS* control of *ori2* replication (Figure 4A) whereas its disruption has only a limited impact on pOR12 copy number (Supplementary Figure S9) indicating that the α -helix is not involved in the negative control of RctB via handcuffing with 39m. When a leucine

of the alpha helix is substituted with a proline (L651P) so as to disrupt its structure (39), replication at *ori2* is no longer controlled by *crtS* (Figure 4A, B). Because of its limited impact on *ori2* initiation negative regulation, we could transpose the RctB_{L651P} mutation in *V. cholerae*. Even though the L651P mutation increases the level of RctB expression by 20-fold that of WT (Supplementary Figure S10), the number of Chr2 copies is only doubled (Figure 4B) suggesting that RctB concentration level is not a major limiting factor of Chr2 replication control since its overexpression does not considerably increase Chr2 copy number. It was shown in *E. coli* that an increase of RctB concentration does not prevent *crtS* enhancing activity on pOR12 copy number (20,23), thus an increase of RctB concentration induced by the L651P mutation does not explain the loss of *crtS* control on *ori2*.

We analysed the binding affinities of the RctB domain IV mutants to 39m and iterons by monitoring their abilities to repress the PrctB and PrctA promoters respectively (Supplementary Figure S1). All domain IV mutants that are no longer regulated by *crtS* (Figures 3C, 4A, B) also fail to repress the PrctB promoter (Figures 3D, 4C) while retaining their PrctA repression activity (Figure 4C, Supplementary Figure S7). This indicates that binding to 39m, but not to iterons, is impaired in *crtS*-independent mutants. We show *in vitro* that RctB Δ IV, RctB Δ 157 and RctB Δ 147 still bind to iteron, 39m and *crtS*, but no longer oligomerize (Figure 5A, Supplementary Figure S12). Thus, it appears that, *in vivo*, binding to 39m or repression of PrctB depends on RctB cooperative binding capacity. It is known that some transcriptional regulators repress gene expression by forming a nucleoprotein filament on their promoter (40). If this was the case for RctB, we would expect PrctA to be repressed as well, but this is not the case, suggesting that only 39m binding is impaired in *crtS* independent mutants. We have previously reported that oligomerization of RctB on DNA, requires protein-protein interactions mediated by its domain IV (10). Here, we show that all *crtS*-independent RctB mutants exhibit a defect in RctB-RctB self-interaction by domain IV (Figure 5B). More specifically, we found that RctB_{L651P}, a *crtS*-independent mutant obtained from a single AA substitution in the extreme C-terminal tail of RctB, is defective for RctB-RctB interactions and PrctB repression. This suggests that these RctB-RctB interactions are involved in the repressing activity of RctB on 29m. However, RctB_{L651P} is weakly involved in the negative regulation of Chr2 copy number, suggesting that these interactions are not involved in *ori2* handcuffing by the 39m.

We can highlight a similarity with the DnaA initiator for which its oligomerization activity is also essential regarding trans-acting sites. In *E. coli*, replication regulation requires a tight control of the cellular concentration of the active initiator DnaA-ATP. Both the reactivation of DnaA-ADP to DnaA-ATP and the hydrolysis of ATP require specific DNA loci (DARS2 and *datA*) and inter-DnaA interactions (41,42). It was shown that mutation in the domain III of DnaA (L290A), which mediates protein self-interaction, severely impairs DARS2-mediated ADP dissociation (41). Thus, it appears that for both chromosome initiators and secondary chromosome initiators (*i.e.* DnaA and RctB),

protein self-interaction is required in the regulation of replication involving a DNA locus other than the origin.

CONCLUDING REMARKS

This study advances our understanding of the coordination of Chr1–Chr2 replication by providing new insights into how Chr2 replication initiation is triggered by *crtS*. Our results indicate that *crtS* acts as an anti-inhibitory site by preventing 39m from negatively regulating *ori2* replication. Thus, we believe that Chr2 replication must be inhibited by RctB binding to 39m for much of the cell cycle (Figure 6A, OFF) and is briefly released when the replication fork passes through the *crtS* site (Figure 6A, ON). The replication of *crtS* would lift the inhibition of Chr2 initiation by the 39m sites, thereby triggering Chr2 replication. Hence, RctB binding to *crtS* may vary depending on its replicative state. Based on past and present results, it is likely that doubling of *crtS* may shift RctB binding from 39m sites to *crtS* allowing *ori2* to be released for replication. We cannot rule out that RctB bound to *crtS* interacts directly with 39m, thereby preventing *ori2* handcuffing, nor can we rule out that a modification of RctB upon binding to *crtS* decreases the affinity of RctB for 39m (10,18,20,23). To verify this, a study of the dynamics of RctB binding at the whole genome level throughout the cell cycle is required.

We also showed that the extreme C-terminal tail of RctB, containing a predicted α -helix, is crucial for the coordination of replication of both chromosomes by *crtS*. Both the C-terminal end of RctB and the activation of *ori2* by *crtS* are found in other *Vibrionaceae* (Supplementary Figure S11; (16)), hence this mechanism is likely to be found in all *Vibrionaceae* species. RctB C-terminal domain has no homologue in other plasmid initiators. The acquisition of an additional domain, here, especially dedicated to *Vibrio* Chr2 replication regulation, illustrates that the development of new functionalities by domain acquisition could be a more general feature of evolution of replication initiators and other bacterial proteins involved in DNA metabolism (Figure 6B). Recently, a unique third domain was identified in TrfA, the Rep initiator of the broad host range RK2 iteron-plasmids. The third domain of TrfA is essential for replication and has led to the definition of a new class of TrfA-like initiators containing three DNA-interacting domains, WH1, WH2 and DBD (43). A similar phenomenon was also observed for integron integrases, where the acquisition of a protein domain led to a functional innovation (44).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Alfonso Soler-Bistué for his help in the development of the Luciferase reporter and sharing of the pASB12 plasmid expressing Renilla Luciferase; Pr. Dhruba Chattoraj for sharing the pTVC11 plasmid; Dr Zeynep Baharoglu, Evelyn Krin and Manon Lang for technical advice on luciferase experiments; Pr. Melanie Blokesch

for help with RNA preparation in *Vibrio* and RT-qPCR; Dr Gouzel Karimova, Pr. Daniel Ladant for sharing pKT25 and pUT18C plasmids; Dr Rocio Lopez-Igual for technical advice on bacterial two-hybrid experiments; Dr Sean Kennedy for help with writing the manuscript, and Stilla Technologies for help with the design of RT-dPCR experiments. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. We also would like to thank all our team members for helpful discussions.

FUNDING

Institut Pasteur; Institut National de la Santé et de la Recherche Médicale (INSERM); Centre National de la Recherche Scientifique [CNRS-UMR 3525]; French National Research Agency, Jeunes Chercheurs [ANR-19-CE12-0001]; Laboratoires d'Excellence [ANR-10-LABX-62-IBEID]; F.F. was supported by ANR-10-LABX-62-IBEID and ANR-19-CE12-0001; T.N. was supported by the Ministère de l'Enseignement Supérieur et de la Recherche; J.C. was supported by Institut Pasteur (Cantarini foundation) [ANR-19-CE12-0001, ANR-10-LABX-62-IBEID]. Funding for open access charge: French National Research Agency [ANR-19-CE12-0001].

Conflict of interest statement. None declared.

REFERENCES

- diCenzo, G.C. and Finan, T.M. (2017) The divided bacterial genome: structure, function, and evolution. *Microbiol. Mol. Biol. Rev.*, **81**, e00019-17.
- Harrison, P.W., Lower, R.P., Kim, N.K. and Young, J.P. (2010) Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol.*, **18**, 141–148.
- Touchon, M. and Rocha, E.P. (2016) Coevolution of the organization and structure of prokaryotic genomes. *Cold Spring Harb. Perspect. Biol.*, **8**, a018168.
- Demarre, G. and Chattoraj, D.K. (2010) DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle. *PLoS Genet.*, **6**, e1000939.
- Duigou, S., Knudsen, K.G., Skovgaard, O., Egan, E.S., Lobner-Olesen, A. and Waldor, M.K. (2006) Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J. Bacteriol.*, **188**, 6419–6424.
- Egan, E.S. and Waldor, M.K. (2003) Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell*, **114**, 521–530.
- Chatterjee, S., Jha, J.K., Ciaccia, P., Venkova, T. and Chattoraj, D.K. (2020) Interactions of replication initiator RctB with single- and double-stranded DNA in origin opening of *Vibrio cholerae* chromosome 2. *Nucleic Acids Res.*, **48**, 11016–11029.
- Koniczny, L., Bury, K., Wawrzycka, A. and Wegryn, K. (2014) Itron Plasmids. *Microbiol. Spectr.*, **2**, <https://doi.org/10.1128/microbiolspec.PLAS-0026-2014>.
- Orlova, N., Gerding, M., Ivashkiv, O., Olinares, P.D.B., Chait, B.T., Waldor, M.K. and Jeruzalmi, D. (2017) The replication initiator of the cholera pathogen's second chromosome shows structural similarity to plasmid initiators. *Nucleic Acids Res.*, **45**, 3724–3737.
- de Lemos Martins, F., Fournes, F., Mazzuoli, M.V., Mazel, D. and Val, M.E. (2018) *Vibrio cholerae* chromosome 2 copy number is controlled by the methylation-independent binding of its monomeric initiator to the chromosome 1 *crtS* site. *Nucleic Acids Res.*, **46**, 10145–10156.

11. Jha, J.K., Demarre, G., Venkova-Canova, T. and Chattoraj, D.K. (2012) Replication regulation of *Vibrio cholerae* chromosome II involves initiator binding to the origin both as monomer and as dimer. *Nucleic Acids Res.*, **40**, 6026–6038.
12. Jha, J.K., Li, M., Ghirlando, R., Miller Jenkins, L.M., Wlodawer, A. and Chattoraj, D. (2017) The DnaK chaperone uses different mechanisms to promote and inhibit replication of *Vibrio cholerae* chromosome 2. *MBio*, **8**, e00427-17.
13. Venkova-Canova, T. and Chattoraj, D.K. (2011) Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 6199–6204.
14. Ramachandran, R., Jha, J., Paulsson, J. and Chattoraj, D. (2017) Random versus cell cycle-regulated replication initiation in bacteria: insights from studying *Vibrio cholerae* chromosome 2. *Microbiol. Mol. Biol. Rev.*, **81**, e00033-16.
15. Fournes, F., Val, M.E., Skovgaard, O. and Mazel, D. (2018) Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes. *Front Microbiol*, **9**, 1833.
16. Kemter, F.S., Messerschmidt, S.J., Schallopp, N., Sobetzko, P., Lang, E., Bunk, B., Sproer, C., Teschler, J.K., Yildiz, F.H., Overmann, J. et al. (2018) Synchronous termination of replication of the two chromosomes is an evolutionary selected feature in Vibrionaceae. *PLoS Genet.*, **14**, e1007251.
17. Rasmussen, T., Jensen, R.B. and Skovgaard, O. (2007) The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J.*, **26**, 3124–3131.
18. Val, M.E., Marbouty, M., de Lemos Martins, F., Kennedy, S.P., Kemble, H., Bland, M.J., Possoz, C., Koszul, R., Skovgaard, O. and Mazel, D. (2016) A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.*, **2**, e1501914.
19. Venkova-Canova, T., Saha, A. and Chattoraj, D.K. (2012) A 29-mer site regulates transcription of the initiator gene as well as function of the replication origin of *Vibrio cholerae* chromosome II. *Plasmid*, **67**, 102–110.
20. Baek, J.H. and Chattoraj, D.K. (2014) Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS Genet.*, **10**, e1004184.
21. Venkova-Canova, T., Srivastava, P. and Chattoraj, D.K. (2006) Transcriptional inactivation of a regulatory site for replication of *Vibrio cholerae* chromosome II. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 12051–12056.
22. Jha, J.K., Ghirlando, R. and Chattoraj, D.K. (2014) Initiator protein dimerization plays a key role in replication control of *Vibrio cholerae* chromosome 2. *Nucleic Acids Res.*, **42**, 10538–10549.
23. Ramachandran, R., Ciaccia, P.N., Filsuf, T.A., Jha, J.K. and Chattoraj, D.K. (2018) Chromosome I licenses chromosome 2 replication in *Vibrio cholerae* by doubling the *crtS* gene dosage. *PLoS Genet.*, **14**, e1007426.
24. Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 5752–5756.
25. Madic, J., Zocivic, A., Senlis, V., Fradet, E., Andre, B., Muller, S., Dangla, R. and Droniou, M.E. (2016) Three-color crystal digital PCR. *Biomol. Detect. Quantif.*, **10**, 34–46.
26. Lo Scudato, M. and Blokesch, M. (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet.*, **8**, e1002778.
27. Soler-Bistue, A., Aguilar-Pierle, S., Garcia-Garcera, M., Val, M.E., Sismeiro, O., Varet, H., Seira, R., Krin, E., Skovgaard, O., Comerci, D.J. et al. (2020) Macromolecular crowding links ribosomal protein gene dosage to growth rate in *Vibrio cholerae*. *BMC Biol.*, **18**, 43.
28. Pal, D., Venkova-Canova, T., Srivastava, P. and Chattoraj, D.K. (2005) Multipartite regulation of *ctb*, the replication initiator gene of *Vibrio cholerae* chromosome II. *J. Bacteriol.*, **187**, 7167–7175.
29. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A. et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
30. Mirdita, M., Ovchinnikov, S. and Steinegger, M. (2021) ColabFold - making protein folding accessible to all. bioRxiv doi: <https://doi.org/10.1101/2021.08.15.456425>, 15 August 2021, preprint: not peer reviewed.
31. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
32. Alias, M., Ayuso-Tejedor, S., Fernandez-Recio, J., Cativiela, C. and Sancho, J. (2010) Helix propensities of conformationally restricted amino acids. Non-natural substitutes for helix breaking proline and helix forming alanine. *Org. Biomol. Chem.*, **8**, 788–792.
33. Srivastava, P. and Chattoraj, D.K. (2007) Selective chromosome amplification in *Vibrio cholerae*. *Mol. Microbiol.*, **66**, 1016–1028.
34. Ciaccia, P.N., Ramachandran, R. and Chattoraj, D.K. (2018) A requirement for global transcription factor Lrp in licensing replication of *Vibrio cholerae* chromosome 2. *Front Microbiol*, **9**, 2103.
35. Wegrzyn, K.E., Gross, M., Uciechowska, U. and Konieczny, I. (2016) Replisome assembly at bacterial chromosomes and iteron plasmids. *Front Mol Biosci*, **3**, 39.
36. Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S.Y., Yokoyama, S., Kurumizaka, H. and Katayama, T. (2008) A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J. Biol. Chem.*, **283**, 8351–8362.
37. Kawasaki, Y., Matsunaga, F., Kano, Y., Yura, T. and Wada, C. (1996) The localized melting of mini-F origin by the combined action of the mini-F initiator protein (RepE) and HU and DnaA of *Escherichia coli*. *Mol. Gen. Genet.*, **253**, 42–49.
38. Kelley, W.L. and Bastia, D. (1992) Activation in vivo of the minimal replication origin beta of plasmid R6K requires a small target sequence essential for DNA looping. *New Biol.*, **4**, 569–580.
39. Lefevre, F., Remy, M.H. and Masson, J.M. (1997) Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function. *Nucleic Acids Res.*, **25**, 447–448.
40. Lim, C.J., Lee, S.Y., Kenney, L.J. and Yan, J. (2012) Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. *Sci. Rep.*, **2**, 509.
41. Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T. and Katayama, T. (2014) Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production and replication initiation. *Nucleic Acids Res.*, **42**, 13134–13149.
42. Kasho, K. and Katayama, T. (2013) DnaA binding locus *datA* promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 936–941.
43. Wegrzyn, K., Zabrocka, E., Bury, K., Tomiczek, B., Wiczor, M., Czub, J., Uciechowska, U., Moreno-Del Alamo, M., Walkow, U., Grochowina, I. et al. (2021) Defining a novel domain that provides an essential contribution to site-specific interaction of Rep protein with DNA. *Nucleic Acids Res.*, **49**, 3394–3408.
44. Escudero, J.A., Nivina, A., Kemble, H.E., Loot, C., Tenailon, O. and Mazel, D. (2020) Primary and promiscuous functions coexist during evolutionary innovation through whole protein domain acquisitions. *Elife*, **9**, e58061.

II. Second article: “Genome-wide binding of secondary chromosome initiator reveals mechanism for replication coordination in *Vibrio cholerae*”

Article en préparation:

Le point de départ de cette étude est le développement de la méthode de ChIPseq à haute résolution sur RctB, l'initiateur du Chromosome 2. Celle-ci nous permet de capturer les interactions protéines ADN *in vivo* sur l'ensemble du génome et de manière quantitative. Ce travail a été inspiré notamment par les résultats de ChIPseq de la protéine de partition ParB chez *Vibrio cholerae* dans l'équipe de Jean-Yves Bouet (Debaugny et al., 2018). L'utilisation de cette méthode leur a permis de bien mieux comprendre le mécanisme de *spreading* de la protéine ParB, et comme nous savions que RctB peut oligomériser sur l'ADN de manière aspécifique nous pensions que le ChIPseq était une méthode pertinente pour mieux comprendre la régulation de l'*ori2* par RctB.

Une fois le ChIPseq de RctB réalisé, nous avons obtenu un résultat frappant, sur l'*ori2* RctB se fixe uniquement sur les sites 39m. Nous avons également détecté de nouveaux sites de fixation et la fixation sur *crtS*. Le ChIPseq étant réalisé sur une population de cellules asynchrones il donne la fréquence de fixation d'une protéine pendant tout le cycle cellulaire. Ainsi pendant la majorité du cycle l'*ori2* est bloquée car RctB se fixe sur les sites inhibiteurs.

Puis nous avons décidé de réaliser ce ChIPseq dans une population synchronisée pour comprendre la temporalité des évènements de fixation de RctB. Pour cela nous avons utilisé une méthode récemment publiée qui permet de tirer avantage de la réponse stringente pour synchroniser les cellules de *Vibrio cholerae*. Cette méthode permet une synchronisation imparfaite mais en l'absence d'autres méthodes disponibles nous avons choisi de réaliser le ChIPseq dans ces conditions. Dans une population synchronisée on observe que RctB est fixé sur l'*ori2* uniquement aux 39mers en début de cycle, puis lorsque le site *crtS* est répliqué on observe un shift de fixation depuis les 39m vers les itérons.

Nous avons également réalisé ce ChIPseq dans une souche $\Delta crtS$, on observe alors une forte augmentation de la fixation de RctB aux 39m de l'*ori2*. Ainsi le rôle de *crtS* semble bien être un anti-inhibiteur en réduisant la fixation aux 39m. Par microscopie électronique on observe que les 39m seuls suffisent à former des boucles dans l'ADN, on propose un nouveau

modèle pour le complexe d'inhibition avec les 39m qui agissent par homo-handcuffing. De plus, on a également observé de possible contact 39mer-*crtS* médié par RctB.

Enfin pour poursuivre cette étude, nous avons réalisé le ChIPseq d'un mutant du domaine IV de RctB L651P. Celui-ci ne se fixe plus à aucun 39mer du génome, sur l'*ori2* on observe une fixation sur les itérons et la région DUE. Nous comprenons maintenant pourquoi l'initiation du Chr2 dans cette souche est indépendante de *crtS* et pourquoi il y a une perte de la synchronie de terminaison.

On a décidé ensuite de travailler en collaboration avec l'équipe d'Abel Garcia Pino pour obtenir la structure du domaine IV de RctB alors encore inconnue. Cette structure révèle que le domaine IV contient deux surfaces de dimérisation. Ces surfaces semblent être essentiels à la fixation aux 39mer mais aussi pour la régulation par *crtS*.

Genome-wide binding of secondary chromosome initiator reveals mechanism for replication coordination in *Vibrio cholerae*

Théophile Niault^{1,2,3} Ariel Talavera⁴, Eric Le Cam⁵, Florian Fournès^{1,2}, Léa Wagner^{1,2}, Abel Garcia Pino⁴, Didier Mazel^{1,2} and Marie-Eve Val^{1,2}

¹ Institut Pasteur, Université de Paris, Unité Plasticité du Génome Bactérien, Département Génomes et Génétique, Paris 75015, France.

² Centre National de la Recherche Scientifique, UMR3525, Paris 75015, France.

³ Sorbonne Université, Collège Doctoral, Paris 75005, France

⁴ Cellular and Molecular Microbiology, Faculté des Sciences, Université libre de Bruxelles (ULB), Boulevard du Triomphe, Brussels, Belgium.

⁵ Genome Integrity and Cancer UMR 9019 CNRS, Université Paris Saclay, Gustave Roussy, Villejuif 94805, France

Contribution personnel: constructions souches et plasmides, ChIPseq (expérimentation et analyse), double hybride, microscopie à fluorescence.

Abstract

Bacterial genomes can expand through the acquisition of megaplasms, which eventually become secondary chromosomes encoding vital functions. The cholera pathogen *Vibrio cholerae* has a secondary chromosome (Chr2) that is stably maintained by tuning its replication to that of the main chromosome (Chr1) through the replication of a non-coding sequence (*crtS*) on Chr1. Our study investigated how this is achieved by studying the genome-wide binding of the Chr2 replication initiator, RctB, at different stages of the *V. cholerae* cell cycle. We found that RctB is predominantly bound to inhibitory sites that prevent Chr2 initiation until *crtS* is replicated. This causes a shift in RctB binding to activating sites, enabling Chr2 replication. The structure of RctB domain IV suggests it plays a role in preventing Chr2 initiation. Our findings support a model where *crtS* destabilizes an inhibition complex that locks the origin of Chr2 replication.

INTRODUCTION

Bacteria can acquire new functions through megaplasms that allow them to adapt to different environments ¹. In some cases, essential housekeeping genes are transferred from the main chromosome to the megaplasmid, creating a secondary chromosome ². This is seen in the Vibrionaceae family, which all have a secondary chromosome ³. The stability of these secondary chromosomes depends on the control of their replication with the host cell cycle ⁴. The initiation of DNA replication is a highly regulated process that controls the entire bacterial cell cycle ^{5,6}. Bacteria with multipartite genomes require additional mechanisms to coordinate the replication of each replicon ⁷. The *Vibrio cholerae* pathogen provides a well-studied example of this coordination, with its secondary chromosome (Chr2) replicating in synchrony with the main chromosome (Chr1) to ensure that both replicate once per cell cycle and terminate together ⁸⁻¹⁰. This synchronization mechanism has been observed in all Vibrionaceae ¹¹ and may also exist in other bacteria with multipartite genomes ¹²⁻¹⁷.

In *V. cholerae*, Chr2 originates from an iteron-type plasmid as evidenced by its replication origin (*ori2*) and initiator (RctB) ¹⁸. RctB is composed of four structural domains (I – IV) ¹⁹ (Supplementary Fig. 1a). Its central domains (II, III) are structurally similar to the Rep iteron plasmid initiators domains (WH1, WH2) while its outer domains (I, IV) are unique to RctB ¹⁹. Each of the domains (I, II and III) interacts with DNA via Helix-Turn-Helix (HTH) motifs and is essential for *ori2* replication ¹⁹. The C-terminal domain (IV) is required to downregulate Chr2 initiation and to coordinate Chr1 and Chr2 replication ^{20,21}. Like Rep, RctB molecules form inactive dimers in the cytoplasm via their dimerization interface in domain II ¹⁹. RctB dimers are monomerized by the DnaK/J chaperones to allow their binding to DNA ^{20,22,23}. RctB monomers cooperatively bind an array of six regularly spaced iterons (12 mer) in *ori2*, which promotes the opening of the DNA helix at a nearby AT-rich region called the DNA Unwinding Element (or DUE) (Supplementary Fig. 1b). RctB can then bind to the bottom single strand of the open DUE on six regularly spaced direct repeats (5'-ATCA) ²⁴. A nucleoprotein complex is then formed, consisting of RctB, IHF and DnaA, which probably enables the recruitment of the replicative helicase DnaB, as is the case for iteron plasmids ^{25 1153}. RctB binds to another type of site (29/39 mer) that differs from iterons in sequence and size (Supplementary Fig. 1a). Three 29/39mer sites flank the *ori2* region and the binding of RctB to these sites inhibit *ori2* initiation by a mechanism called handcuffing : folding of the origin by protein-protein interactions that cause steric hindrance ^{26,27}. Furthermore, the 29mer site is located in the promoter of *rctB* which can thereby repress its own expression ²⁷. RctB binds to a third kind of site on Chr1 ⁹. This unique 62-mer site called *crtS* is located midway between the origin and terminus of Chr1 and acts as a checkpoint to coordinate the replication of Chr2 with Chr1 ^{10,28}.

When *crtS* is deleted, bacteria undergo lethal loss of Chr2 due to an under-initiation of *ori2* replication¹⁰. Upon replication, *crtS* triggers the initiation of Chr2 by a mechanism that is still largely unknown¹⁰. Recently, we showed that the action of *crtS* was associated with the inhibitory action of the 29/39mer sites since alteration of their activity allowed Chr2 to initiate independently of *crtS*²¹.

In this study, we performed a genome-wide analysis of RctB binding using ChIP-seq to better understand the mechanism of *crtS* activation. Our results showed that RctB predominantly binds to 29/39mer inhibitory sites on *ori2*, hindering the replication of Chr2. However, following *crtS* replication, we observed a transition in RctB binding from 29/39mers to iterons, activating Chr2 replication. This activation appears to be direct between *crtS* and *ori2* and involves domain IV of RctB. We determined the structure of RctB domain IV, which contains two dimerization interfaces that are essential for 29/39mer binding in vivo. Our findings support a model in which *crtS* destabilizes an inhibition complex mediated by RctB domain IV and 29/39mer sites that locks the origin of Chr2 replication.

RESULTS

RctB binds predominantly to replication inhibitory sites at *ori2*

A former ChIP-on-chip of RctB in *V. cholerae* identified 5 RctB binding regions outside *ori2*: 4 sites on Chr2 and a unique site on Chr1 (*crtS*)⁹. To increase the resolution of RctB binding profile at *ori2*, we developed the ChIP-seq of RctB in *V. cholerae*, which also provided us with a binding profile of RctB on the rest of the genome. Quite surprisingly, we found that RctB bound to many more regions along the genome than was known: 9 on Chr1 and 6 on Chr2, mostly within CDSs (Fig.1a, Supplementary Fig. 2). Peak analysis using the MEME suite²⁹ identifies mainly two patterns: the iteron-like containing a central inverted direct 5'-TGATCA repeat and the 39mer-like containing a conserved 5'-TTACGG motif (Fig.1b). In addition, we identified three sites that shared no sequence similarity with iterons and 39mer, including *crtS* (Supplementary Table 1). Since the ChIP-seq was performed from exponential culture of non-synchronized replicating bacteria, the height of the peaks gives a good indication of the frequency of RctB binding during the cell cycle. Zooming in on *ori2* (Fig.1c), we observe that the ChIP signal is predominantly at the 29/39mer sites and absent from the iteron sites, meaning that RctB is more frequently bound to the inhibitory sites, suggesting that Chr2 replication initiation is quenched for most of the cell cycle. The signal at the 29mer site is about 3-fold higher than that at the two 39mer sites. One reason is that the ParB2 protein, which is essential for Chr2 segregation, binds and spreads in the *rctA* region and also binds directly to the 39mer, thus entering into competition with RctB³⁰. A second explanation is that the

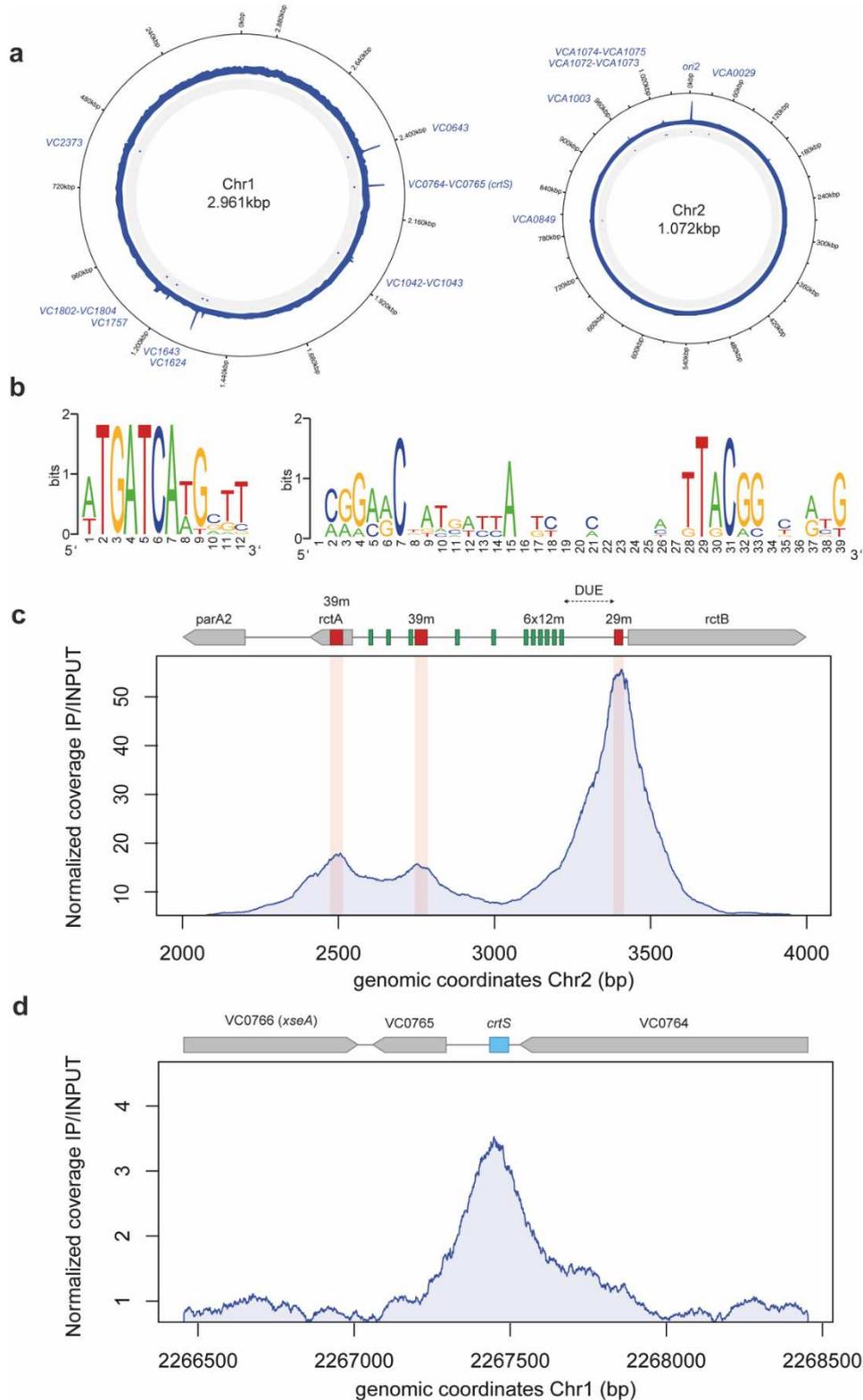


Figure 1: Chr2 initiator RctB binds preferentially to inhibitory sites on *ori2*.

a. Circular map of *Vibrio cholerae* genome using shinyCircos³¹, RctB ChIP signal is depicted in blue, significant peaks enrichment are marked with a blue dot and proximal genes are depicted. **b.** ChIP peaks motif analysis using MEME suite²⁹, two sequences are significantly enriched and correspond to already known iterons (left) and 39mer sites (right). **c & d.** on y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference genome (CP028827.1 and CP028828.1). A 2.000bp window is represented and centered around the peak. Genetic context and known RctB binding sites are depicted on top of the graph: genes as grey arrow, 39mer site as red box, iterons a green triangle, *crtS* as cyan box. The AT-rich region also called DNA unwinding element (DUE) is also depicted.

probability of RctB binding to the 29mer is greater because the protein is synthesized near its binding site. Zooming in on *crtS* (Fig.1d), the ChIP signal extends over a wider area (~800 bp) than for the other peaks (~400 bp) and in an asymmetric manner with a shoulder downstream of *crtS*. This extended signal outside of *crtS* could be a signature of RctB's propensity to oligomerize non-specifically on DNA, as previously seen in vitro ²⁸.

We tested the role of the newly identified RctB binding sites on Chr2 replication. For this, we used a plasmid *pORI2*, whose replication depends solely on *ori2* ²⁸. By quantitative digital PCR (dPCR), we monitored the copy number variation of *pORI2* in *E. coli* MG1655 without and with a chromosomal copy of the RctB binding site to be tested. Using this system, we showed that the *crtS* site increases *pORI2* copy number ²⁸. None of the new sites tested had a significant effect on *pORI2* replication (Supplementary Fig. 3a). Similarly, the inactivation of three of the sites (VC0643, VC01643 and VC1042-VC1043) showed no impact on Chr2 copy number in *V. cholerae* (Supplementary Fig. 3b). Because RctB can repress its own transcription by binding to its gene promoter ²¹, we monitored the expression of the VC1803 and VC1042 genes having an RctB binding site in their promoter. No significant differences in the expression of these two genes were observed by RT-dPCR in the presence and absence of RctB (Supplementary Fig. 3c-d). Taken together, these results show that the newly identified RctB binding sites have no function in controlling Chr2 replication. Moreover, except for its own promoter, RctB cannot be considered as a transcription factor as it is the case for DnaA ³².

To examine the binding profile of RctB when cells are not replicating, we performed the ChIPseq of RctB in a stationary phase culture of *V. cholerae*. The binding profile of RctB to *ori2* was found to be the same as in exponential phase (Supplementary Fig. 4a), suggesting that inhibition of initiation at *ori2* by RctB binding to 29/39mer is carried out in both replicative and non-replicative phases. RctB thus behaves differently from DnaA, which detaches from *oriC* in stationary phase ³³. Similarly, RctB is bound to *crtS* even in stationary phase indicating that RctB is not recruited to *crtS* through the passage of the replication fork (Supplementary Fig. 4b).

Following *crtS* replication, RctB binding at *ori2* shifts towards replication activation sites

Based on the above results, we hypothesized that in order to activate replication at *ori2*, the binding of RctB to iterons must be transient, while the rest of the cell cycle, RctB is bound to the inhibitory 29/39mer sites. Our previous approach did not allow us to capture brief events. Therefore, we implemented the ChIPseq of RctB in a replication-synchronized population to monitor RctB binding profile over the course of the cell cycle. As for *E. coli*, an exponentially growing population of *V. cholerae* can be synchronized via induction of the stringent response by adding a serine analog (serine hydroxamate, SHX) to the growth medium ^{11,34,35}. Upon

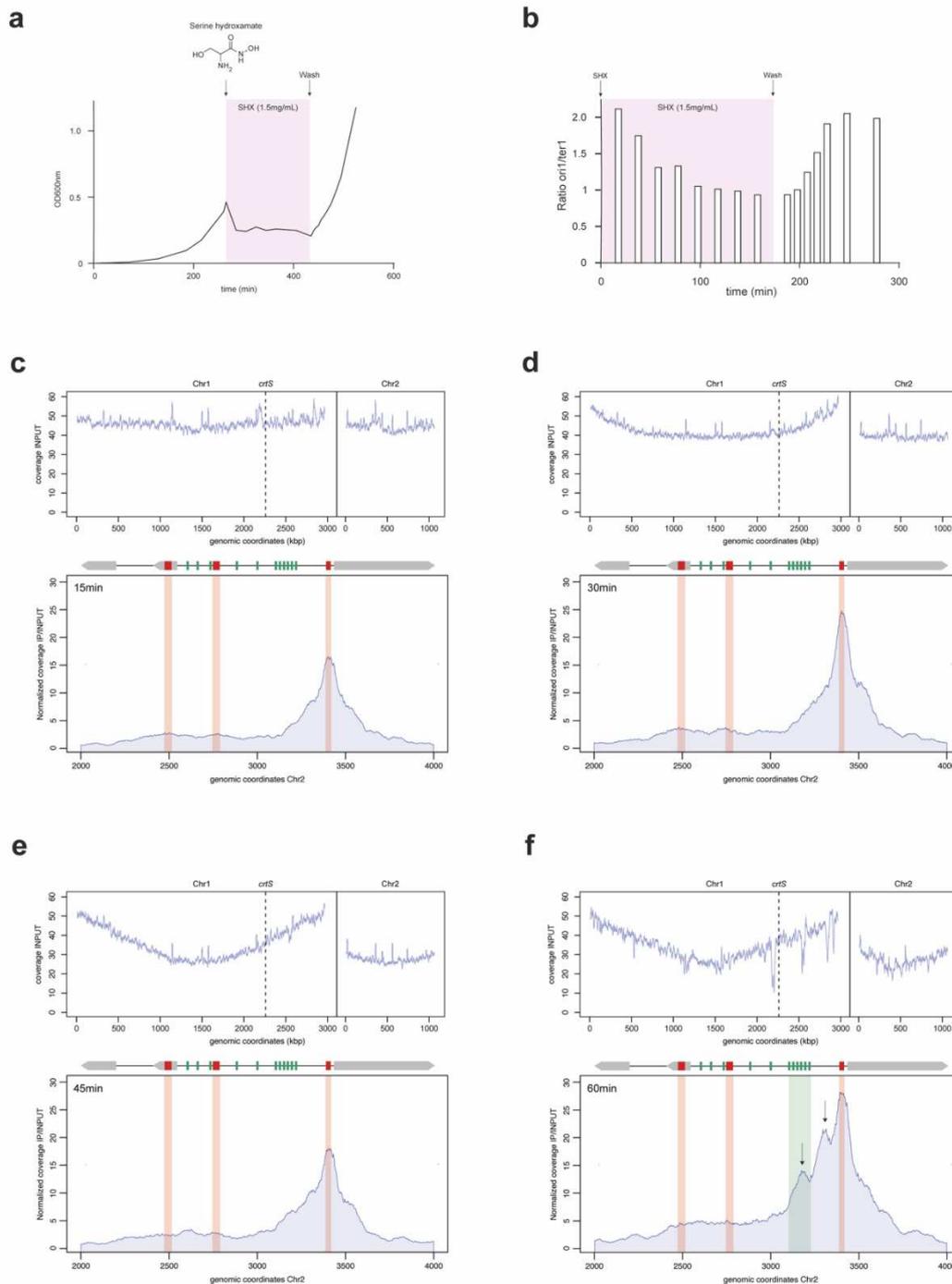


Figure 2: RctB binding pattern on *ori2* shift toward iterons after *crtS* replication. **a.** *Vibrio cholerae* growth curve in M9 glucose 0.2% casamino acids 0.4% 30°C. At OD₆₀₀=0.5 DL-serine hydroxamate (SHX) is added to the media at 1.5mg/mL final concentration. After 120min of SHX treatment, cells are pelleted and resuspend in a fresh M9 media. **b.** *ori1/ter1* measurement by digital droplet PCR (ddPCR) on samples from a growing *Vibrio cholerae* culture at different time points during SHX treatment and after SHX removal. **c, d, e, f.** ChIPseq experiments on synchronized population at different time points after SHX removal (R+15min, R+30min, R+45min, R+60min). Top panel: Marker frequency analysis on the input sample (not immunoprecipitated) from ChIP experiments, coverage is represented on y axis and genomic coordinates of both chromosomes (in kilo base pair) are present on the x axis. Curves have been smoothed with a 10kb sliding window along the genome. Bottom panel: on y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference Chr2 (CP028828.1). A 2.000bp window is represented. Genetic context of *ori2* and known RctB binding sites are depicted on top of the graph (genes in grey arrows, iterons in green squares and 39mers in red squares).

treatment with SHX the replication forks continue to process DNA but re-initiation is impossible because of the specific degradation of DnaA-ADP by the Lon protease³⁶. Once their cell cycle is completed, the bacteria stalled in G1 phase are transferred to a medium without SHX so that they theoretically restart their replication in synchrony. In *V. cholerae*, after treatment with SHX, it has been shown that Chr1 re-initiates first and then Chr2¹¹ presumably due to control by *crtS*¹⁰. In our hands, addition of SHX at 1.5mg/mL final concentration to an exponential culture of *V. cholerae* caused an instantaneous growth arrest (Fig. 2a). To ensure completion of ongoing replication, we measured the ratio of loci at the origin and terminus of Chr1 (*ori1/ter1*) by dPCR (Fig. 2b). After 120min of SHX treatment, all cells successfully completed replication with a ratio of *ori1/ter1* = 1 meaning that Chr1 is fully replicated. After washing, the cells resumed replication within a time window of 20-30 minutes, similar to³⁵. The synchronization by SHX in *V. cholerae* is therefore not absolutely perfect as the re-start window is quite long. In the absence of a better method to synchronize *V. cholerae* replication, we decided to use this synchronization method to perform ChIP-seq of RctB at different time points (15, 30, 45, 60 minutes) after removal of SHX. Using input DNA reads (control without immunoprecipitation), we performed a marker frequency analysis (MFA) to track replication fork progression on both chromosomes (Methods) (Fig.2c-2f, top panel). At 15 minutes after washing (Fig.2c), replication has not yet restarted, as shown by the flat MFA profile of Chr1 and Chr2. At *ori2*, RctB is predominantly bound to 29/39mer. After 30 minutes (Fig.2d), Chr1 starts replicating and at 45 minutes (Fig.2e) the *crtS* site seems to be just replicated (the Ter region of Chr1 still appears quite flat), while Chr2 replication has not yet started. At both time points, the ChIP-seq pattern at *ori2* is nearly unchanged (Fig. 2d-2e). At 60 minutes (Fig.2f), Chr2 is finally replicating. On the ChIP-seq profile, we observe two additional peaks: one on the six-iteron array and one on the DUE (Fig.2f, black arrows). Although the signal is increasingly fuzzy due to a desynchronization of the replication within the population, we were able to capture a release of RctB from the inhibitory sites (29/39m) towards the activating sites (iteron) and towards the DUE indicating the opening of *ori2*. This is the first time we have been able to observe a temporal shift in the binding mode of RctB (from inhibitory to activating sites) at *ori2* resulting in the initiation of Chr2 replication. As expected, *ori2* activation follows *crtS* replication on Chr1. Importantly, we did not observe any fluctuation in the binding of RctB at *crtS* before and after the passage of the replication fork (Supplementary Fig. 5).

***crtS* acts at *ori2* by reducing the binding of RctB to inhibitory sites.**

crtS is an enhancer of Chr2 replication. In a Δ *crtS* strain, Chr2 replication occurs rarely and randomly, causing cell filamentation¹⁰ and the overall Chr2 copy number is much lower than Chr1 (Supplementary Fig. 6). The mechanism behind the enhancing effect of *crtS* on *ori2*

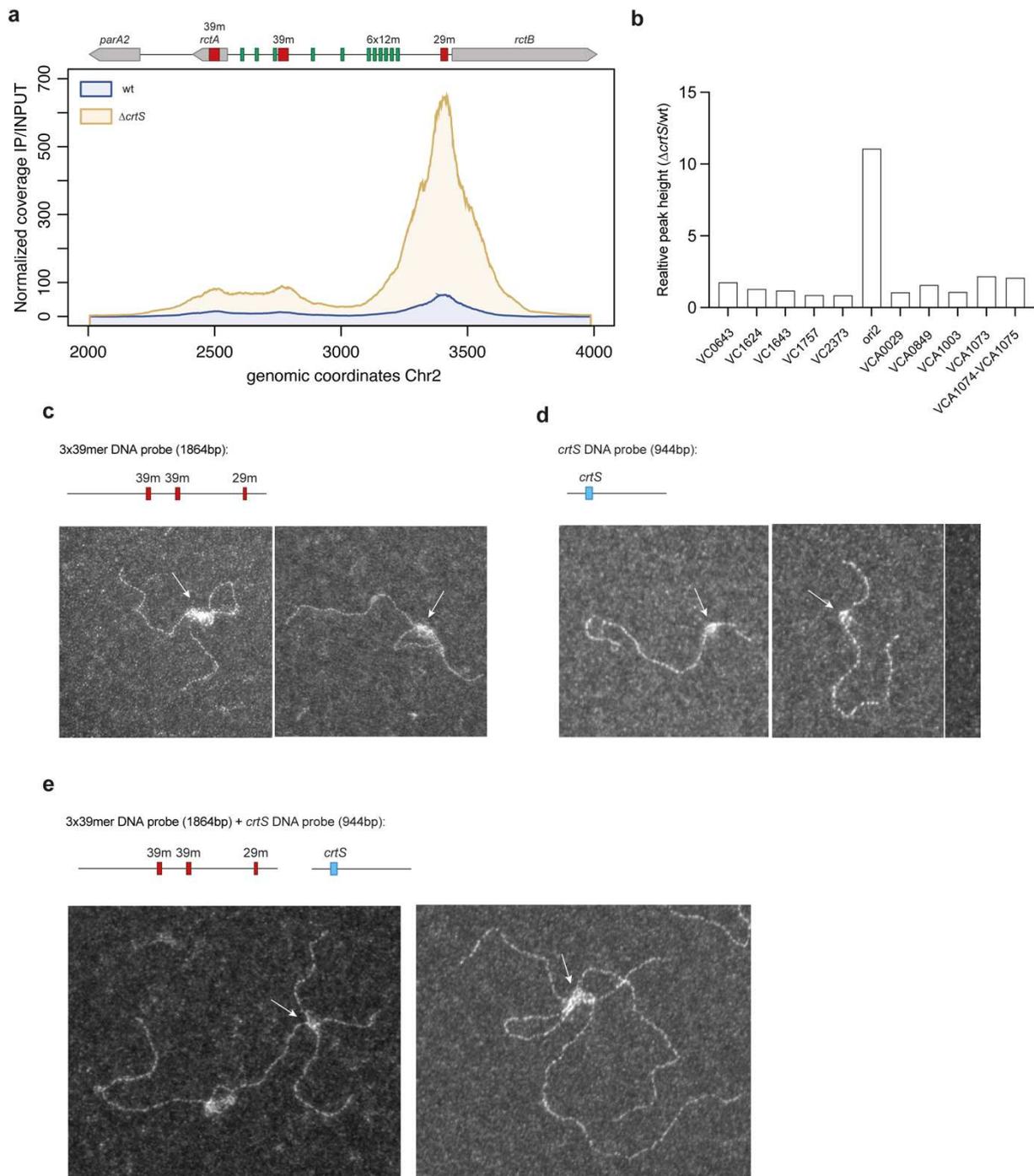


Figure 3: *crtS* enhanced *ori2* initiation by diminishing inhibitory sites fixation. **a.** on y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates of Chr2 from *Vibrio cholerae* N16961 reference genome (CP028828.1). A 2.000bp window is represented. The genetic context of *ori2* and known RctB binding sites are depicted on top of the graph. wt ChIP signal is in blue, $\Delta crtS$ ChIP signal is in sand. **b.** Barplot representing the maximum height of RctB ChIP peaks along the genome (logarithmic scale). Peak maximal height was calculated from the average of two independent ChIP experiments, for *ori2* the peak height represented is the one on the 29mer. **c.** TEM observations of nucleoprotein complexes form by an *ori2* DNA probe containing three 39m sites incubated with RctB. RctB forms loops inside *ori2* probes indicated with white arrow. **d.** TEM on *crtS* containing probe, RctB binding is punctual and indicated with white arrows. **E.** TEM on *crtS* and 39m containing probes mixed with RctB. Some contacts mediated by RctB between the two molecules are observed (white arrows).

remains poorly understood. In a previous study, we observed that by impeding 29/39mer inhibitory activity, Chr2 replication became independent of *crtS*, suggesting that *crtS* must act as an anti-inhibitor of 29/39mer sites²¹. Here, we performed RctB ChIPseq in a mutant deleted for *crtS* (Δ *crtS*) to observe the effect on RctB whole genome binding (Fig. 3a). As in the wild-type (wt), in the Δ *crtS* mutant, RctB binds predominantly to the 29/39mer inhibitory sites at *ori2*. However, the signal increases strikingly by about 10-fold, suggesting a much stronger inhibition of replication in the absence of *crtS*. This is consistent with the Chr2 under-initiation phenotype observed in Δ *crtS* mutants. Thus the role of *crtS* would be to reduce the binding of RctB to 29/39mer. As we have seen that RctB binds to other 39mer-like sites outside of *ori2* (Fig. 1b), we wondered whether in the absence of *crtS*, RctB binding to these sites was also increased. Remarkably, no increase in RctB binding to other sites, as seen at *ori2*, was observed in the Δ *crtS* mutant. (Fig. 3b). This suggests that *crtS* acts specifically at *ori2* by reducing the binding of RctB to the inhibitory sites. This result is consistent with the high frequency of contacts observed between *crtS* and *ori2* by Hi-C¹⁰.

Then we wanted to directly visualize the inhibition complex formed by RctB bound to 39m on *ori2*. We generated an *ori2* DNA fragment which is not methylated in which iterons only 39m are bound by RctB. This DNA probe was then mixed with RctB pre-treated with DnaK/J chaperones and observed by electron microscopy (TEM). We observe that with only three 39m sites, RctB can form loops within the origin. We think that those loops are the inhibitory complex that block *ori2* initiation. We also generated a *crtS* containing probe and observed RctB binding on it. RctB binding on *crtS* is punctual, RctB seems to be able to introduce a kink in the DNA. Then we mixed the two probes together and incubated them with RctB. We detected few contacts between 39m and *crtS* mediated by RctB, which confirm the possibility of a direct action of *crtS* on *ori2* through DNA bridging.

The replication of one *crtS* site triggers the initiation of one *ori2*

To understand how *crtS* replication leads to *ori2* activation, we tested whether an increase in *crtS* gene dosage is sufficient to trigger Chr2 initiation. As observed previously^{10,28,37}, the synthetic addition of a second chromosomal copy of *crtS* results in an increase in Chr2 copy number. This led to speculation that the increase in *crtS* gene dosage might be what triggers Chr2 replication³⁷. According to this model, a single *crtS* would have a significant but insufficient activity for *ori2* activation and the doubling of its copy number (by replication) would raise its activity above the threshold required to initiate Chr2 replication³⁷. If so, in a *V. cholerae* mutant containing two *crtS* sites, we would expect to have an altered RctB binding profile at *ori2* compared with the wild type. We therefore performed the ChIP-seq of RctB in a strain

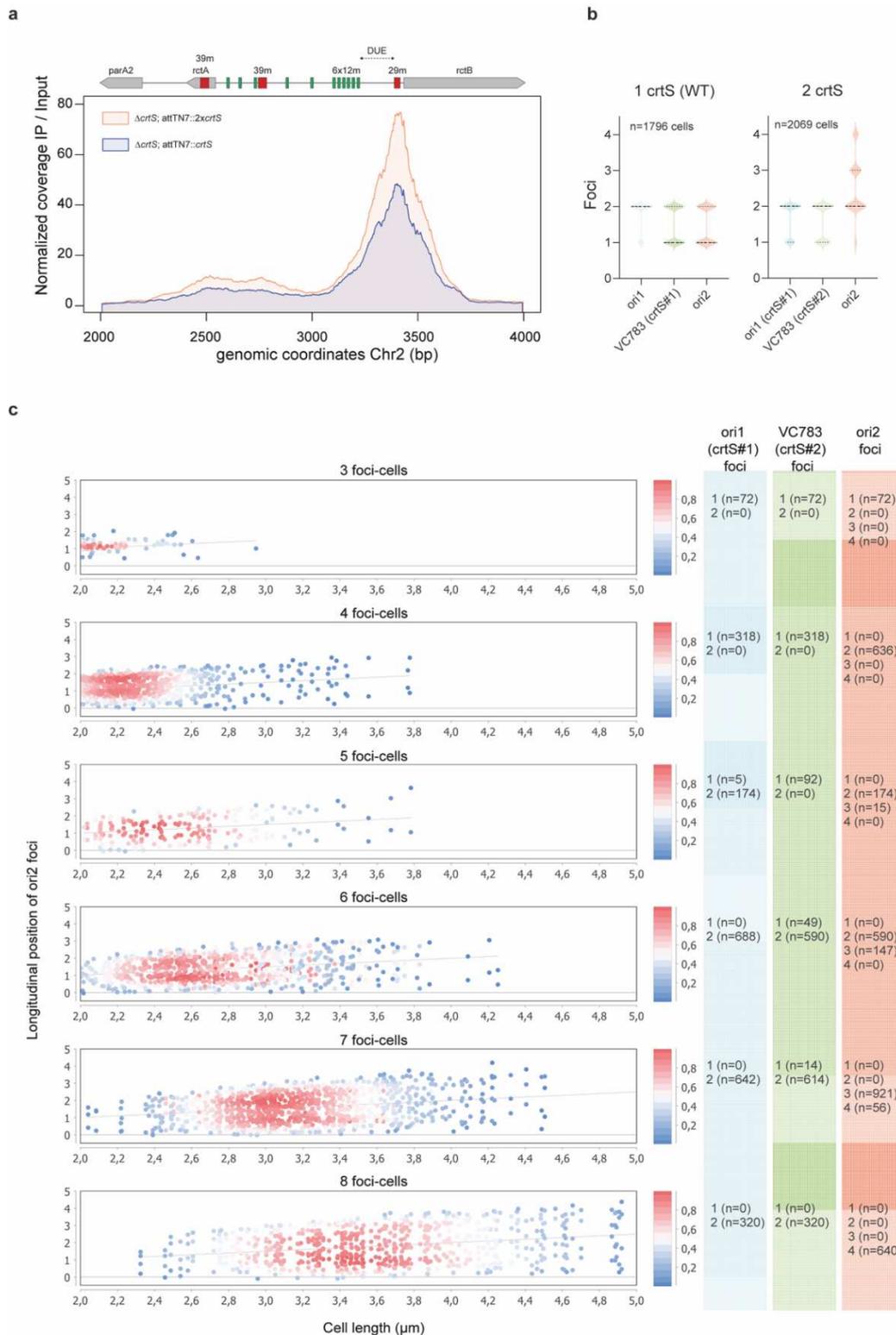


Figure 4: The replication of one *crtS* site triggers the initiation of one *ori2*. **a.** On y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference genome (CP028828.1). A 2.000bp window is represented. ChIP signal is represented in yellow for the strain carrying a single copy of *crtS* (*attTN7::crtS*) and in orange for the strain carrying two copies of *crtS* (*attTN7::2xcrtS*). Genetic context and known RctB binding sites are depicted on top of the graph: genes as grey arrow, 39mer site as red box, iterons a green triangle, *crtS* as cyan box. The AT-rich region also called DNA unwinding element (DUE) is also depicted. **b.** Epi-fluorescent microscopy on two *V. cholerae* strains two *crtS* using three different fluorescent proteins : ParB_{P1}-CFP (*ori1*), ParB_{MT1}-yGFP (native *crtS*), lacl-RFP (*ori2*) (see Methods)³⁸.

containing either one or two chromosomal copies of *crtS* inserted in the Chr1 native attTn7 site²⁸. As expected, we observed that Chr2 copy number increases with two copies of *crtS* (Supplementary Fig. 7). However, we observed no significant difference in the binding profile of RctB between a strain with one or two copies of *crtS*, neither the binding to 29/39m decreased nor the binding to iterons increased at *ori2* (Fig. 4a). In both cases, RctB binds predominantly to the inhibitory 29/39m sites. This suggests that a doubling of the *crtS* gene dosage is not what causes *ori2* initiation. To investigate this further, we observed at the single cell level using fluorescence microscopy, a *V. cholerae* mutant in which a second *crtS* site was added next to *ori1* so as to temporally offset the replication timing of the two *crtS* sites. In this mutant, we had previously observed that the majority of cells are born with two *ori2* and have four *ori2* before cell division¹⁰. We also observed that a small portion of the population had one or three *ori2* foci suggesting an asynchrony in Chr2 initiation or an asymmetry in cell division. To further extend this study, we followed the dynamics of both *crtS* sites and of the *ori2* using three different fluorescent proteins : ParB_{P1}-CFP (*ori1*), ParB_{MT1}-yGFP (native *crtS*), lacI-RFP (*ori2*) (see Methods)³⁸. Compared to the wild type, which contains only one or two foci of *ori1*, *crtS*, and *ori2*, in the mutant having two *crtS* sites, the cells could have up to four *ori2* foci (Fig. 4b). We tracked the longitudinal position of *ori2* foci relative to the two *crtS* sites (near *ori1* and near VC783) in growing cells (Fig. 4c). Apart from a minor portion (n=72) of cells with 3 total foci (probably resulting from asymmetric divisions), most cells are born with 4 foci (1 *ori1*, 1 VC783, 2 *ori2*). As cell size increases, there is a sequential increase in the number of *ori2* foci with three and then four foci. These results suggest that the replication of the first *crtS* (near *ori1*) activates the replication of one of the two *ori2* and that the replication of the second *crtS* (near VC783) activates the initiation of the other *ori2*. These results are consistent with a model in which the replicated *crtS* site would be able to activate a single *ori2*, possibly by direct interaction.

RctB domain IV harbors two dimerization interfaces involved in *ori2* replication inhibition

We recently showed that the C-terminal tail of RctB domain IV is crucial for *crtS* mediated control of Chr2 replication²¹. RctB domain IV, which is not essential for replication, is however critical to down-regulate Chr2 copy number through its involvement in handcuffing activities and in 29/39m binding^{20,21,39}. In vitro, RctB is also able to oligomerize non-specifically onto adjacent DNA through domain IV-mediated interactions^{21,28}. Thus, we speculate that the coordination of Chr1 and Chr2 replication most likely relies on the ability of RctB molecules to interact via their domain IV. However, the full-length structure of RctB has remained refringent to structural studies, likely due to the presence of intrinsically unstructured regions in the protein. The use of truncated versions of RctB, domain I alone and domains II-III together^{19,23}

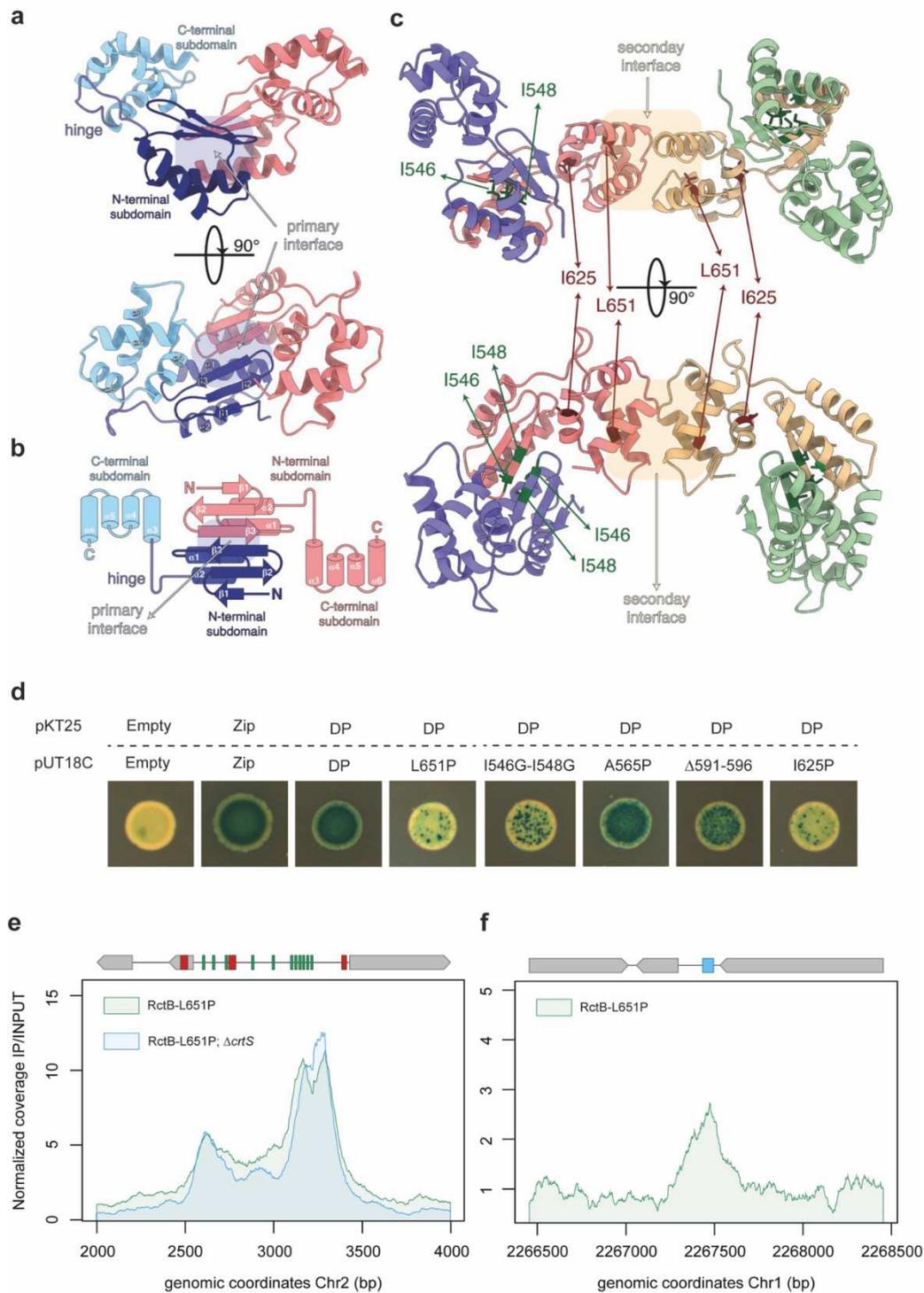


Figure 5: RctB domain IV structure reveals two dimerization interface involved in 39m-mediated inhibition
a. Crystal structure of the RctB domain IV dimer. Each monomer is composed of an $\alpha\beta$ N-terminal subdomain (shown in dark blue) connected to a four α -helices bundle (in light blue) via a hinge region. The primary dimer interface is formed by the extension of the central β -sheet of each monomer through $\beta 3$ and the antiparallel interactions of the stacking $\alpha 1$. Important residues contributing to the interface are coloured in dark green. **b.** Topological representation of the domain IV dimer. **c.** Crystallographic tetramer formed through lattice contacts mediated by the C-terminal subdomain of neighboring dimers. This secondary interface is shadowed in light yellow with the I625 and L651 functionally relevant residues shown in dark red. **d.** Bacterial-Two-hybrid of RctB D314P (DP) vs. RctB domain IV mutants. Empty (negative control), Zip (positive control). **e, f.** on y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference genome (CP028827.1 and CP028828.1). A 2.000bp window is represented and centered around the peak. Genetic context and known RctB binding sites are depicted on top of the graph. RctB-L651P ChIP signal is represented in green, RctBL651P Δ crtS ChIP signal is represented in light blue.

facilitated structure determination of these regions. Yet the structure of domain IV has remained unknown.

In this context, we focus on generating a stable fragment of domain IV that could be used in structural biology resulting in the crystallization and structure determination of a fragment of domain IV encompassing amino acids 532 to 658. RctB domain IV is organized into two structural subdomains connected by an extended linker or hinge region. The N-terminal subdomain consists of a 3-stranded β -sheet stacking two α -helices and the C-terminal subdomain has 4 α -helices in a bundle (Fig. 5a). In the crystal, Domain IV retains a dimeric association predicted by Size Exclusion Chromatography (SEC). This primary interaction involves an interface of 392.4 \AA^2 and is mediated by a β -sheet (β 3) from residue 545 to 549 (Fig. 5a-b). Interestingly, the analysis of the crystal lattice contacts of the C-terminal subdomain suggests a potential secondary interaction interface involving 315.4 \AA^2 , which could mediate further oligomerization via this region (Fig. 5c). Indeed, a DALI search identified the C-terminal subdomain as a protein recognition domain that could mediate the formation of transient homodimeric interfaces, with Z-scores between 3.7 and 5.4. This secondary interface recreates a similar interface to that of a representative structural homologue detected by DALI (a putative transcription regulator from *Neisseria gonorrhoeae*)⁴⁰. This suggests that the domain IV could potentially associate in more than one way or even form higher oligomeric structures when recruited to the DNA.

To further characterize these interfaces, we used the bacterial two-hybrid assay in *E. coli*⁴¹. We performed a structure guided point mutation approach aiming at probing potential interfaces involved in oligomerization of domain IV. Since RctB has another dimerization interface in domain II¹⁹, we mutated residue D314 with a proline (DP) to prevent domain II-mediated interactions that would hide domain IV-mediated interactions. In Fig. 5d, blue colonies (DP/DP) indicate that RctB can interact with other RctB through another domain, as already reported to be via domain IV²¹. Mutations in the β 3 sheet cause the formation of white colonies (DP/I546G-I548G), indicating that the β 3 sheet is important for RctB interactions while substitutions that destabilize the α 1 helix appear to have little impact on dimerization as evidenced by the formation of blue colonies (DP/A565P). Deletion of the linker connecting the two subdomains (591-596), gives blue colonies (DP/ Δ 591-596) but to a lesser extent than the A565P substitution, suggesting that RctB self-association may be sensitive to local structural rearrangements between the two subdomains. The I625P and L651P substitutions in two different α -helices (α 3 and α 6) of the C-terminal subdomain also abrogate RctB interactions confirming our previous results²¹. Thus, RctB molecules can interact with each other in at least two different areas of domain IV: at a primary interface formed between the β 3 sheets in an antiparallel arrangement and at the secondary interface between α -helices in a parallel arrangement. In summary, RctB domain IV could be an anchor not only for dimerization but

also for the formation of transient oligomeric structures when bound to DNA or even to bridge two chromosomal regions.

We were able to introduce the RctB-L651P mutation in *V. cholerae* as it allows Chr2 replication in the absence of *crtS* while causing only a moderate 2-fold increase in its copy number, which is not too detrimental to the bacteria²¹. To better understand the role of domain IV-mediated interactions on RctB genome-wide binding as well as their importance for *crtS* activity, we performed the ChIP-seq of RctB-L651P in a +/- *crtS* genetic background. As expected, no difference in the replication profile is observed with or without *crtS* when RctB carries the L651P mutation (Supplementary Fig. 8a). Compared to RctB wt (Fig. 1c), the ChIP-seq of RctB-L651P adopts a completely different binding profile at *ori2*, regardless of the presence or absence of *crtS* (Fig. 5e). RctB-L651P binding to *ori2* is fully shifted to the iterons array and the DUE and not at all on the 29/39m sites. This shift from inhibitory to activating sites explains why Chr2 replication can initiate independently of *crtS* in the RctB-L651P mutant. Furthermore, we observe the same binding profile at *crtS* between RctB-L651P (Fig. 5f) and RctBwt (Fig. 1d). Thus, although RctB-L651P binds to *crtS*, there is no change in RctB activity at *ori2* in the presence or absence of *crtS*. Note that only 39mer-like sites (Fig. 1b) are not bound by the RctB-L651P mutant (Supplementary Fig. 8b), suggesting that, although lacking direct DNA-binding activity, domain IV of RctB may mediate differential recognition of 29/39mer inhibitory sites relative to iterons or *crtS* via the formation of alternative oligomeric structures. We speculate that domain IV-mediated interactions between RctB are crucial for the inhibition of replication at *ori2* through the 29/39m sites and that *crtS* triggers Chr2 initiation by destabilizing these domain IV interactions.

We believe that local changes after replication of *crtS* are likely to affect the structure or stability of domain IV and the relative orientation of the DNA binding motifs leading to changes in the DNA recognition pattern

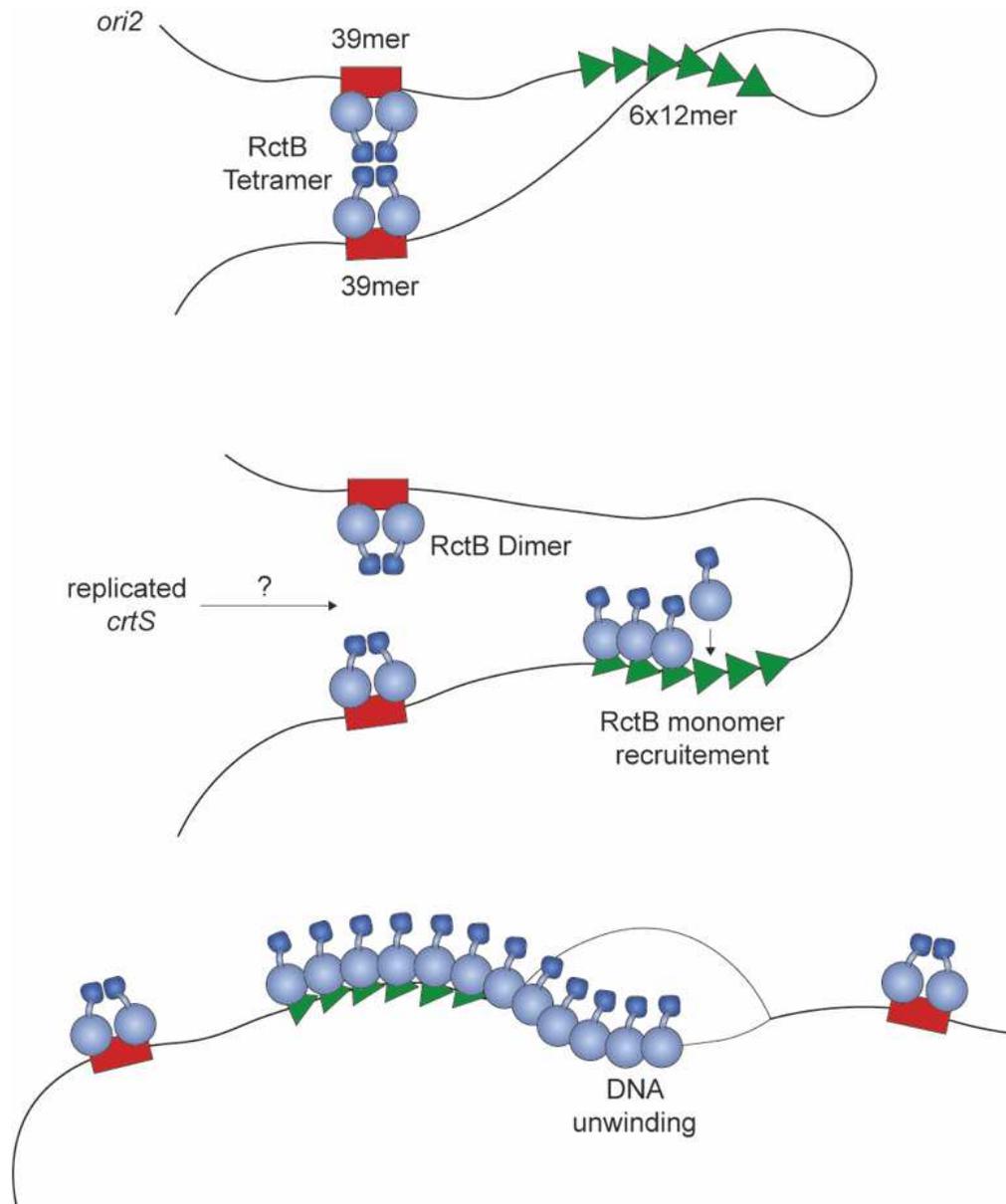


Figure 6. Model of *crtS*-mediated *ori2* firing. In this model, RctB only binds to 39m sites on *ori2* before *crtS* replication forming an inhibitory complex, probably through its domain IV. RctB binds to 39m as a dimer, interacting through its domain IV. Thus domain IV has a double role, stable binding to 39m and handcuffing. The replication of *crtS* acts on domain IV, maybe through direct interaction between *crtS*-bound RctB and 39m-bound RctB. This event unlocks the inhibitory complex and allows the recruitment of RctB on iterons which will lead to DNA unwinding.

DISCUSSION

High-resolution ChIPseq study in different genetic backgrounds allowed us to dissect the *ori2* regulatory network. We propose a cell cycle integrated model for the initiation of Chr2 (Fig. 6). In this model, RctB binds to *ori2* only at the three 39mer sites thereby inhibiting replication initiation during most of the cell cycle. RctB binds to 39mer sites as a homodimer interacting via their domain IV. This domain IV- domain IV interaction is essential for stable attachment to the 39mers. When the *crtS* site is duplicated, it acts locally on *ori2* and disrupts this domain IV domain IV interaction of RctBs attached to 39mer. Thus, the inhibitory complex is destabilized exposing the *ori2* for a short time. RctB is then recruited to the 6x-iteron sites, oligomerizing along the DUE and recruiting the replisome to initiate Chr2 replication.

METHODS

High Resolution ChIP-seq

ChIP-seq of RctB was performed as described in ⁴² in *V. cholerae* expressing RctB-3xFLAG under its native promoter. We confirmed that 3xFLAG did not impair RctB function by digital PCR (dPCR) measurement of Chr2 copy number (Supplementary Fig.1). Cells were grown in 100mL Mueller Hinton (MH) media with shaking at 37°C until they reach mid-log growth phase (OD₆₀₀ 0.5). Then cells were fixed by crosslinking with 1% formaldehyde for 30 min at room temperature under agitation, followed by quenching with 0,5 M glycine for 15 min. Bacteria were harvested, washed two times in ice-cold Tris-Buffered saline (20 mM Tris/HCl pH 7.5, 150 mM NaCl) and pellets were stored at -80°C. Pellets were lysed in buffer containing lysozyme and protease inhibitor pill (cOmplete Protease Inhibitor, Roche). Chromatin was then fragmented by sonication (Covaris S220) for in milliTUBE (1ml with AFA Fiber). Correct DNA fragmentation centered around 300bp was confirmed using 1.8% agarose gel electrophoresis. Following sonication, RctB-DNA complexes were immunoprecipitated with anti-FLAG magnetic beads (M8823 Sigma). Correct immunoprecipitation was confirmed by western blot using anti-FLAG antibodies (F7425 Sigma). Once recovered, immunoprecipitated DNA was de-crosslink at 65°C overnight. Then DNA was blunt-ended, A-tailed, ligated to adaptors, amplified by PCR and purified using TruSeq ChIP Library Preparation Kit (illumina) and AMPureXP beads (Beckman Coulter). ChIP-seq libraries quality was measured by Agilent Bioanalyzer 2100 instrument using a DNA High Sensitivity chip. Finally, libraries were sequenced using a MiniSeq Mid Output Kit on a Miniseq sequencing machine (Illumina). Paired-end reads from two independent RctB ChIP-seq experiments were mapped to *Vibrio cholerae* N16961 reference genome (chromosome 1: CP028827.1 and chromosome 2: CP028828.1) with Bowtie2 using galaxy.pasteur platform. Peak calling was performed using

MACS2, where aligned reads were significantly enriched compared to a control (strain without FLAG). Finally, identified peaks were checked by hand, using IGV genome browser to confirm correct peak assignment. Average read counts were normalized according to the coverage per base for each experiment. MEME-ChIP (version 5.4.1) online browser was used for motif enrichment finding.

Marker Frequency Analysis

Using ChIP-seq input data, we performed marker frequency analysis (MFA) to follow the replication fork progression on both chromosomes. Like most bacterial chromosomes, *Vibrio cholerae* initiates DNA replication at the *ori* site by recruiting two replication complexes (replisomes). The chromosome is then bidirectionally replicated by these replisomes toward the terminus, *ter*. This causes a copy number gradient in an exponentially growing population, making origin-proximal sequences more prevalent than origin-distal sequences⁴³.

Digital PCR quantification

Quantifications of (*ori1* and *ori2*) in *V. cholerae* and (*oriC* and *pORI2*) in *E. coli* were performed as described in²¹ using multiplex dPCR (Stilla Technologies). dPCR was performed directly on cell lysate: 1mL of overnight culture was washed twice with 1mL PBS, pellets were then frozen at -20°C and resuspended in 200uL PBS, samples were then boiled for 10min. Samples were centrifuge for 10min, supernatant was recovered and DNA content was measured using Qubit device (ThermoFisher). PCR reactions were performed with 0.1 ng of DNA using the PerfeCTa® MultiPlex qPCR ToughMix® (Quantabio) on a Sapphire chip (Stilla Technologies). Digital PCR was conducted on a Naica Geode and Image acquisition on the Naica Prism3 reader. Images were analyzed Crystal Miner software (Stilla Technologies). The dPCR run was performed using the following steps: droplet partition (40°C, atmospheric pressure AP to + 950mbar, 12 min), initial denaturation (95°C, +950 mbar, for 2 min), followed by 45 cycles at (95°C for 10 s and 60°C for 30 s), droplet release (down 25°C, down to AP, 33 min).

Digital PCR quantification of RNA (RT-dPCR)

Quantification of RctB mRNA was performed in *V. cholerae* from exponentially growing cultures (OD600 0.5) using multiplex digital RT-dPCR (Stilla Technologies)⁴⁴. Primers and probes are listed in Supplementary Table 3. *V. cholerae* RNA was prepared as described in⁴⁵. PCR reactions were performed with 1 ng of RNA using the qScript™ XLT One-Step RT-qPCR ToughMix® (Quantabio). RT-dPCR and data analysis were conducted as described above for dPCR. The RT-dPCR run was performed in the following steps: droplet partition (40°C, AP to +950 mbar, 12 min), cDNA synthesis (50°C, +950 mbar, 10 min), initial denaturation (95°C, +950 mbar, for 1 min), followed by 45 cycles at (95°C for 10 s and 60°C for 15 s), droplet release (down 25°C, down to AP, 33 min). Expression values were normalized to the

expression of the housekeeping gene *gyrA* as described in ⁴⁵. More details are provided in Supplementary Methods.

SHX *Vibrio cholerae* synchronization

Cells were grown in 40mL of M9 glucose 0.2% casamino acids 0.4% media at 30°C. When cells reached $OD_{600nm}=0.5$, they were treated with serine hydroxamate (DL-serine hydroxamate, Sigma) at 1.5mg/mL final concentration for 120min. Cells were centrifuged at 6.000g for 15min, pellets were resuspended in 40ml M9 fresh media and but back to grow at 30°C. At several time points cells were fixed with 1% formaldehyde: 15, 30, 45 and 60min after the wash. Samples were then treated with the ChIP-seq protocol.

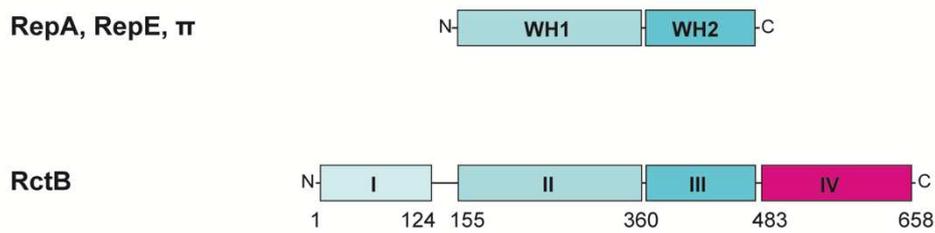
Transmission electronic microscopy

For observation of DNA-protein complexes, 2 μ L of the remaining 7 μ L of reaction were rapidly diluted 120-fold in buffer containing 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 3 mM MgCl₂ and observed by electron microcopy (DNA-protein samples).

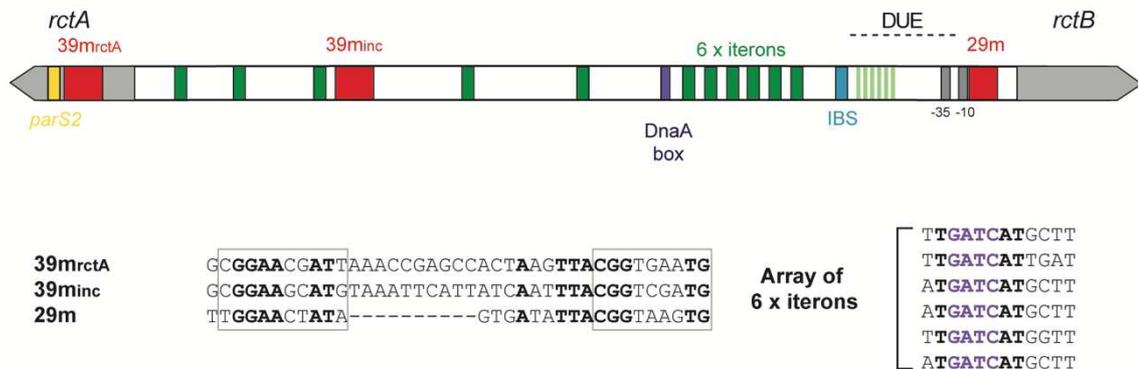
In both cases, for 1 min, a 5 μ L drop of the dilution was deposited on a 600-mesh copper grid previously covered with a thin carbon film and preactivated by glow discharge in the presence of amylamine (Sigma-Aldrich, France) 66,67. The grids were rinsed and positively stained with 2% (w/v) aqueous uranyl acetate, carefully dried with filter paper, and observed in annular dark-field mode in lossless filtered imaging using a Zeiss 902 transmission electron microscope. Images were captured at 85,000 \times magnification with a Veleta MegaviewIII CCD camera and analyzed with iTEM software (Olympus Soft Imaging Solution).

SUPPLEMENTARY FIGURES AND LEGENDS

a



b

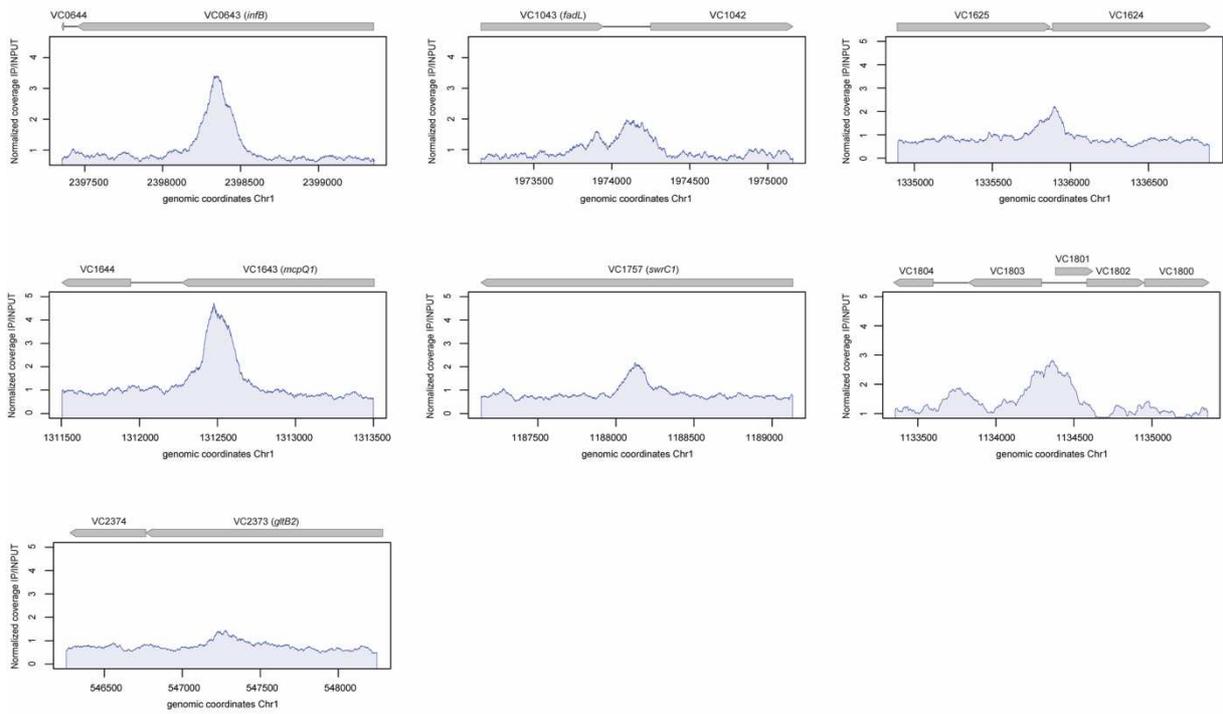


Supplementary Figure 1

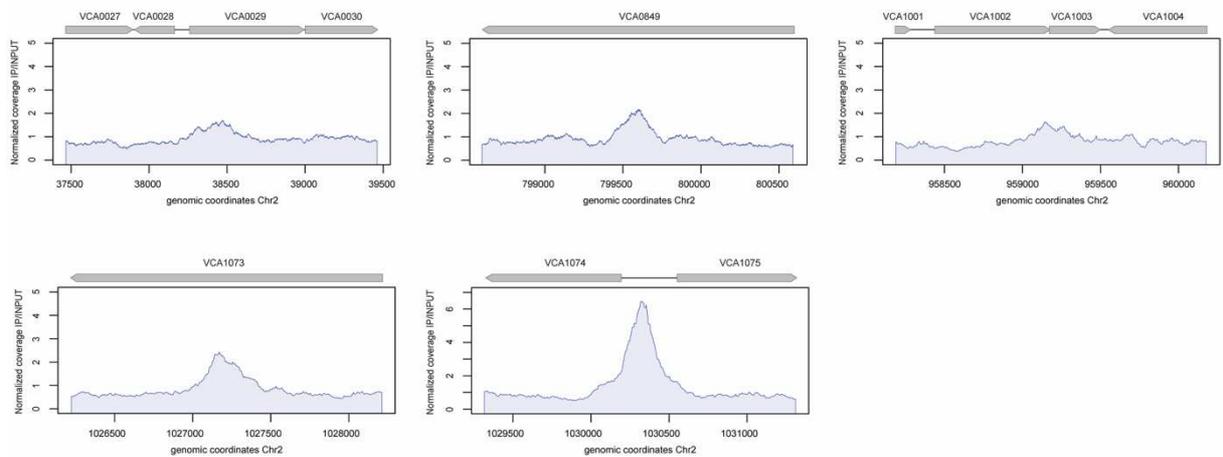
a. Comparison of protein domain organization between Chr2 initiator (RctB) and iteron-type plasmid initiators (RepA, RepE, π). The different structural domains are indicated (WH = winged helix).

b. Representation of Chr2 origin of replication (*ori2*) and sequence alignment of iterons and 39m sites. RctB binding sites are indicated by green (iterons) and red (39m) rectangles. Eleven iterons are present in *ori2*, six of which are essential for replication, while the others have a regulatory function. The three 29m/39m sites act as strong negative regulators of *ori2* initiation. The single-stranded binding sites of RctB (ATCA) in the DUE (DNA unwinding element) are represented as light green bars. Other DNA binding sites are indicated: IBS (IHF binding site), DnaA box, *parS2* (ParB2 binding site). In the sequence alignment, conserved bases are indicated in bold. On the iterons, Dam-methylated GATC motifs are shown in purple. The imperfect GC-rich repeats are RctB binding sites are outlined in a gray rectangle.

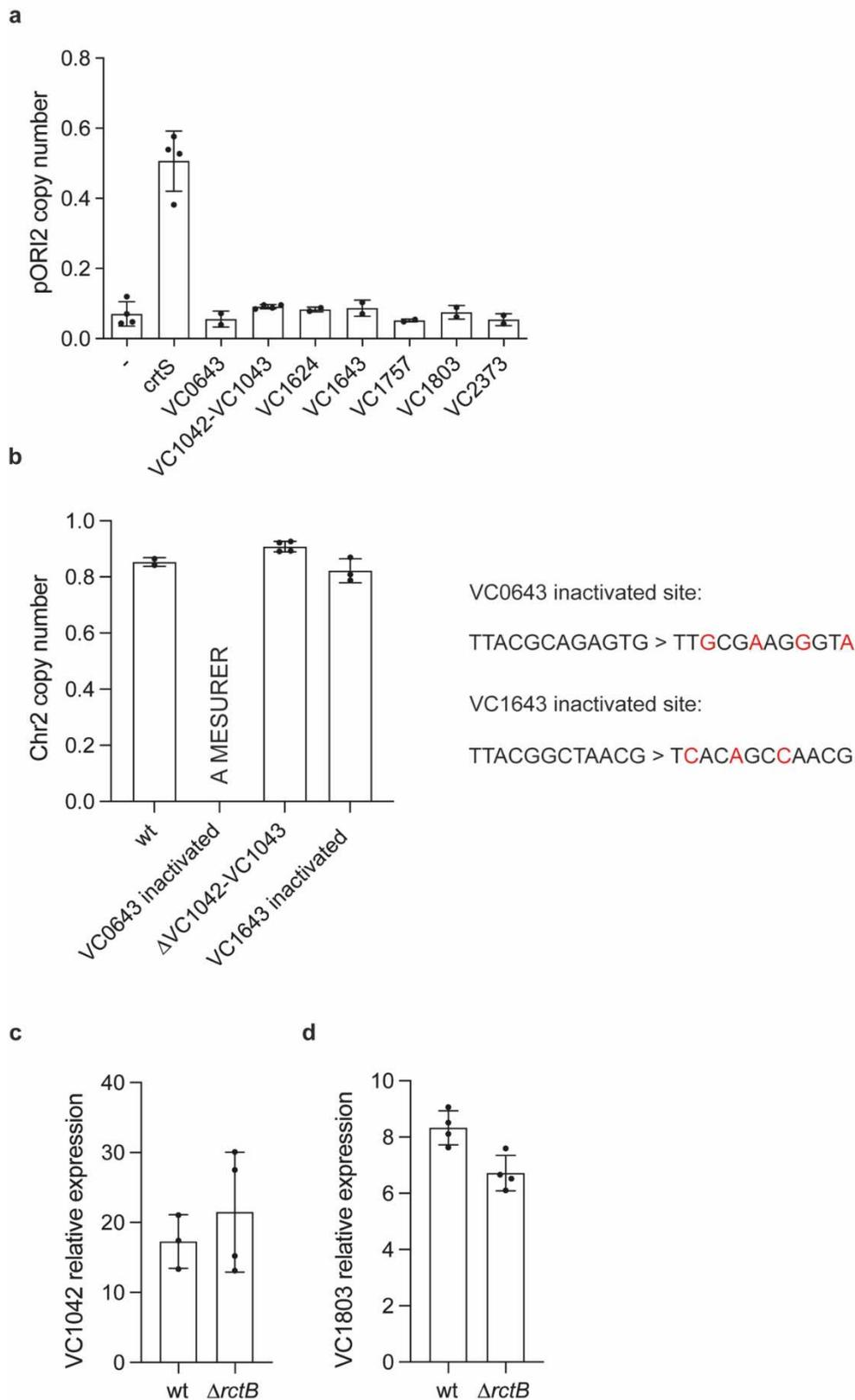
Peaks on Chr1



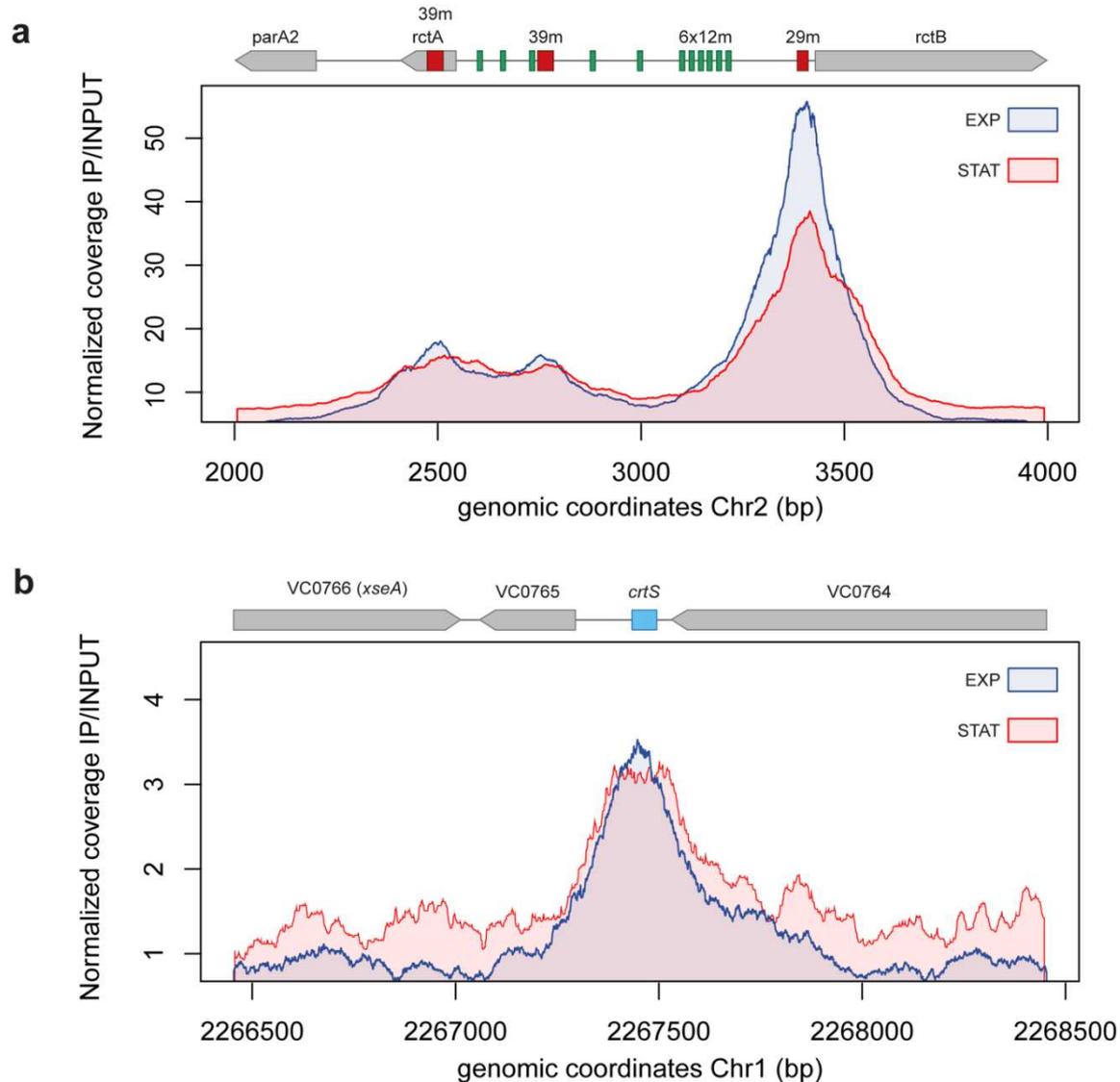
Peaks on Chr2



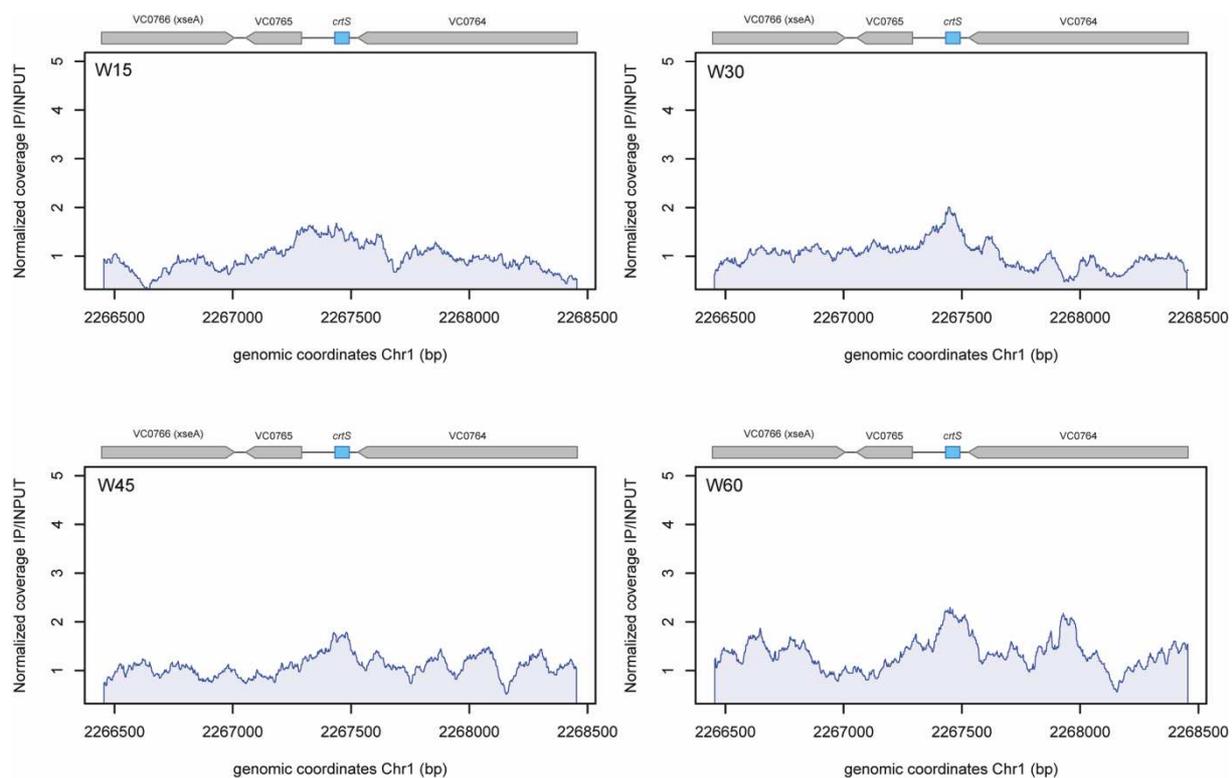
Supplementary Figure 2: RctB secondary binding sites. on y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference genome (CP028827.1 and CP028828.1). A 2.000bp window is represented and centered around the peak. ChIP signals are represented in blue. Genetic context is depicted on top of the graph: genes as grey arrow.



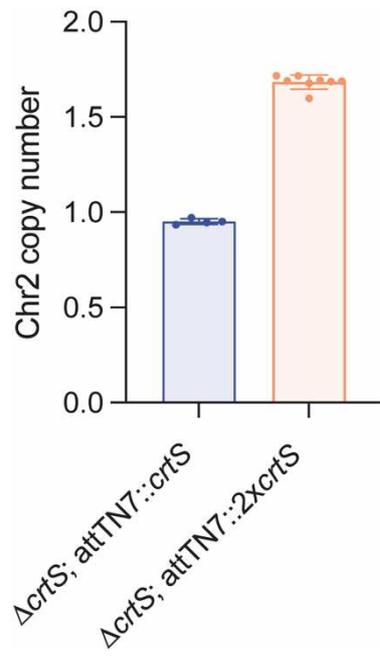
Supplementary Figure 3: assessing the role of RctB secondary binding sites. **a.** pOR12 copy number (CN) relative to the chromosome (pOR12/oriC) in *E. coli* MG1655 strains with and without a chromosomal RctB binding. **b.** Chr2 copy number relative to Chr1 (*ori2/ori1*) in non-replicating *V. cholerae*. For VC0643 and VC1643 contained in CDS RctB binding sites have been inactivated without perturbing codons sequence, for VC1042-1043 the intergenic region containing RctB has been deleted. **c** and **d.** Expression of VC1042 and VC1803 genes relative to the housekeeping gene *gyrA* in the N16961 wild-type strain (WT), and MCH2 ($\Delta rctB$) *V. cholerae* mutant carrying fused chromosomes.



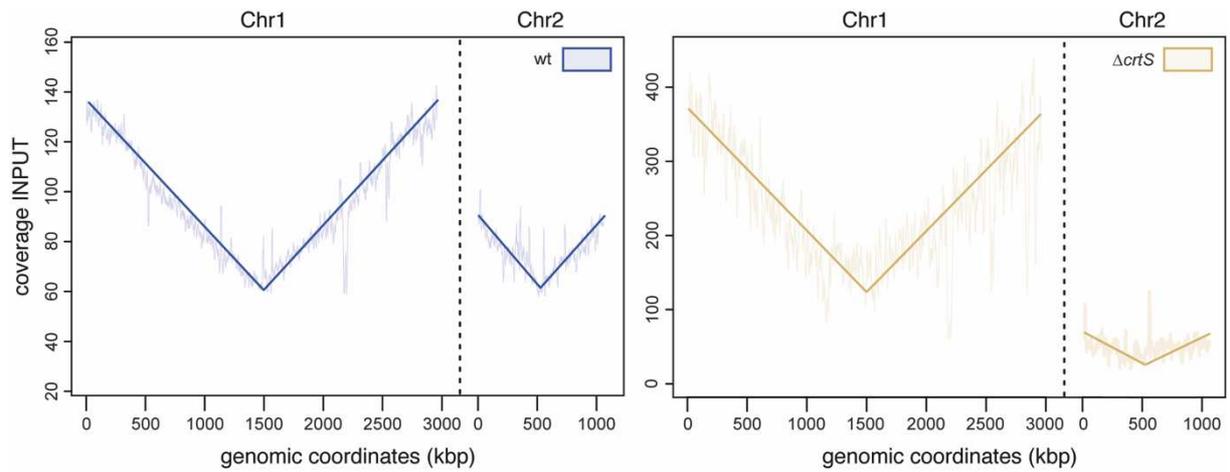
Supplementary Figure 4: RctB ChIP profil remains similar between growing and not growing phases. a & b. on y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference genome (CP028827.1 and CP028828.1). A 2.000bp window is represented and centered around the peak. ChIP signals are represented in blue for exponential phase and in red for stationary phase. Genetic context and known RctB binding sites are depicted on top of the graph: genes as grey arrow, 39mer site as red box, iterons as green triangle, *crtS* as cyan box. The AT-rich region also called DNA unwinding element (DUE) is also depicted.



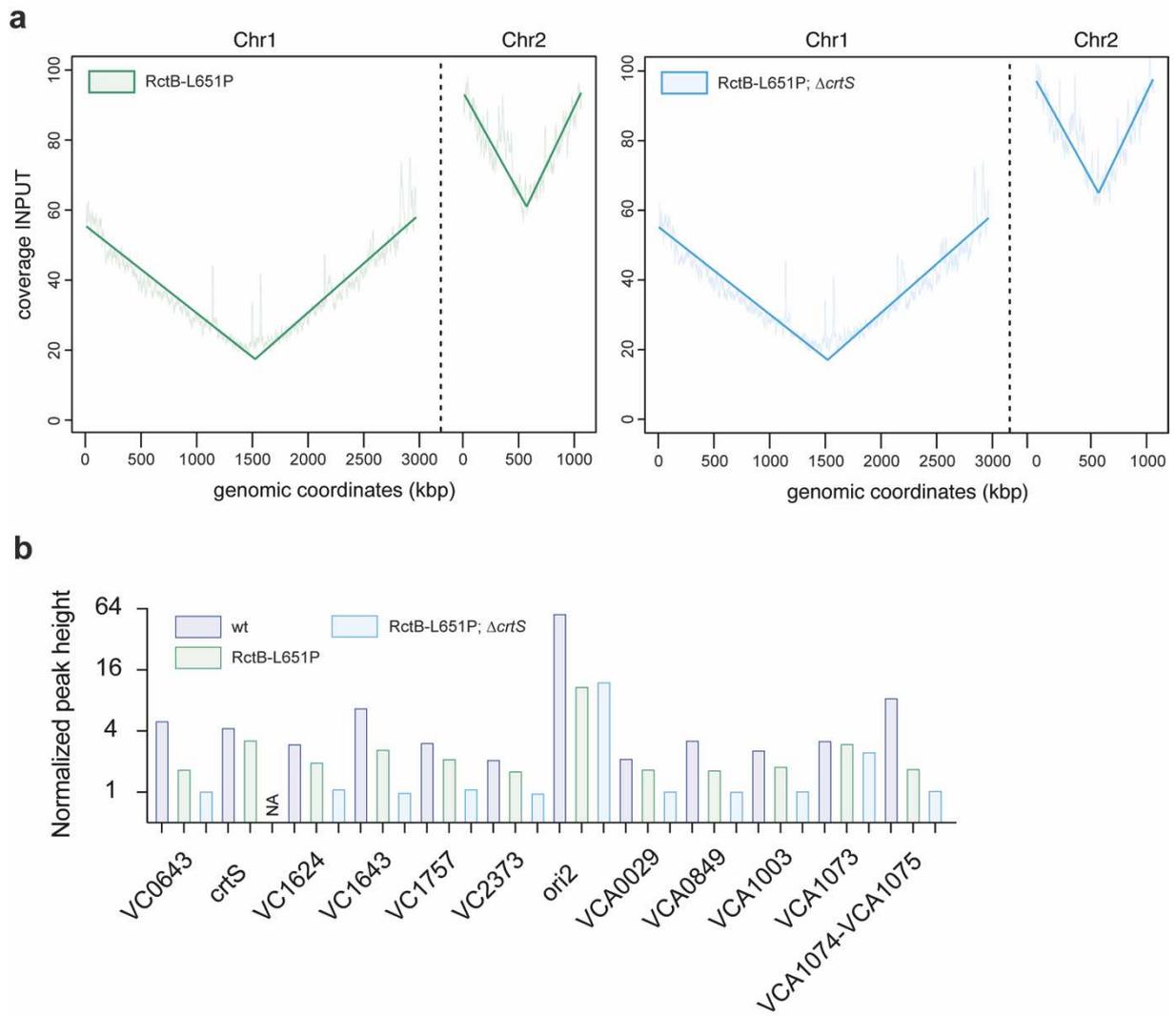
Supplementary Figure 5: RctB binding pattern on *crtS* in synchronized population. ChIPseq experiments on synchronized population at different time points after SHX removal (R+15min, R+30min, R+45min, R+60min). On y axis: RctB normalized ChIP signal (IP coverage/INPUT coverage), on x axis: genomic coordinates on *Vibrio cholerae* N16961 reference Chr1 (CP028827.1). A 2.000bp window is represented centered around *crtS*. Genetic is depicted on top of the graph (genes in grey arrows, *crtS* in light blue square).



Supplementary Figure 6: RctB ChIP profil remains similar in a strain carrying two chromosomal copies of *crtS*. Chr2 copy number relative to Chr1 (*ori2/ori1*) in non-replicating *V. cholerae*, in strains with one ($\Delta crtS; attTN7::crtS$) or two chromosomal copies of *crtS* ($\Delta crtS; attTN7::2xcrtS$).



Supplementary Figure 7: MFA on wt and $\Delta crtS$ strains. Marker frequency analysis on the input sample (not immunoprecipitated) from the ChIP experiment, coverage is represented on y axis and genomic coordinates of both chromosomes (in kilo base pair) are present on the x axis. Curves have been smoothed with a 10kb sliding window along the genome. Data from the wt input are represented in blue, those from the $\Delta crtS$ input in sand.



Supplementary figure 8: MFA RctB L651P with or without *crtS* site. **a.** Marker frequency analysis on the input sample (not immunoprecipitated) from the ChIP experiment, coverage is represented on y axis and genomic coordinates of both chromosomes (in kilo base pair) are present on the x axis. Curves have been smoothed with a 10kb sliding window along the genome. Data from the RctB-L651P input are represented in green, those from the double mutant RctB-L651P $\Delta crtS$ input in light blue. **b.** Maximum value of normalized peak height (=IP coverage/INPUT coverage) for RctBwt and RctBL651P, calculated on biological duplicate ChIP experiment.

Chromosome	Peak summit	-LOG10(pvalue)	fold enrichment	Locus name	Identified motif	site type	genetic context
CP028827.1	547255	21.84341	2.16975	VC2373	atgttctctccggaactatggtagcggttcggttcggg	39mer-like	CDS
CP028827.1	1134355	28.65055	2.39152	VC1802-VC1803-VC1803-VC1804	NA	NA	intergenic
CP028827.1	1188134	24.52891	2.26743	VC1757	ctactaccggaaggcggttaaccaagttacggccaact	39mer-like	CDS
CP028827.1	1312502	224.57054	5.99008	VC1643	gtttggcagggaaggcgctgtctcaaatcaggctaacg	39mer-like	CDS
CP028827.1	1335897	35.08026	2.57201	VC1624	aaggaaactgtctattctacagattggtagggcggtag	39mer-like	CDS
CP028827.1	1974161	15.83979	1.97413	VC1042-VC1043	NA	NA	intergenic
CP028827.1	2267454	117.16525	4.10547	VC0764-VC0765 (crtS)	ttgttttagattttatttttttaaaatggaatgcagaataatgtaacttaacttgccttgg	crtS	intergenic
CP028827.1	2398355	237.47833	5.83823	VC0643	tcttggcacaagaacaatgctacccttacgcaaggtag	39mer-like	CDS
Chromosome	Peak summit	-LOG10(pvalue)	fold enrichment	Locus name	Identified motif	site type	genetic context
CP028828.1	3414	7720.91699	66.55206	ori2	NA	NA	ori
CP028828.1	38465	29.31495	2.44963	VCA0029	atgataaaggtt	12mer	CDS
CP028828.1	799599	69.89726	3.53533	VCA0849	acaagaagcaccattactcagcagaataattacgggctatg	39mer-like	CDS
CP028828.1	959185	26.46835	2.40253	VCA1003	atgataatgctt	12mer	CDS
CP028828.1	1027217	78.14582	3.43762	VCA1073	atgataaaggtt, ttgatacgtt	2x12mer	CDS
CP028828.1	1030315	396.46924	7.49721	VCA1074-VCA1075	tcggaacaatgataaagtagcacaataatacgggtaatg	39mer	intergenic

Supplementary table 1: RctB binding sites identified from ChIPseq experiment. ChIP reads data were aligned on *Vibrio cholerae* N16961 reference genome (CP028827.1 for Chr1 and CP028828.1 for Chr2). Peak summit, p-value and fold enrichment are calculated with MACS2.0 program by comparing RctB-FLAG versus wt (no FLAG) ChIPseq data. ChIP peaks motif are found using the MEME suite, two sequences are significantly enriched and correspond to already known iterons (left) and 39mer sites (right).

- 1 Hall, J. P. J., Botelho, J., Cazares, A. & Baltrus, D. A. What makes a megaplasmid? *Philos Trans R Soc Lond B Biol Sci* **377**, 20200472, doi:10.1098/rstb.2020.0472 (2022).
- 2 diCenzo, G. C. & Finan, T. M. The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiology and molecular biology reviews : MMBR* **81**, doi:10.1128/MMBR.00019-17 (2017).
- 3 Sonnenberg, C. B., Kahlke, T. & Haugen, P. Vibronaceae core, shell and cloud genes are non-randomly distributed on Chr 1: An hypothesis that links the genomic location of genes with their intracellular placement. *BMC genomics* **21**, 695, doi:10.1186/s12864-020-07117-5 (2020).
- 4 Fournes, F., Val, M. E., Skovgaard, O. & Mazel, D. Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes. *Frontiers in microbiology* **9**, 1833, doi:10.3389/fmicb.2018.01833 (2018).
- 5 Reyes-Lamothe, R. & Sherratt, D. J. The bacterial cell cycle, chromosome inheritance and cell growth. *Nature reviews. Microbiology* **17**, 467-478, doi:10.1038/s41579-019-0212-7 (2019).
- 6 Beaufay, F., Coppine, J. & Hallez, R. When the metabolism meets the cell cycle in bacteria. *Current opinion in microbiology* **60**, 104-113, doi:10.1016/j.mib.2021.02.006 (2021).
- 7 Ramachandran, R., Jha, J., Paulsson, J. & Chatteraj, D. Random versus Cell Cycle-Regulated Replication Initiation in Bacteria: Insights from Studying *Vibrio cholerae* Chromosome 2. *Microbiology and molecular biology reviews : MMBR* **81**, doi:10.1128/MMBR.00033-16 (2017).
- 8 Rasmussen, T., Jensen, R. B. & Skovgaard, O. The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *The EMBO journal* **26**, 3124-3131, doi:10.1038/sj.emboj.7601747 (2007).
- 9 Baek, J. H. & Chatteraj, D. K. Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS genetics* **10**, e1004184, doi:10.1371/journal.pgen.1004184 (2014).
- 10 Val, M. E. *et al.* A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Science advances* **2**, e1501914, doi:10.1126/sciadv.1501914 (2016).
- 11 Kemter, F. S. *et al.* Synchronous termination of replication of the two chromosomes is an evolutionary selected feature in Vibronaceae. *PLoS genetics* **14**, e1007251, doi:10.1371/journal.pgen.1007251 (2018).
- 12 Du, W. L. *et al.* Orderly Replication and Segregation of the Four Replicons of *Burkholderia cenocepacia* J2315. *PLoS genetics* **12**, e1006172, doi:10.1371/journal.pgen.1006172 (2016).
- 13 Frage, B. *et al.* Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Molecular microbiology* **100**, 808-823, doi:10.1111/mmi.13351 (2016).
- 14 Deghelt, M. *et al.* G1-arrested newborn cells are the predominant infectious form of the pathogen *Brucella abortus*. *Nature communications* **5**, 4366, doi:10.1038/ncomms5366 (2014).
- 15 Dubarry, N., Willis, C. R., Ball, G., Lesterlin, C. & Armitage, J. P. In Vivo Imaging of the Segregation of the 2 Chromosomes and the Cell Division Proteins of *Rhodobacter sphaeroides* Reveals an Unexpected Role for MipZ. *mBio* **10**, doi:10.1128/mBio.02515-18 (2019).
- 16 Xie, B. B. *et al.* Evolutionary Trajectory of the Replication Mode of Bacterial Replicons. *mBio* **12**, doi:10.1128/mBio.02745-20 (2021).
- 17 Ren, Z. *et al.* Conformation and dynamic interactions of the multipartite genome in *Agrobacterium tumefaciens*. *Proceedings of the National Academy of Sciences of the United States of America* **119**, doi:10.1073/pnas.2115854119 (2022).
- 18 Egan, E. S. & Waldor, M. K. Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* **114**, 521-530, doi:10.1016/S0092-8674(03)00611-1 (2003).
- 19 Orlova, N. *et al.* The replication initiator of the cholera pathogen's second chromosome shows structural similarity to plasmid initiators. *Nucleic acids research* **45**, 3724-3737, doi:10.1093/nar/gkw1288 (2017).
- 20 Jha, J. K., Demarre, G., Venkova-Canova, T. & Chatteraj, D. K. Replication regulation of *Vibrio cholerae* chromosome II involves initiator binding to the origin both as monomer and as dimer. *Nucleic acids research* **40**, 6026-6038, doi:10.1093/nar/gks260 (2012).
- 21 Fournes, F. *et al.* The coordinated replication of *Vibrio cholerae*'s two chromosomes required the acquisition of a unique domain by the RctB initiator. *Nucleic acids research* **49**, 11119-11133, doi:10.1093/nar/gkab903 (2021).
- 22 Kothapalli, R. *et al.* The dimerization interface of initiator RctB governs chaperone and enhancer dependence of *Vibrio cholerae* chromosome 2 replication. *Nucleic acids research*, doi:10.1093/nar/gkac210 (2022).
- 23 Jha, J. K. *et al.* The DnaK Chaperone Uses Different Mechanisms To Promote and Inhibit Replication of *Vibrio cholerae* Chromosome 2. *mBio* **8**, doi:10.1128/mBio.00427-17 (2017).
- 24 Chatterjee, S., Jha, J. K., Ciaccia, P., Venkova, T. & Chatteraj, D. K. Interactions of replication initiator RctB with single- and double-stranded DNA in origin opening of *Vibrio cholerae* chromosome 2. *Nucleic acids research*, doi:10.1093/nar/gkaa826 (2020).
- 25 Wegrzyn, K. E., Gross, M., Uciechowska, U. & Konieczny, I. Replisome Assembly at Bacterial Chromosomes and Iteron Plasmids. *Frontiers in molecular biosciences* **3**, 39, doi:10.3389/fmolb.2016.00039 (2016).
- 26 Venkova-Canova, T. & Chatteraj, D. K. Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6199-6204, doi:10.1073/pnas.1013244108 (2011).
- 27 Venkova-Canova, T., Saha, A. & Chatteraj, D. K. A 29-mer site regulates transcription of the initiator gene as well as function of the replication origin of *Vibrio cholerae* chromosome II. *Plasmid* **67**, 102-110, doi:10.1016/j.plasmid.2011.12.009 (2012).

- 28 de Lemos Martins, F., Fournes, F., Mazzuoli, M. V., Mazel, D. & Val, M. E. Vibrio cholerae chromosome 2
copy number is controlled by the methylation-independent binding of its monomeric initiator to the
chromosome 1 *crtS* site. *Nucleic acids research* **46**, 10145-10156, doi:10.1093/nar/gky790 (2018).
- 29 Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic acids research* **43**, W39-
49, doi:10.1093/nar/gkv416 (2015).
- 30 Venkova-Canova, T., Baek, J. H., Fitzgerald, P. C., Blokesch, M. & Chattoraj, D. K. Evidence for two
different regulatory mechanisms linking replication and segregation of vibrio cholerae chromosome II.
PLoS genetics **9**, e1003579, doi:10.1371/journal.pgen.1003579 (2013).
- 31 Yu, Y., Ouyang, Y. & Yao, W. shinyCircos: an R/Shiny application for interactive creation of Circos plot.
Bioinformatics **34**, 1229-1231, doi:10.1093/bioinformatics/btx763 (2018).
- 32 Menikpurage, I. P., Woo, K. & Mera, P. E. Transcriptional Activity of the Bacterial Replication Initiator
DnaA. *Frontiers in microbiology* **12**, 662317, doi:10.3389/fmicb.2021.662317 (2021).
- 33 Zhang, Q. *et al.* Reversible lysine acetylation is involved in DNA replication initiation by regulating activities
of initiator DnaA in Escherichia coli. *Scientific reports* **6**, 30837, doi:10.1038/srep30837 (2016).
- 34 Ferullo, D. J., Cooper, D. L., Moore, H. R. & Lovett, S. T. Cell cycle synchronization of Escherichia coli
using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods*
48, 8-13, doi:10.1016/j.ymeth.2009.02.010 (2009).
- 35 Kemter, F. S. *et al.* Stringent response leads to continued cell division and a temporal restart of DNA
replication after initial shutdown in Vibrio cholerae. *Molecular microbiology* **111**, 1617-1637,
doi:10.1111/mmi.14241 (2019).
- 36 Gross, M. H. & Konieczny, I. Polyphosphate induces the proteolysis of ADP-bound fraction of initiator to
inhibit DNA replication initiation upon stress in Escherichia coli. *Nucleic acids research*,
doi:10.1093/nar/gkaa217 (2020).
- 37 Ramachandran, R., Ciaccia, P. N., Filsuf, T. A., Jha, J. K. & Chattoraj, D. K. Chromosome 1 licenses
chromosome 2 replication in Vibrio cholerae by doubling the *crtS* gene dosage. *PLoS genetics* **14**,
e1007426, doi:10.1371/journal.pgen.1007426 (2018).
- 38 Dalia, A. B. & Dalia, T. N. Spatiotemporal Analysis of DNA Integration during Natural Transformation
Reveals a Mode of Nongenetic Inheritance in Bacteria. *Cell* **179**, 1499-1511 e1410,
doi:10.1016/j.cell.2019.11.021 (2019).
- 39 Jha, J. K., Ghirlando, R. & Chattoraj, D. K. Initiator protein dimerization plays a key role in replication
control of Vibrio cholerae chromosome 2. *Nucleic acids research* **42**, 10538-10549,
doi:10.1093/nar/gku771 (2014).
- 40 Das, D. *et al.* The structure of the first representative of Pfam family PF09836 reveals a two-domain
organization and suggests involvement in transcriptional regulation. *Acta Crystallogr Sect F Struct Biol
Cryst Commun* **66**, 1174-1181, doi:10.1107/S1744309109022672 (2010).
- 41 Karimova, G., Pidoux, J., Ullmann, A. & Ladant, D. A bacterial two-hybrid system based on a reconstituted
signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of
America* **95**, 5752-5756, doi:10.1073/pnas.95.10.5752 (1998).
- 42 Diaz, R. E., Sanchez, A., Anton Le Berre, V. & Bouet, J. Y. High-Resolution Chromatin
Immunoprecipitation: ChIP-Sequencing. *Methods in molecular biology* **1624**, 61-73, doi:10.1007/978-1-
4939-7098-8_6 (2017).
- 43 Skovgaard, O., Bak, M., Lobner-Olesen, A. & Tommerup, N. Genome-wide detection of chromosomal
rearrangements, indels, and mutations in circular chromosomes by short read sequencing. *Genome
research* **21**, 1388-1393, doi:10.1101/gr.117416.110 (2011).
- 44 Madic, J. *et al.* Three-color crystal digital PCR. *Biomolecular detection and quantification* **10**, 34-46,
doi:10.1016/j.bdq.2016.10.002 (2016).
- 45 Lo Scudato, M. & Blokesch, M. The regulatory network of natural competence and transformation of Vibrio
cholerae. *PLoS genetics* **8**, e1002778, doi:10.1371/journal.pgen.1002778 (2012).

III. Résultats non publiés : étude du rôle de Lrp dans le mécanisme *crtS-ori2*

1. Introduction : Lrp un régulateur global

La transcription bactérienne est contrôlée de manière hiérarchique : un petit nombre de régulateurs globaux fortement exprimés exercent un contrôle sur l'expression de centaines de facteurs de transcription secondaires en aval. *E.coli* possède sept régulateurs globaux de ce type : ArcA, CRP, FNR, IHF, Fis, H-NS, et Lrp (Martinez-Antonio and Collado-Vides, 2003).

Parmi ces régulateurs globaux, la protéine Lrp (Leucine-responsive Regulatory Protein) est très conservée chez les procaryotes ; près de la moitié des bactéries séquencées et presque toutes les archées séquencées contiennent un ou plusieurs homologues de Lrp (Brinkman et al., 2003).

Lrp est une petite protéine (18,8 kDa), très abondante (environ 3 000 dimères/cellule), qui se lie à l'ADN. La protéine Lrp régule jusqu'à 32% des gènes chez *E.coli*, directement ou indirectement (Kroner et al., 2019). On pense qu'elle détecte l'état nutritionnel de l'environnement par le biais des niveaux de L-leucine exogène et qu'elle régule l'expression génétique en conséquence.

Le régulon Lrp d'*E. coli* comprend des gènes impliqués dans le métabolisme des acides aminés, le métabolisme d'un carbone, le transport des peptides et des acides aminés, la biosynthèse des flagelles et des fimbriae, le stress osmotique, le stress acide et d'autres facteurs de virulence. On sait que Lrp régule 70 % des gènes impliqués dans la transition de la croissance logarithmique à la phase stationnaire lorsque les nutriments sont épuisés (Tani et al., 2002).

La protéine Lrp s'organise en deux domaines structuraux : un domaine HTH qui permet la liaison à l'ADN, un domaine RAM (regulation of amino acid metabolism) qui permet l'oligomérisation et la régulation par la leucine. Lrp peut adopter différents niveaux d'oligomérisation : on le retrouve en dimère lorsqu'il est dans le cytoplasme, il se fixe sur l'ADN sous une forme d'octamère (Figure 41).

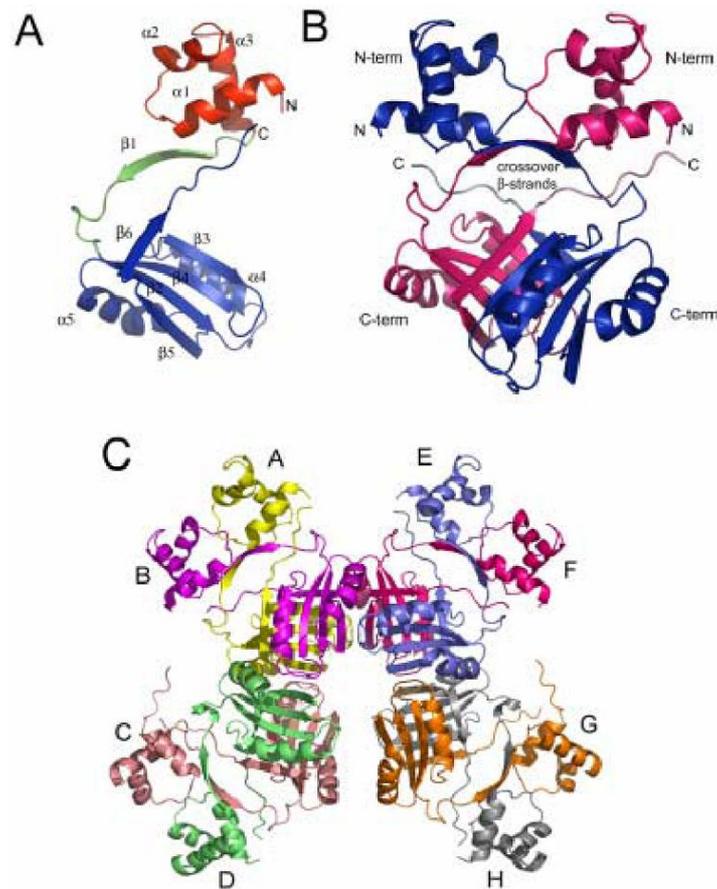


Figure 41. Lrp et ses différents niveaux d'oligomérisation. A. Monomère de Lrp. B. Dimère de Lrp. C. Octamère de Lrp (de los Rios and Perona, 2007).

Lorsque Lrp se fixe à l'ADN sous une forme d'octamère il enroule autour de lui l'ADN sur une longueur d'environ 120pb formant une structure « histone-like ». De plus, deux octamères fixés à l'ADN peuvent interagir et former un hexadécamère (16mer) (Figure 42). Il apparaît que dans la cellule, Lrp est trouvé dans un équilibre octamère et hexadécamère, cet équilibre semble pouvoir être influencé notamment vis-à-vis de la concentration en leucine du milieu. Une étude récente par ChIPseq de Lrp dans différentes conditions démontre que Lrp est capable de former des ponts (*bridge*) entre différentes régions d'ADN probablement en formant ces hexadécamères (Ziegler and Freddolino, 2021).

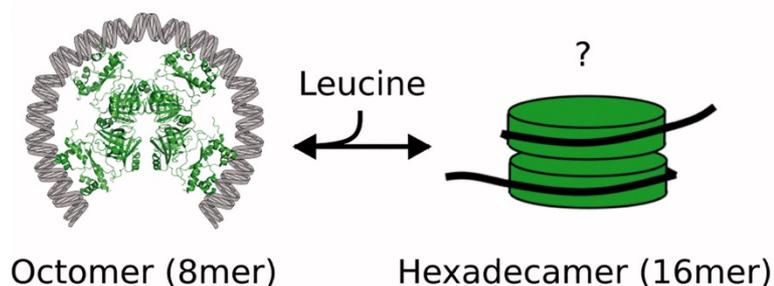


Figure 42. Lrp existe dans la cellule dans un équilibre d'octamère et d'hexadécamère.

2. Rôle de Lrp (Leucine responsive Regulatory Protein) dans le mécanisme *crtS-ori2*

Dans une étude récente réalisée dans l'équipe de Dhruva Chatteraj publiée en 2018, un crible génétique a permis de révéler que le régulateur global Lrp se fixe au site *crtS* (Ciaccia *et al.*, 2018). Ils ont ensuite démontré que la délétion du gène *lrp* chez *E.coli* abolit l'effet positif de *crtS* sur le nombre de copies d'un plasmide p*ORI2*. Puis ils ont montré par microscopie à fluorescence qu'une souche Δlrp chez *V.cholerae* sous-initie son Chr2. Ainsi il apparait qu'*in vivo* Lrp est nécessaire pour l'action de *crtS*. Enfin, par gel retard ils ont montré que Lrp facilite la fixation de RctB à *crtS* en formant un complexe tripartite RctB-Lrp-*crtS*. La leucine semble n'avoir aucun effet sur la formation de ce complexe. Leur hypothèse principale est que Lrp change localement la conformation de l'ADN permettant la reconnaissance de *crtS* par RctB.

Nous avons décidé de reproduire ces résultats de notre côté et de poursuivre ces recherches pour comprendre le rôle de Lrp dans le processus *crtS-ori2*. Premièrement, nous avons effectué la délétion du gène *lrp* (VC1904) dans notre souche de *V.cholerae*. Puis nous avons étudié le phénotype sous le microscope du mutant Δlrp . Celle-ci présente un phénotype de filamentation qui sous-tend une perturbation de la division cellulaire (Figure 43). Ce phénotype ressemble fortement à celui observé pour la souche $\Delta crtS$.

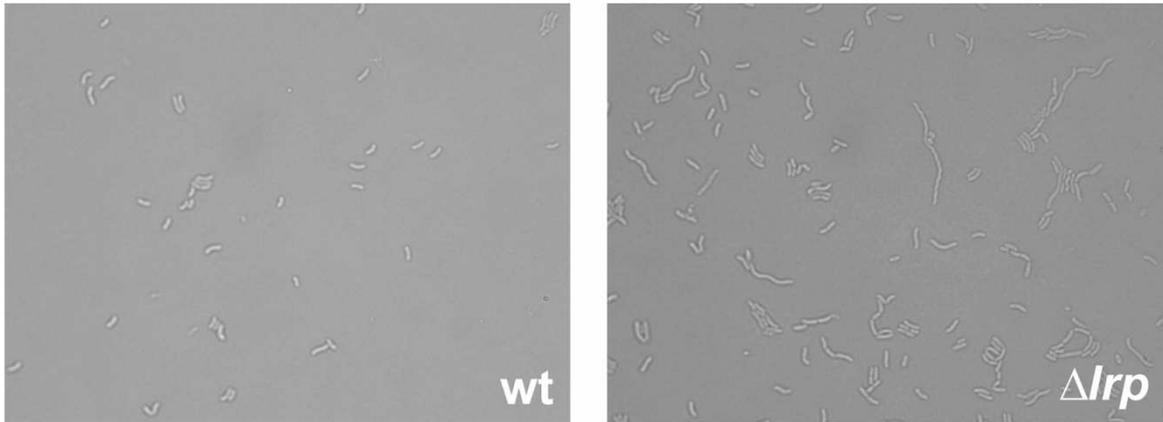


Figure 43. La délétion de *lrp* entraîne un phénotype de filamentation chez *V.cholerae*. Microscopie sur pad d'agarose M9 glucose1%. Observation d'une culture bactérienne de milieu liquide M9 glucose. Panel de gauche : *Vibrio cholerae* N16961 sauvage. Panel de droite : *Vibrio cholerae* N16961 Δlrp (VC1904).

Nous avons ensuite mesuré par PCR digitale le nombre de copies du Chr2 par rapport au Chr1 dans la souche Δlrp . Le nombre de copie du Chr2 par cellule est d'environ 0.3 alors qu'il est de 1 pour la souche sauvage. Ainsi *lrp* est bien nécessaire pour la réplication du Chr2. A noté que ce défaut de réplication du Chr2 est le même qu'observé dans la souche $\Delta crtS$ (Figure 44).

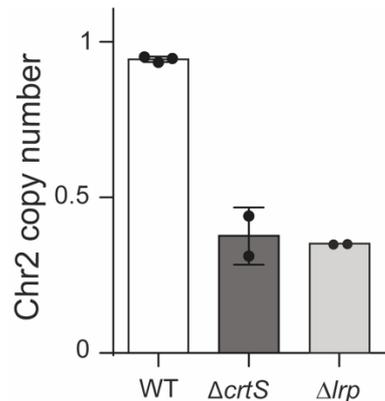


Figure 44. La délétion de *lrp* ou de *crtS* entraîne le même défaut de réplication du Chr2. Histogrammes représentant le nombre de copies de Chr2 par rapport à Chr1 (*ori2/ori1*) dans les souches de *V. cholerae*. Le nombre de copies de *ori1* et *ori2* a été mesuré par dPCR sur de l'ADN génomique provenant de cellules non répliquées (phase stationnaire) après 16 heures de croissance. Toutes les données représentent la moyenne de trois expériences indépendantes (\pm écart-type).

Nous avons également isolé quatre souches Δlrp qui ne présentait plus de défaut de réplication du Chr2 et nous avons séquencé leur génome pour trouver d'éventuelles mutations compensatoires. Parmi ces quatre souches, une porte une mutation dans le gène *rctB* plus précisément dans la partie C-terminale de RctB. Cette mutation est une délétion en phase de 9 nucléotides (acides aminés) : I577-K578-E579. Nous savons que cette région de RctB est liée à la régulation négative de l'*ori2*. Pour les trois autres mutants on observe une fusion des

deux chromosomes. Ces mutations confirment le lien direct entre Lrp et la réplication du Chr2, en effet il apparaît que le site *crtS* ne fonctionne pas dans un contexte Δlrp .

3. Lrp et RctB sur *crtS*: une fixation coopérative

Afin de mieux comprendre pourquoi la protéine Lrp est importante pour l'activité de *crtS* nous avons décidé de caractériser l'interaction Lrp-*crtS* de manière *in vitro* en utilisant différentes approches biochimiques. Dans un premier temps, nous avons réalisé un gel retard de RctB et de Lrp sur une sonde d'ADN contenant *crtS* (Figure 45). Dans les conditions utilisées ici Lrp est capable de se fixer de manière stable sur *crtS*. A l'inverse, nous n'observons pas de fixation de RctB seul à *crtS*. Cependant lorsque on ajoute les deux protéines Lrp et RctB simultanément on observe l'apparition d'un complexe supérieur sur le gel. Lrp facilite donc le recrutement de RctB à *crtS* et il y a formation d'un complexe tripartite *crtS*-RctB-Lrp.

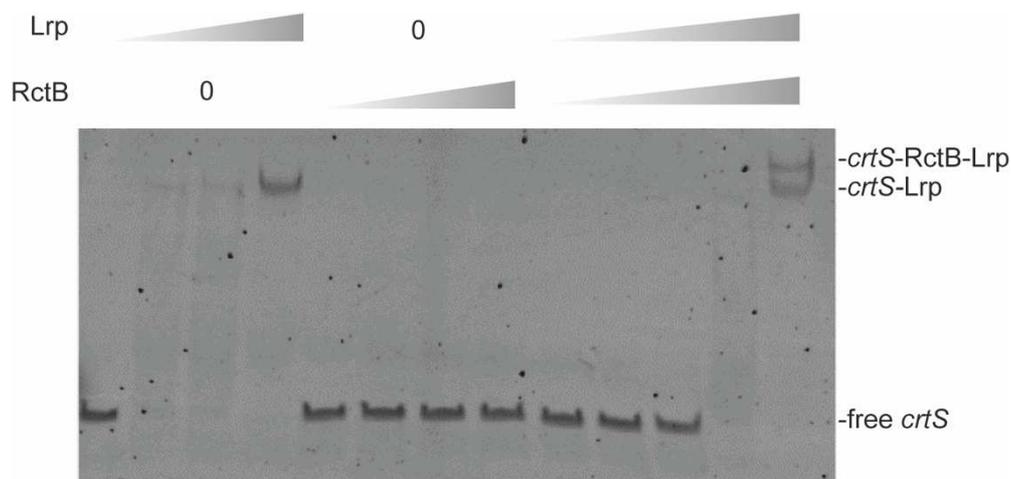


Figure 45. Gel retard de RctB et Lrp sur le site ADN *crtS*. Gel retard réalisé avec les protéines Lrp et RctB sur une sonde d'ADN de 170pb contenant le site *crtS* et marqué en 5' au Cy5. Des quantités croissantes de RctB (puits 1 à 4 et 9 à 13), et de Lrp (puits 5 à 13) ont été ajoutés à des réactions contenant chacune 5nM de la sonde d'ADN fluorescente. A noter qu'une seconde bande pour Lrp apparaît au-dessus de la première en présence de RctB.

Nous avons pu déterminer des conditions pour former un complexe stable de Lrp sur *crtS*, nous avons alors décidé d'observer ce complexe directement par microscopie électronique. Nous avons visualisé le complexe Lrp-*crtS* via microscopie électronique à transmission en le marquant positivement à l'acétate d'uranium en collaboration avec Eric Lecam à l'institut Gustave Roussy.

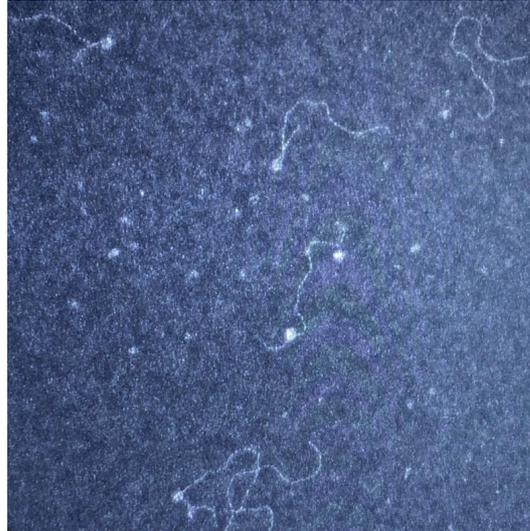


Figure 46. Visualisation du complexe nucléoprotéique Lrp-*crtS* via TEM. A. Observation par microscopie électronique à transmission du complexe Lrp-*crtS* sur une couche de carbone après coloration positive à l'acétate d'uranium. La sonde d'ADN de 944pb contient le site *crtS* (placé à environ un quart de la sonde), elle a été mélangée dans un milieu Tris NaCl et magnésium avec la protéine Lrp de *Vibrio cholerae* (50nM). Les flèches blanches indiquent les complexes Lrp-*crtS*. B. Modèle de l'interaction d'un octamère de Lrp (tétramère de dimère) avec l'ADN : l'octamère de Lrp enroule l'ADN sur environ 130pb (de los Rios and Perona, 2007).

Nous observons une fixation spécifique de Lrp sur *crtS* (situé à environ $\frac{1}{4}$ de la sonde ADN), l'ADN s'entoure autour de Lrp au niveau de *crtS* en accord avec les modèles prédits d'interaction Lrp avec l'ADN (Figure 46). Nous savons dorénavant que Lrp change localement la conformation de l'ADN au niveau de *crtS*. Ce changement de conformation pourrait permettre à RctB de reconnaître *crtS*.

4. CryoEM : observations du complexe RctB-Lrp-*crtS*

Pour poursuivre ce projet nous avons décidé d'étudier directement le complexe formé par Lrp-RctB-*crtS* par CryoEM. Cette étude est réalisée en collaboration avec l'équipe d'Abel Garcia Pino de l'Université Libre de Bruxelles (ULB) et plus particulièrement par Ariel Talavera. La prise d'image est réalisée à la plateforme de Cryo-EM de l'Institut Pasteur. Nous avons obtenu les premiers résultats de structure avec la protéine RctB seule dans un premier temps.

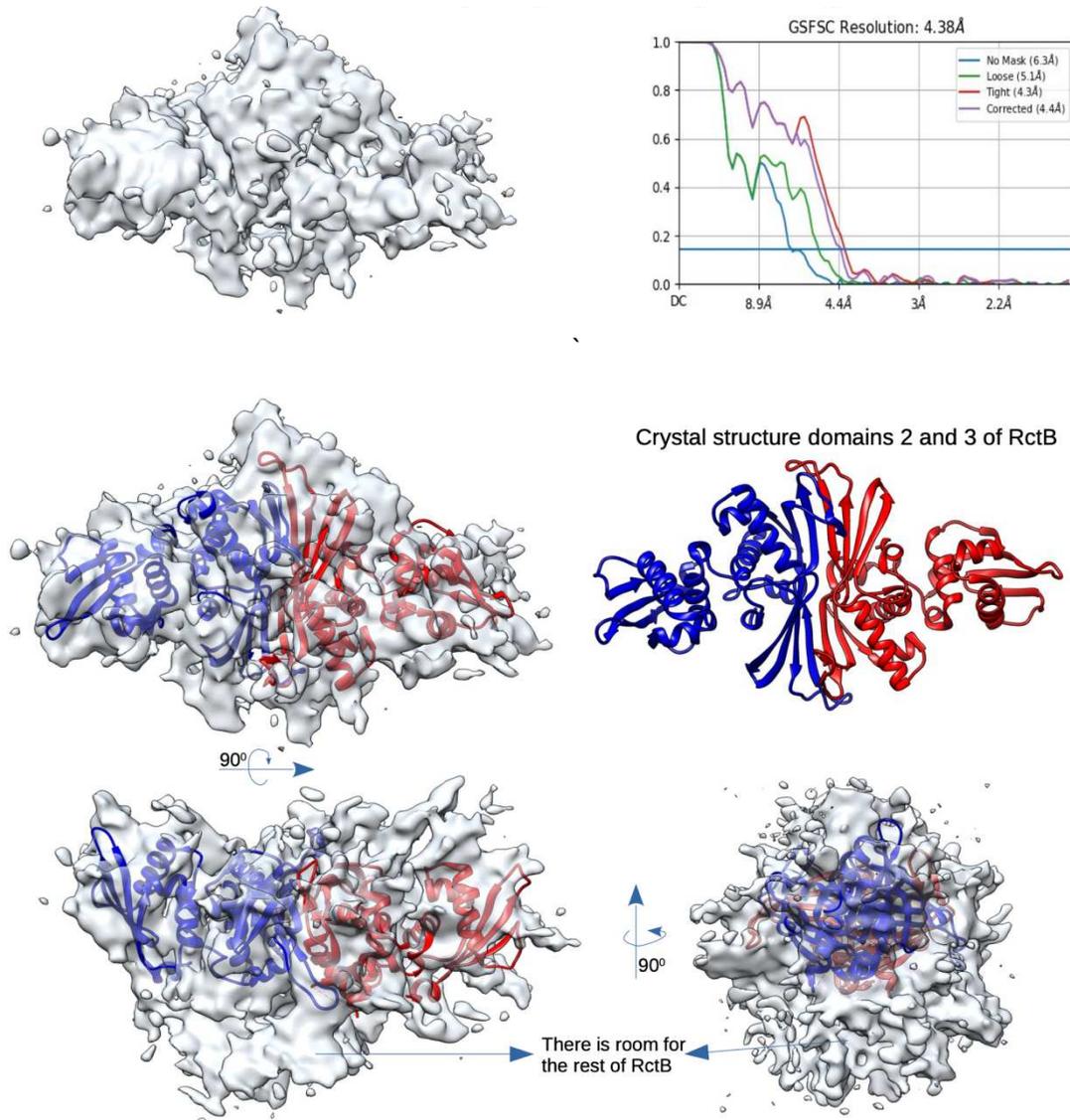


Figure 47. Structure de la protéine RctB par CryoEM. La protéine RctB a été observée par CryoEM, la résolution obtenue est de 4.4Å ce qui correspond à une basse résolution. RctB adopte sur la grille une orientation préférentielle qui perturbe la prise d'image complète de toutes les faces de RctB. La structure obtenue par cryoEM peut se superposer avec la structure connue du dimère des domaines II-III de RctB.

Pour le moment nous avons obtenu uniquement la structure de RctB avec une basse résolution (4.4Å) (Figure 47). Lorsqu'on compare la structure déjà connue des domaines II-III de RctB (formant un dimère) avec notre structure de cryoEM, on observe une bonne superposition et qu'il reste également de la place pour les deux autres domaines de RctB. Nous réaliserons bientôt d'autres observations de la protéine RctB seule pour affiner sa structure. Par la suite nous observerons la structure du complexe Lrp-*crtS*-RctB, nous espérons que celui-ci nous livrera des informations clés pour la compréhension du mécanisme *crtS-ori2*.

Pour conclure cette partie d'étude *in vitro*, nous savons que Lrp se fixe spécifiquement à *crtS* et enroule l'ADN formant une structure ressemblant à celle des histones chez les eucaryotes. Notre hypothèse est que le changement de conformation local de l'ADN induit par Lrp permet à RctB de reconnaître *crtS*, des études complémentaires et notamment la cryoEM nous permettront d'affiner ce modèle.

5. ChIPseq de RctB dans un contexte Δlrp

Pour comprendre le rôle de *lrp* sur la réplication du Chr2 de manière *in vivo* nous avons réalisé le ChIPseq de RctB dans un contexte génétique Δlrp . Ce ChIPseq a été fait dans une population en phase exponentielle de croissance ($DO_{600}=0.5$).

Le premier résultat de ce ChIPseq est que RctB ne se fixe plus au site *crtS* lorsque Lrp est absent (Figure 48), Lrp est donc essentiel pour le recrutement de RctB à *crtS*. De plus, lorsqu'on observe le signal de ChIP à l'*ori2*, celui-ci augmente d'environ 10 fois sur les sites 39mer. En superposant le signal de ChIP de RctB dans un contexte Δlrp ou $\Delta crtS$ on observe exactement le même pattern. Lorsque Lrp est absent l'initiation de l'*ori2* est fortement inhibée.

Pour conclure, ces résultats permettent d'expliquer pourquoi la délétion de *lrp* a un effet négatif sur l'initiation du Chr2. Lrp est nécessaire au fonctionnement de *crtS* car il permet le recrutement de RctB *in vivo*. A noter qu'il est possible de détecter la fixation de RctB seul (sans Lrp) sur *crtS* de manière *in vitro* par EMSA mais que cette fixation est plutôt faible (de Lemos Martins *et al.*, 2018).

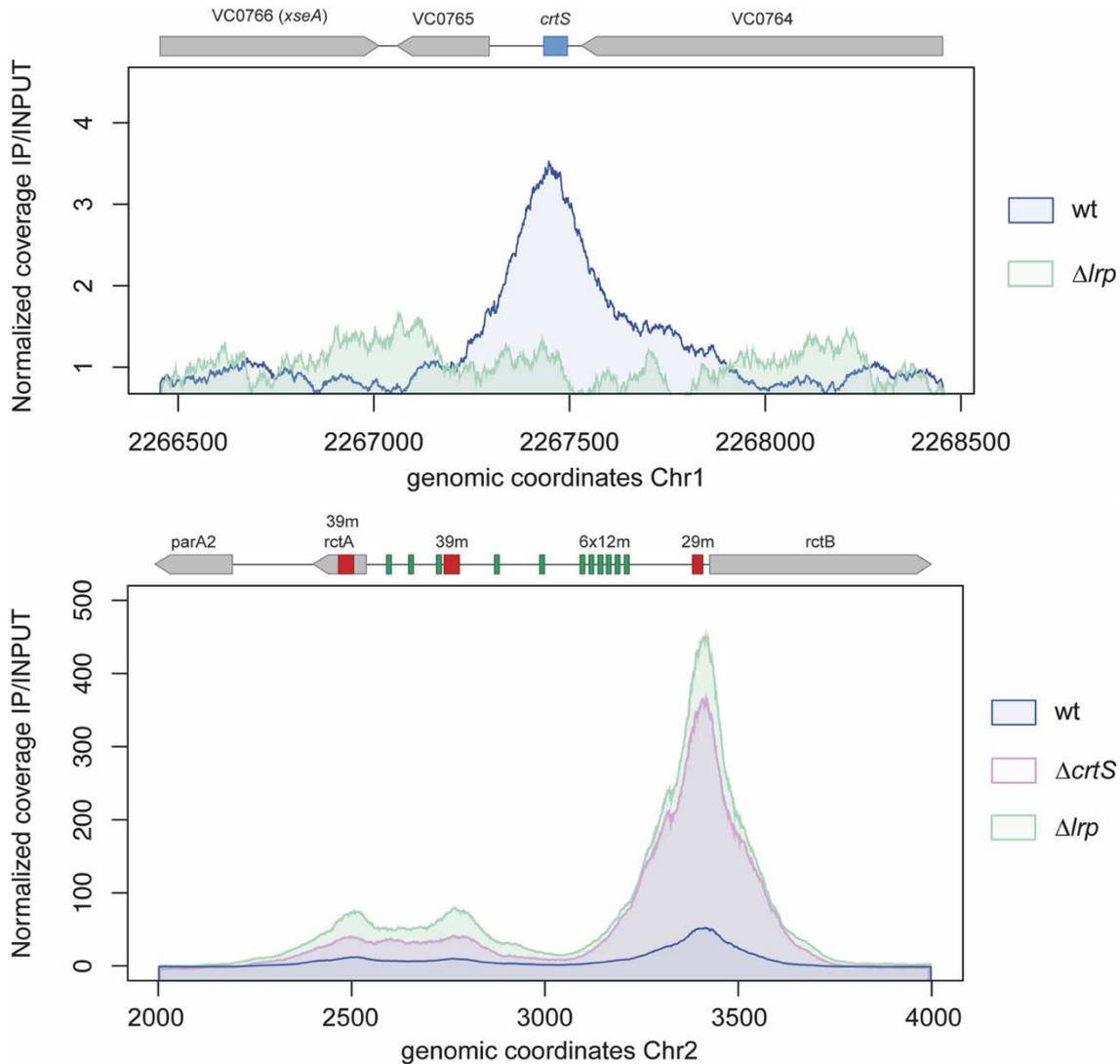


Figure 48. Influence de Lrp sur le profil de fixation de RctB. A. Signal de ChIPseq de RctB au locus *crtS* dans un contexte wt (bleu) ou Δlrp (vert). B Signal de ChIPseq de RctB au locus *ori2* dans un contexte wt (bleu) ou Δlrp (vert) ou $\Delta crtS$ (rose). Les détails du contexte génétique sont représentés en haut du graphique.

6. ChIPseq de Lrp chez *Vibrio cholerae*

Pour mieux comprendre le rôle de Lrp dans le mécanisme de coordination de la réplication par *crtS* nous avons décidé de réaliser l'analyse par ChIPseq de la fixation de Lrp sur le génome de *V.cholerae*. Pour cela nous avons remplacé directement sur le génome le gène *lrp* (VC1904) par un allèle *lrp-3xFLAG* (tag en Cter) puis vérifié que ce tag n'a pas d'impact sur l'action de Lrp en mesurant le nombre de copies du Chr2.

Nous avons réalisé ce ChIPseq en phase exponentiel (DO600=0.5). Après analyse des pics de fixation avec le programme MACS2.0 nous avons trouvé 193 pics sur le Chr1 et 44 sur le Chr2 (Figure 49).

Lrp CHIPseq

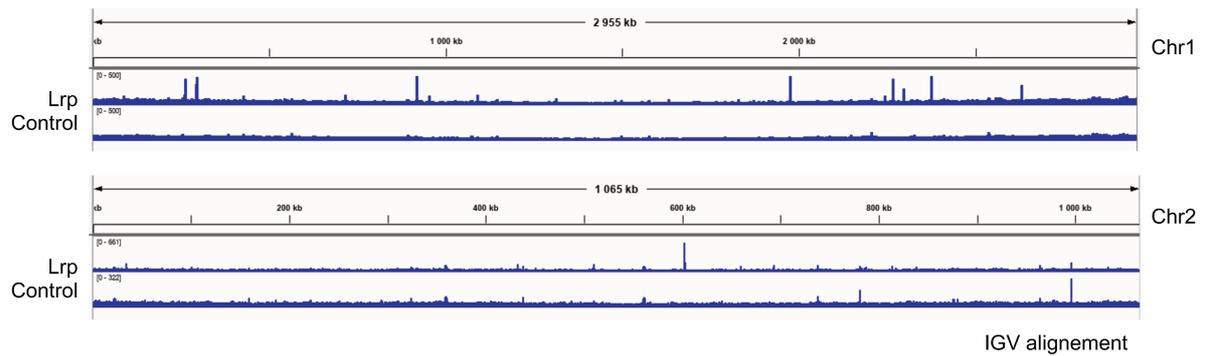


Figure 49. Profil de fixation de Lrp sur le génome de *V.cholerae*

Puis nous avons cherché un motif récurrent parmi ces sites en utilisant le programme MEME-CHIP, et 132 pics possèdent un motif très proche de celui chez *E.coli* (Figure 50).

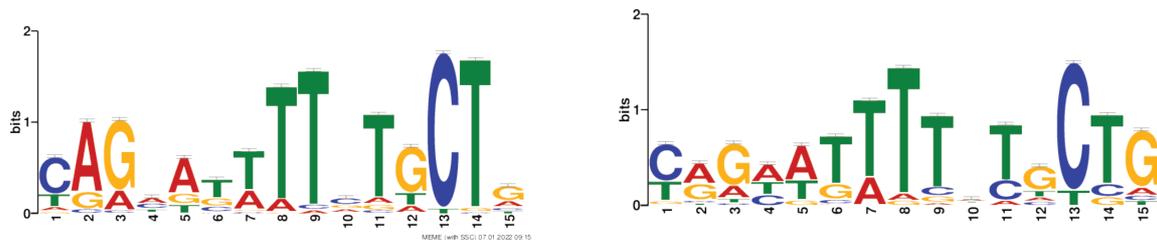


Figure 50. Séquence consensus des sites de fixation de Lrp. Gauche. Weblogo de la séquence consensus de fixation de Lrp à partir de l'analyse des pics du ChIPseq de Lrp. Droite. Weblogo publié de Lrp chez *E. coli*.

On détecte très bien la fixation de Lrp sur *crtS* (Figure 51), c'est le 6^{ème} site le plus enrichi sur l'ensemble du génome. Le pic est centré sur *crtS* mais semble aussi s'étendre du côté droit sur l'ADN, nous pensons qu'il pourrait s'agir de l'oligomérisation de Lrp sur l'ADN de manière non spécifique.

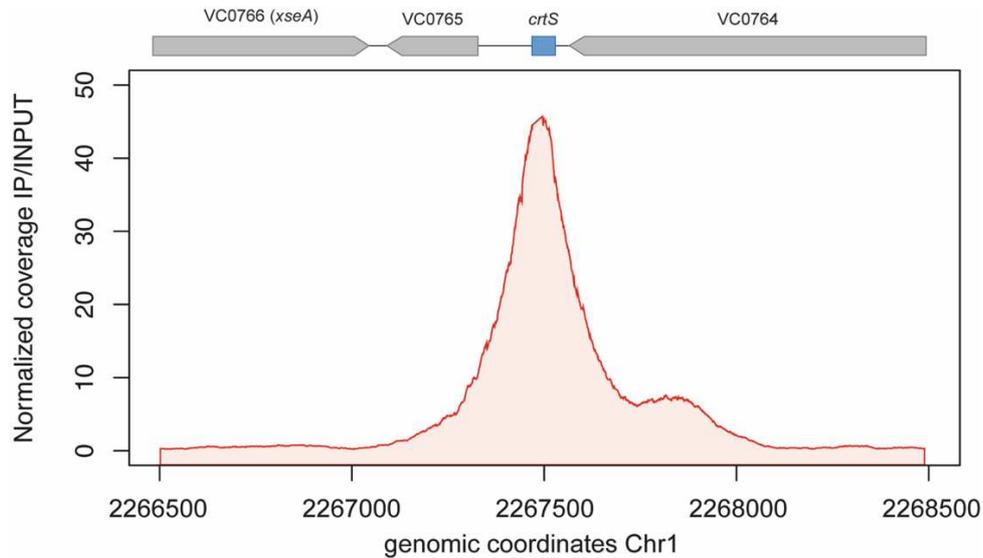


Figure 51. Données de ChIPseq de Lrp sur *crtS*. En rouge est représenté le signal normalisé de ChIPseq pour Lrp dans une culture de *V.cholerae* en phase exponentielle de croissance, la fenêtre est centrée autour du site *crtS*.

7. Effet de la phase de croissance sur le pattern de fixation de Lrp

D'après les résultats de notre article sur le ChIPseq de RctB nous savons que la fixation de RctB sur *crtS* ne varie pas selon la phase de croissance. RctB est toujours présent à *crtS* que ce soit en phase exponentielle ou en phase non réplivative. De plus, à partir de l'étude d'une population synchronisée il semble que la fixation de RctB à *crtS* ne varie pas non plus le long du cycle cellulaire. Nous nous sommes demandé si c'était également le cas pour Lrp, pour cela nous avons réalisé le ChIPseq de Lrp sur une culture en phase stationnaire (environ 16h de culture).

De manière globale, on observe que la fixation de Lrp le long du génome change selon la phase de croissance. Certains pics disparaissent ou diminuent, d'autres apparaissent. On observe que la fixation de Lrp à *crtS* est réduite en phase stationnaire par rapport à la phase exponentielle. Pourtant le niveau de protéine semble être similaire entre ces 2 phases de croissance comme l'atteste un western blot entre ces 2 conditions (Figure 52).

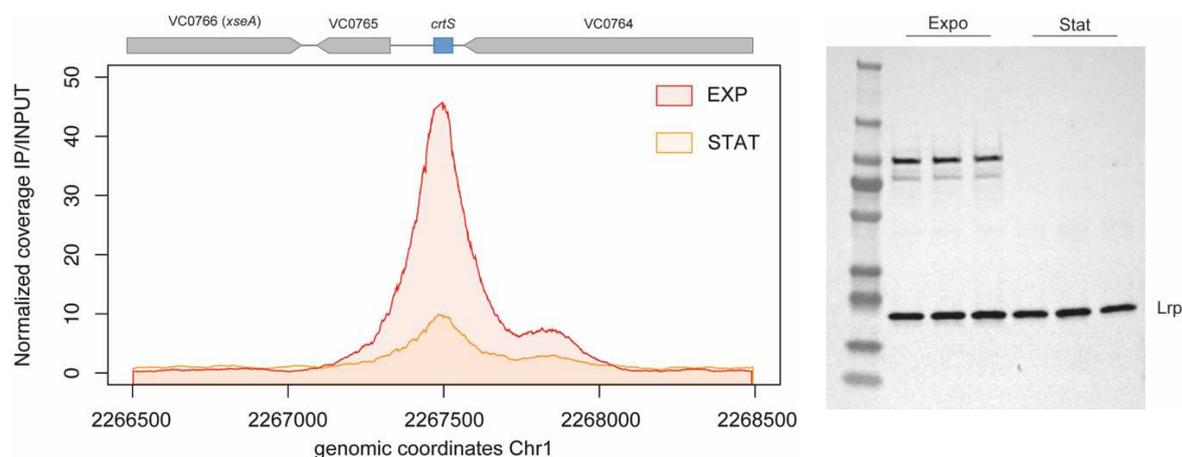


Figure 52. La fixation de Lrp est régulée par la phase de croissance. A) ChIPseq de Lrp chez *Vibrio cholerae* en phase exponentielle (rouge) et stationnaire (orange) centré sur le locus *crtS*. B) Western Blot anti-Lrp en phase exponentielle et stationnaire.

La différence que l'on observe entre phase exponentielle et phase stationnaire peut suggérer l'existence d'un mécanisme de régulation de la fixation de Lrp. Pour savoir si cette différence de fixation est liée à la réplication, nous avons décidé de réaliser le ChIPseq de Lrp dans une population synchronisée.

8. ChIPseq de Lrp en population synchronisée

Nous avons observé que contrairement à RctB la fixation de Lrp sur *crtS* varie selon la phase de croissance (diminuée en phase stationnaire). Nous nous sommes demandé si la fixation de Lrp à *crtS* pouvait également varier au cours du cycle cellulaire, pour étudier cela nous devons nous placer dans une population synchronisée pour la réplication.

Pour réaliser le ChIPseq de Lrp dans une population bactérienne synchronisée nous avons décidé de nous placer dans un système hétérologue. En effet, nous avons utilisé une souche d'*E.coli* MG1655 *dnaC2* que l'on peut synchroniser avec des changements successifs de températures.

dnaC2 est un allèle thermosensible de *dnaC*, loader de l'hélicase DnaB (PMC107071). DnaC est essentielle uniquement lors de l'initiation de la réplication mais elle n'est pas nécessaire pour la progression des fourches de réplication. Ainsi lorsque l'on fait passer une population bactérienne asynchrone en phase exponentielle de croissance (DO=0.2) à une température restrictive (39°C) les cellules terminent toutes leur réplication mais ne peuvent plus initier. Après 2H de culture on estime que toutes les cellules ont terminé leur réplication, on les passe ensuite à température permissive (30°C) et elles initient toutes leur réplication de manière synchrone.

Dans cette souche synchronisable nous avons introduit le site *crtS* au locus *lacZ* situé à 1.080.000pb de l'*oriC*. Nous avons ensuite réalisé le ChIPseq de Lrp d'*E.coli* en remplaçant préalablement l'allèle *lrp* sauvage par l'allèle taggué *lrp-3xflag*. Nous avons réalisé le ChIPseq de Lrp à différents temps : 10, 20 et 40min après initiation. En sachant que le temps de génération de la souche dans ces conditions est d'environ 40min. On devrait être capable de reconstituer l'ensemble du cycle cellulaire avec cette expérience.

Avec les données contrôles non immunoprécipitées du ChIPseq on peut réaliser un MFA (Marker Frequency Analysis), qui nous permet de suivre l'état d'avancement des fourches de réplication le long du génome du chromosome de *E.coli*. Pour le premier point en phase non réplivative, on observe un profil de MFA plat attestant de l'absence de fourche de réplication active (Figure 53). Au temps T10min, on observe la présence de deux fourches de réplication situées à environ 300kpb de l'*oriC*. Au temps T20min, les fourches se trouvent à environ 900kpb de l'*oriC*. Puis au temps T40min, les fourches ont quasiment terminé la réplication et se trouvent dans la région Ter à une distance d'environ 2.300.000pb de l'*oriC*. Ainsi, entre le temps T20 et T40, le locus qui contient *crtS* a été répliqué.

L'ensemble de ces résultats suggère que la fixation de Lrp est régulée par la réplication. En cela Lrp pourrait être le senseur du passage de la fourche dans le système *crtS-ori2*. Nous proposons que la réplication influence l'état d'oligomérisation de Lrp lors du passage de la fourche sur un site fixé par ce dernier.

9. L'état d'oligomérisation de Lrp dans l'action de *crtS*

La protéine Lrp semble pouvoir s'organiser spontanément en différents niveaux d'oligomérisation : -dimère (2mer) -octamère (8mer) -hexadécamère (16mer). Lrp serait dans la cellule en équilibre entre l'état octamère et hexadécamère avec une prédominance d'hexadécamère (Chen et al., 2001). Il est possible que ces différents états oligomériques permettent d'influencer la régulation des gènes par Lrp, cependant les preuves expérimentales restent faibles.

Afin d'étudier plus en détails comment l'état oligomérique de Lrp affectent la régulation de l'*ori2* par *crtS*, nous avons construit différents mutants déjà caractérisés de Lrp : $\Delta C11$, D114E, L136R. Le mutant $\Delta C11$ forme uniquement des dimères, les mutants D114E et L136R forment majoritairement des octamères et moins d'hexadécamères (Platko and Calvo, 1993). Puis nous avons étudié l'impact de ces mutations sur la réplication d'un *pORI2* chez *E.coli*.

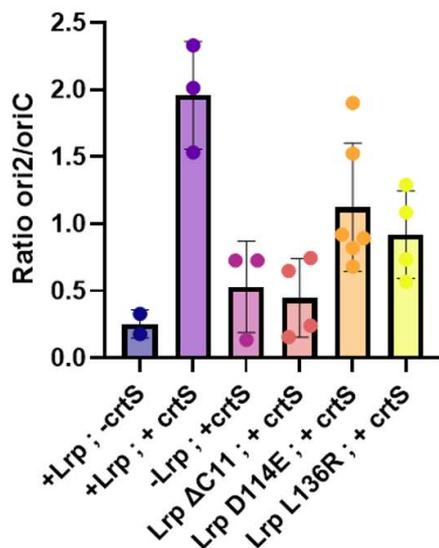


Figure 54. L'état d'oligomérisation de Lrp affect la réplication d'un *pORI2*. Mesure par dPCR du nombre de copies d'un *pORI2* dans une souche d'*E.coli* contenant une copie chromosomique du site *crtS* au locus *lacZ* ; ainsi qu'un allèle de Lrp (Δ , wt, $\Delta C11$, D114E, L136R).

En absence du site *crtS* ou de Lrp, le nombre de copies du *pORI2* est faible (environ 0,25), lorsqu'il y a *crtS* et Lrp le nombre de copies est d'environ 2. La protéine Lrp tronquée de 11 acides aminés *Lrp $\Delta C11$* réduit fortement la réplication du *pORI2*, suggérant dans ce contexte *crtS* est non fonctionnel. Chez les mutants *LrpD114E* et *LrpL136R*, réduisant la

formation d'hexadécamères, la réplication du *pORI2* est significativement diminuée par rapport à la souche Lrp-WT. Ainsi, il apparaît que le mécanisme *crtS-ori2* nécessite la formation d'hexadécamères de Lrp sur *crtS*. Des études supplémentaires seront nécessaires pour poursuivre l'étude du rôle de Lrp.

10. Lrp régule un tiers des gènes de *Vibrio cholerae*

Pour étudier le réseau de régulation de Lrp chez *Vibrio cholerae*, nous avons réalisé une analyse RNAseq dans les mêmes conditions que le ChIPseq : MH, culture exponentielle, DO=0.5. On trouve que le nombre de gènes dont l'expression est modifiée significativement (Fold Change > 2) dans une souche Δ lrp est 1019, 733 sont moins exprimés, 286 sont plus exprimés (Figure 55). Sur les 3620 gènes du génome 28,14% sont donc régulés par Lrp à la fois de manière directe et indirecte. Ce qui concorde avec ce qui a été démontré chez *E. coli* avec près d'un tiers du génome régulé par Lrp (Kroner *et al.*, 2019).

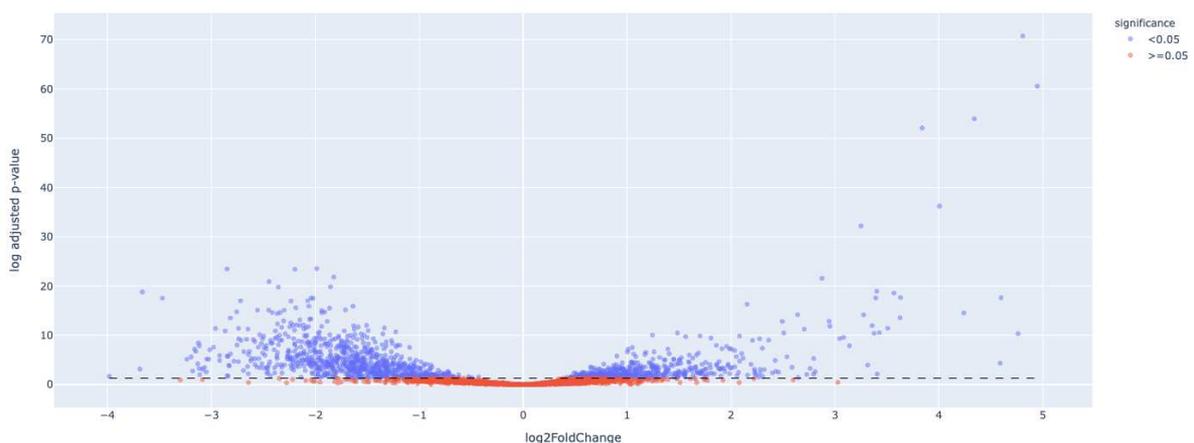


Figure 55. Analyse différentielle de RNAseq d'une souche Δ lrp. Analyse de RNAseq d'une souche Δ lrp comparé à une souche wt de *Vibrio cholerae*. L'expression différentielle des 3620 gènes du génome est représentée sous forme d'un volcano plot. En bleu sont représentés les gènes avec une différence d'expression significative ($p < 0.05$).

En combinant les deux jeux de données RNAseq et ChIPseq nous allons pouvoir déterminer finement le réseau de régulation de Lrp. Ces analyses bio-informatiques sont en cours, elles vont nous permettre de distinguer les régulations directes et indirectes de Lrp sur le génome de *Vibrio cholerae*.

Matériel et méthodes

PCR digitale

Les quantifications de (ori1 et ori2) chez *V. cholerae* et de (oriC et pORI2) chez *E. coli* ont été réalisées comme décrit dans (Fournes *et al.*, 2021) en utilisant la dPCR multiplex (Stilla Technologies). La dPCR a été réalisée directement sur le lysat cellulaire : 1 mL de culture d'une nuit a été lavé deux fois avec 1 mL de PBS, les culots ont été congelés à -20°C et remis en suspension dans 200 µL de PBS, les échantillons ont ensuite été bouillis pendant 10 min. Les échantillons ont été centrifugés pendant 10 minutes, le surnageant a été récupéré et la concentration en ADN a été mesurée à l'aide d'un appareil Qubit (ThermoFisher). Les réactions PCR ont été réalisées avec 0,1 ng d'ADN en utilisant le PerfeCTa® MultiPlex qPCR ToughMix® (Quantabio) sur une puce Sapphire (Stilla Technologies). La PCR numérique a été réalisée sur un Naica Geode et l'acquisition des images sur le lecteur Naica Prism3. Les images ont été analysées avec le logiciel Crystal Miner (Stilla Technologies). La dPCR a été réalisée en suivant les étapes suivantes : partition des gouttelettes (40°C, pression atmosphérique AP à + 950 mbar, 12 min), dénaturation initiale (95°C, +950 mbar, pendant 2 min), suivie de 45 cycles à (95°C pendant 10 s et 60°C pendant 30 s), libération des gouttelettes (down 25°C, down to AP, 33 min).

Microscopie électronique

La sonde d'ADN contenant le site *crtS* (944bp) a été produite par PCR à partir d'un plasmide pUC18::*crtS* en utilisant les amorces MV577/MV485 (voir tableau de l'article en écriture). Les fragments d'ADN ont ensuite été purifiés à l'aide de billes AMPure XP (Beckman Coulter). La protéine Lrp a été produite et purifiée comme dans PMC6212839 via un tag 6xHis-TEV-thioredoxine. Après purification nous avons obtenu la protéine Lrp à 178 µM avec une purification supérieure à 95%. Lrp a d'abord été dilué 180 fois dans un tampon 100 mM NaCl, 50 mM Hepes, 4 mM acétate de magnésium. Puis, pour la formation des complexes ADN-protéines, 200 nM de Lrp et 4 nM de la sonde ADN ont été mélangés dans un tampon Tris-HCl 10 mM pH 7.5, 50 mM NaCl, 1 mM MgCl₂ (volume final de 20 µL) pendant 10 min à température ambiante. Pour l'observation des complexes ADN-protéines, une goutte de 5 µL de l'incubation a été déposée sur une grille de cuivre de 600 mailles préalablement recouverte d'un mince film de carbone et préactivée par décharge lumineuse en présence d'amylamine (Sigma-Aldrich, France). Les grilles ont été rincées et colorées positivement avec de l'acétate d'uranyle aqueux à 2% (w/v), soigneusement séchées avec du papier filtre, et observées en mode champ sombre annulaire en imagerie filtrée sans perte à l'aide d'un microscope électronique à transmission Zeiss 902. Les images ont été capturées à un grossissement de 85 000x avec une caméra CCD Veleta MegaviewIII et analysées avec le logiciel iTEM (Olympus Soft Imaging Solution).

ChIPseq

ChIP-seq de Lrp a été réalisé comme décrit dans (Diaz *et al.*, 2017) dans *V. cholerae* exprimant Lrp-3xFLAG sous son promoteur natif. Nous avons confirmé que 3xFLAG ne perturbait pas l'activité de Lrp en mesurant par PCR digitale (dPCR) le nombre de copies du Chr2 (Fig.1 supplémentaire). Les cellules ont été cultivées dans 100mL de milieu Mueller Hinton (MH) sous agitation à 37°C jusqu'à ce qu'elles atteignent la phase de croissance mi-log (OD600 0,5). Les cellules ont ensuite été fixées par *crosslinking* avec du formaldéhyde à 1% pendant 30 minutes à température ambiante sous agitation, suivie d'un *quench* avec de la glycine 0,5 M pendant 15 minutes. Les bactéries ont été récoltées, lavées deux fois dans une solution saline tamponnée Tris (20 mM Tris/HCl pH 7,5, 150 mM NaCl) et les culots ont été conservés à -80°C. Les culots ont été lysés dans un tampon contenant du lysozyme et un inhibiteur de protéase (cOmplete Protease Inhibitor, Roche). La chromatine a ensuite été fragmentée par sonication (Covaris S220) dans un milliTUBE (1ml avec AFA Fiber). La fragmentation correcte de l'ADN, centrée sur 300 pb, a été confirmée par électrophorèse sur gel d'agarose à 1,8 %. Après sonication, les complexes Lrp-ADN ont été immunoprécipités avec des billes magnétiques anti-FLAG (M8823 Sigma). L'immunoprécipitation correcte a été confirmée par western blot en utilisant des anticorps anti-FLAG (F7425 Sigma). Une fois récupéré, l'ADN immunoprécipité a été dé-crosslinké à 65°C pendant une nuit. L'ADN a ensuite été réparé, ligué à des adaptateurs, amplifié par PCR et purifié à l'aide du kit TruSeq ChIP Library Preparation Kit (illumina) et des billes AMPureXP (Beckman Coulter). La qualité des librairies ChIP-seq a été mesurée par l'instrument Agilent Bioanalyzer 2100 à l'aide d'une puce DNA High Sensitivity. Enfin, les librairies ont été séquencées à l'aide d'un kit de sortie MiniSeq Mid sur une machine de séquençage Miniseq (Illumina). Les *reads* en pair-end de deux expériences indépendantes de ChIP-seq de RctB ont été mappés sur le génome de référence de *Vibrio cholerae* N16961 (chromosome 1 : CP028827.1 et chromosome 2 : CP028828.1) à l'aide de Bowtie2 en utilisant la plateforme galaxy.pasteur. L'identification des pics a été réalisée à l'aide de MACS2, où les lectures alignées étaient significativement enrichies par rapport à un contrôle (souche sans FLAG). Les pics identifiés ont été vérifiés à la main, en utilisant le navigateur de génome IGV pour confirmer l'attribution correcte des pics. Le navigateur en ligne MEME-ChIP (version 5.4.1) a été utilisé pour la recherche d'enrichissement de motifs. Pour la normalisation, les fichiers IP et INPUT ont d'abord été ajustés au même nombre de lectures à l'aide de R. Ensuite, la valeur de couverture pour chaque position dans le fichier IP a été divisée par la valeur de couverture moyenne d'une fenêtre coulissante de 2 kb du fichier INPUT.

Synchronisation de la réplication avec la souche *E.coli dnaC2*

Culture de nuit d'une souche de *Escherichia coli* MG1655 ; *dnaC2* ; *lrp ::lrp-3xFLAG* ; *lacZ ::crtS* à 30°C en M9 glucose 0,4%, acides aminés 0,5%. Culture de jour à 30°C au 1/100ème dans un milieu frais jusqu'à ce que les cellules atteignent une OD600 ~0.2, puis passage dans un bain marie avec agitation à 39°C pendant 2 heures. Après 2H, ajout de 16,5mL de milieu froid (4°C) et transfert dans un bain marie à 30°C pendant 6min. Après 6min, second transfert à 39°C pour permettre uniquement un seul cycle de réplication de repartir. Les cellules synchronisées sont alors fixées en suivant le protocole de ChIPseq à différents temps.

RNAseq

Des cultures de *Vibrio cholerae* ont été réalisées en milieu Lysogeny Broth (LB) à 37°C sous agitation (180 rpm) jusqu'à atteindre la phase exponentielle (OD600 ~0.6) en triplicat. Les cellules ont été récoltées par centrifugation. Les cellules ont ensuite été traitées avec du Trizol pour lyser les cellules et stabiliser l'ARN. L'extraction de l'ARN total a été effectuée en utilisant le RNeasy Mini Kit de Qiagen conformément aux instructions du fabricant. L'ARN extrait a ensuite été utilisé pour préparer les bibliothèques à la plateforme Biomix de l'institut Pasteur. Brièvement, les ARN ribosomiques ont été éliminés à l'aide d'une méthode custom de déplétion via oligonucléotides complémentaires, puis les ARNm restants ont été fragmentés et convertis en cDNA par transcription inverse. Les cDNA ont été ligués à des adaptateurs puis amplifiés par PCR. Les bibliothèques ont été séquencées sur un Illumina NextSeq, générant des reads d'une longueur de 150 pb. Les reads brutes ont été nettoyées et filtrées pour éliminer les adaptateurs et les reads de mauvaise qualité. Pour l'analyse des données, l'équipe du Hub de bioinformatique a utilisé le package R DESeq2. Les reads ont d'abord été alignés sur le génome de référence de *V. cholerae* à l'aide de BOWTIE2. DESeq2 a ensuite été utilisé pour normaliser et identifier les gènes différentiellement exprimés.

IV. Résultats non publiés : RctB et modification post-traductionnelle

La modification post-traductionnelle (ou PTM) est un processus qui se produit après la traduction et la synthèse d'une protéine. Elle implique la modification d'un ou plusieurs acides aminés de la protéine, ce qui peut altérer sa structure et sa fonction. Il peut s'agir de processus tels que la glycosylation, l'acétylation, la phosphorylation et l'ubiquitination, entre autres. Ces modifications sont importantes pour une variété de processus cellulaires, y compris le repliement des protéines, l'adressage vers des emplacements cellulaires spécifiques et les interactions protéine-protéine.

Pendant une longue période le consensus était qu'il y a peu de PTM chez les bactéries, surtout en comparaison avec les eucaryotes où celles-ci sont au cœur de très nombreux mécanismes cellulaires. Cependant, on détecte de plus en plus de PTM chez les bactéries grâce aux progrès des méthodes de spectrométrie de masse. Ces PTM interviennent dans des mécanismes majeurs de la cellule comme le métabolisme, la synthèse de protéines et le cycle cellulaire (Macek et al., 2019). Les modifications principales que l'on retrouve chez les bactéries sont : la phosphorylation, l'acétylation et la succinylation (Figure 56).

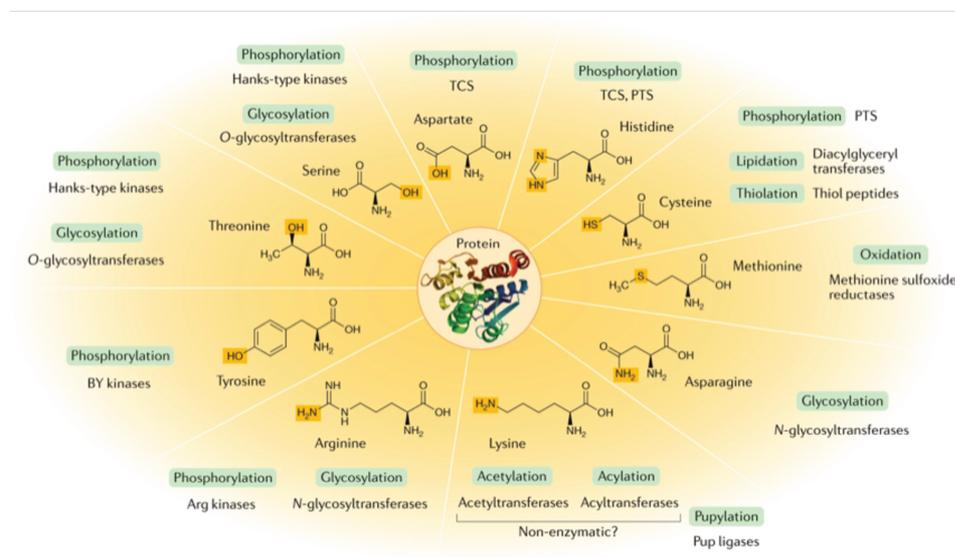


Figure 56. Modification Post Traductionnelle chez les Bactéries. De multiples résidus de protéines peuvent être modifiés de manière covalente. La phosphorylation des protéines (His, Asp, Arg, Ser, Thr, Tyr), l'acétylation (Lys, extrémités amino des protéines), la glycosylation (Asn, Arg, Ser, Thr), la lipylation (Cys), l'acylation (Lys), la thiolation (Cys) et l'oxydation (Met) sont les modifications post-traductionnelles des protéines les plus courantes (Macek et al., 2019).

Dans une étude récente, il a été démontré que l'initiateur DnaA est régulée par PTM chez *E.coli* (Zhang et al., 2016). En effet, la lysine K178 contenue dans le motif walker-A est acétylée de manière réversible. Lorsque la lysine K178 est acétylée DnaA ne fixe pas l'ATP ce qui inhibe sa fixation à *oriC*, c'est un mécanisme permettant de réguler négativement l'initiation de la réplication. Le niveau d'acétylation de DnaA est le plus fort lors de la phase stationnaire, là où les cellules ne se répliquent pas.

L'acétylation est une PTM particulière au sens où elle permet de changer un acide aminé chargé positivement en un acide aminé neutre. Ainsi cette PTM peut perturber les interactions protéine-protéine, protéine-ADN ou encore la conformation de la protéine (Figure 57).

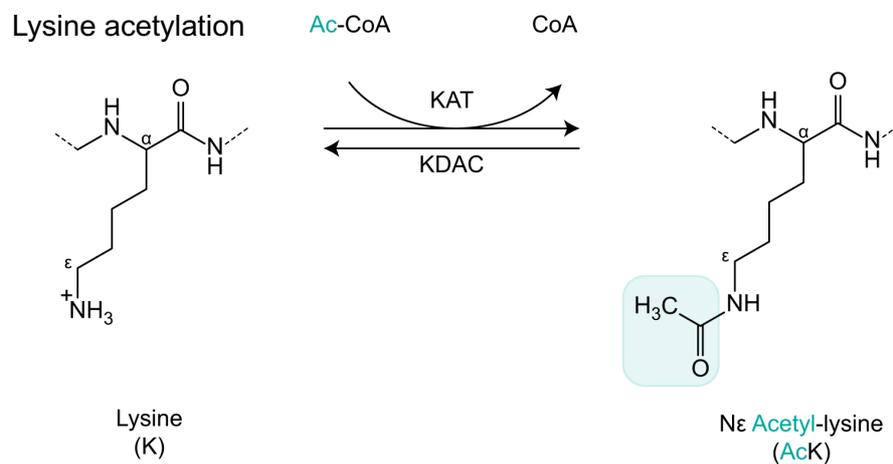


Figure 57. Réaction d'acétylation d'une lysine. Les acétyltransférases de lysine (KATs) catalysent le transfert d'un groupe acétyle (CH_3O , turquoise) de l'acétyl-CoA (Ac-CoA) au groupe ϵ -amino des chaînes latérales de lysine (K). La fixation covalente d'un groupe acétyle élimine la charge positive (+) du groupe amine. Le groupement acétyle peut être éliminé par les lysines déacétyltransférases (KDAC), c'est une modification réversible (Ree et al., 2018).

Nous nous sommes demandé si de manière analogue à DnaA, la protéine RctB pouvait subir des PTM. Notamment si celles-ci pouvaient avoir lieu lors de la duplication du site *crtS*, ce qui modifierait l'activité de RctB.

Une étude de l'acétylome global de *Vibrio cholerae* a été réalisé en 2017. Dans celle-ci, RctB apparait comme acétylé sur le résidu K404 (Jers et al., 2017). Ce résidu K404 a déjà été étudié car il médie un contact intra-moléculaire entre la partie C-ter et la région d'interaction avec les protéines chaperonnes DnaK/J (pos :150-163) (Jha et al., 2017), son acétylation peut donc potentiellement changer l'activité de RctB et notamment influencer la dimérisation. De plus,

l'acide aminé K404 est conservée à 100% parmi les *Vibrionaceae* (24 espèces). L'acétylation de ce résidu K404 semble particulièrement intéressant à caractériser.

Dans un premier temps, nous avons voulu confirmer les résultats de l'acétylome publié. Pour cela, nous avons réalisé un Western Blot anti-acétyl-lysine sur une immunoprécipitation anti FLAG de RctB (Figure 58).

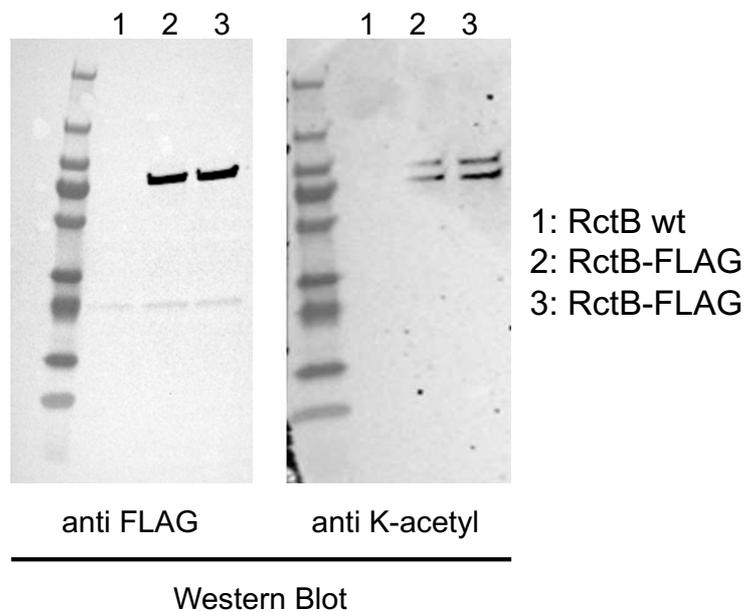
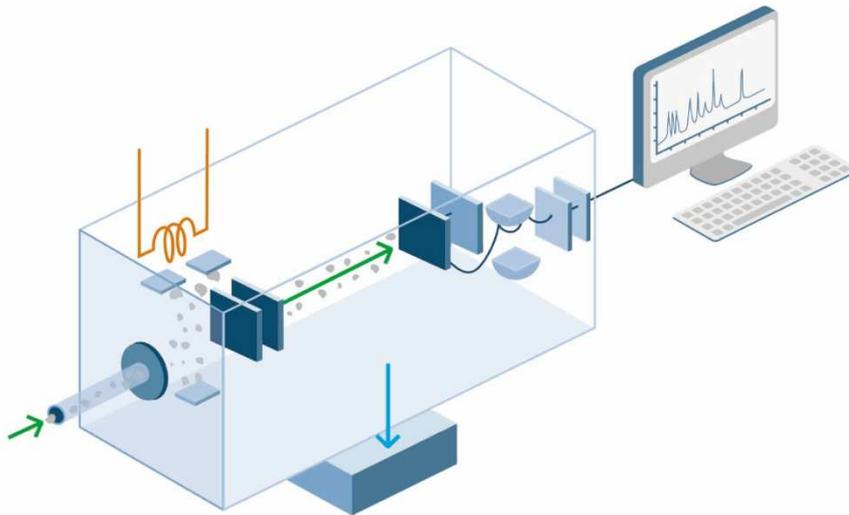


Figure 58. Western Blot pour détecter la potentielle acétylation de RctB. Western Blot anti FLAG (gel de gauche) ou anti acétyl-lysine (gel de droite) sur une immunoprécipitation réalisée sur une culture exponentielle d'une souche sauvage de *Vibrio cholerae* ou de deux souches exprimant RctB-FLAG. Dans le gel de droite, une bande sort au niveau de RctB suggérant que RctB contient une ou plusieurs lysines acétylées.

D'après les résultats de ce western blot, la protéine RctB est acétylée *in vivo*. Dans un second temps, nous avons décidé d'analyser RctB par spectrométrie de masse (MS) à la plateforme protéomique de l'Institut Pasteur dans le but de localiser la position de ces acétylations dans la séquence d'acides aminés. Pour cela nous avons réalisé une immunoprécipitation de RctB suivie d'une purification par découpe de bande dans un gel de protéine. Cet échantillon fut ensuite analysé par MS.



Acetyl-K sites: positions and intensity (from Acetyl (K)Sites table)

Proteins	AA-Positions	Localization prob	Acetyl (K) Probabilities	Intensity B-2	Intensity B-3	Intensity B-4
RctB-FLAG;Q9KNG2	K-10	1	LIK(1)LPR	6,61E+06	6,00E+06	6,72E+06
RctB-FLAG;Q9KNG2	K-42	1	GVTK(1)SIVELLNLI SLR	0,00E+00	0,00E+00	4,73E+06
RctB-FLAG;Q9KNG2	K-95	1	AAVAVGLFK(1)QIPVR	1,31E+07	1,72E+07	1,15E+07
RctB-FLAG;Q9KNG2	K-108	1	FE EGLAGK(1)TMLHR	2,04E+06	0,00E+00	0,00E+00
RctB-FLAG;Q9KNG2	K-144	1	SK(1)ALAQTVNQQR	5,92E+06	8,22E+06	8,85E+06
RctB-FLAG;Q9KNG2	K-200	1	SYQK(1)SMVSI SGTR	0,00E+00	1,15E+06	0,00E+00
RctB-FLAG;Q9KNG2	K-257	1	QAPNK(1)TPLYTDI LSLR	0,00E+00	3,87E+06	6,14E+06
RctB-FLAG;Q9KNG2	K-404	1	HTDVMMLSELNQK(1)LAR	0,00E+00	0,00E+00	7,12E+06
RctB-FLAG;Q9KNG2	K-454	1	GK(1)VVDYQVDIK	8,05E+06	4,86E+06	6,76E+06
RctB-FLAG;Q9KNG2	K-500	1	NELVSK(1)QK	5,85E+06	6,62E+06	7,73E+06

Figure 59. Spectrométrie de masse sur RctB pour détection des acétylations. Analyse de trois réplicas d'immunoprécipitation de RctB digérés à la trypsine. L'analyse des peptides recouvre 75% de la protéine (excellente couverture). 10 lysines de RctB sont détectées comme acétylées. De manière globale l'intensité des peptides acétylés a une valeur relativement faible en comparaison des peptides non modifiés de RctB.

A partir de l'analyse données de spectrométrie de masse sur RctB, 10 sites d'acétylation ont été détecté : K-10; K-42; K-95; K-108; K-144; K-200; K-257; K-404; K-454; K-500 (Figure 59).

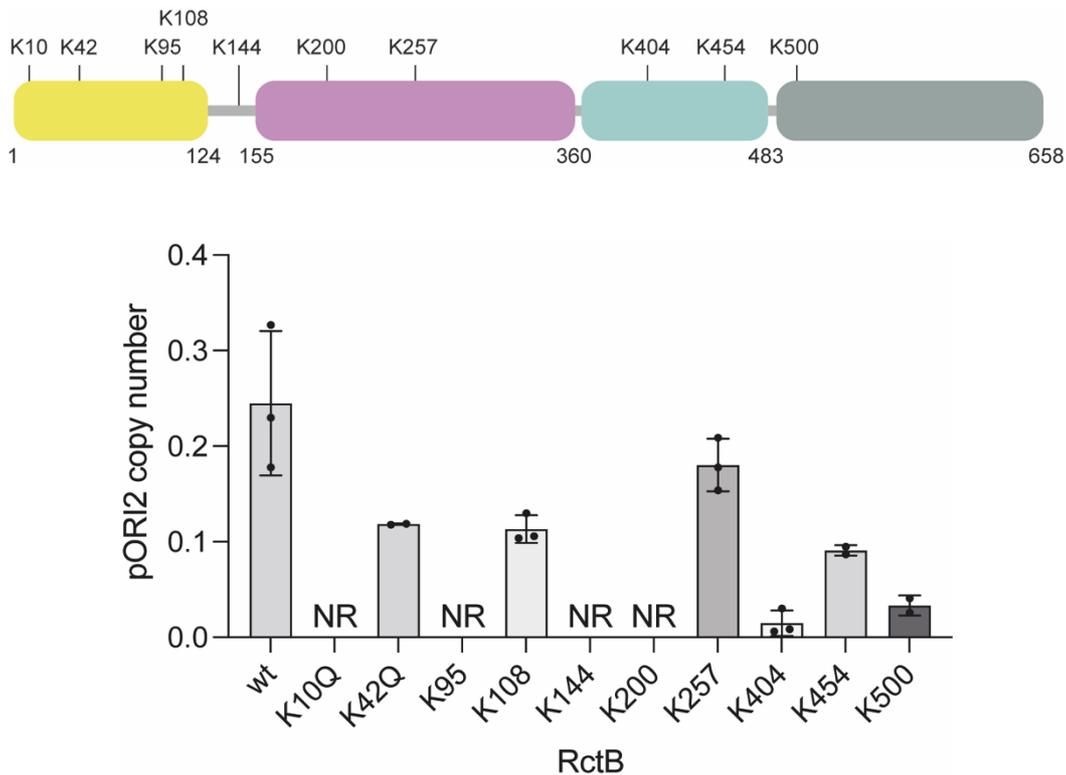


Figure 60. Distribution des différents sites d'acétylation sur les différents domaines de RctB. A. Les dix sites d'acétylation sont répartis dans les quatre domaines de RctB représentés avec des couleurs différentes. B. Mesure par dPCR du nombre de copies d'un plasmide pORI2 chez *E.coli* avec des allèles mutés de RctB.

Pour poursuivre cette analyse nous avons réalisé une mutagenèse dirigée de toutes les lysines identifiées sur RctB en les remplaçant une à une par une glutamine. En effet, une façon d'étudier l'acétylation des protéines est de muter les résidus de lysine en glutamine. Cette substitution permet de mimer l'effet de l'acétylation d'une lysine, c'est-à-dire de passer d'un résidu avec une charge positive à un avec une charge neutre, sans changer l'encombrement stérique. Ainsi ces mutants K>Q permettent de se placer dans un état « acétylation permanente » et d'étudier leur rôle dans la régulation de la fonction de la protéine.

Nous avons mesuré le nombre de copies d'un plasmide pORI2 portant le gène *rctB* sauvage ou les différentes versions mutées K>Q (Figure 60). Globalement ces mutations affectent négativement l'activité de RctB : soit elles diminuent le nombre de copies du plasmide (K42Q, K108Q, K257Q, K404Q, K454Q, K500Q) soit elles inactivent complètement RctB (K10Q, K95Q, K144Q, K200Q). On peut donc conclure que le rôle de l'acétylation de RctB est d'inhiber son activité, il pourrait donc s'agir d'un mécanisme analogue à celui constaté pour DnaA chez *E. coli*.

Lorsque nous avons commencé cette étude le but était de comprendre comment *crtS* active RctB. L'hypothèse était qu'après la duplication de *crtS*, RctB pouvait être modifié via PTM et ensuite activer l'*ori2*. Cependant, ici nous avons trouvé uniquement des modifications négatives qui ne nous permettront pas de résoudre le mécanisme *crtS-ori2*. Nous avons décidé de mettre cette étude en attente.

Matériels et méthodes

Western Blot anti acétyl-lysine

Les protéines totales ont d'abord été extraites d'une culture de *Vibrio cholerae* en phase exponentielle de croissance (DO=0.6). Les échantillons ont été préparés dans un tampon laemli et dénaturés à 95°C pendant 5min puis séparés par électrophorèse sur gel de polyacrylamide SDS-PAGE 4-16% (BioRad), puis transférés sur une membrane de nitrocellulose. Une fois le transfert effectué, la membrane a été bloquée pendant une heure à température ambiante dans une solution de PBS contenant 0,05% de Tween20 et 1% de Régilait. Pour la détection des protéines acétylées sur la lysine, nous avons utilisé l'anticorps primaire Anti-Acetylated Lysine produit chez le lapin (SAB5200090, Sigma-Aldrich) dilué à une concentration de 1:5000 (6µL dans 30mL) dans du PBS 0.05% Tween20 1% Régilait. La membrane a été incubée avec cet anticorps pendant la nuit à 4°C sous agitation constante. Dans un deuxième temps, après plusieurs lavages avec le PBS 0.05% Tween20, la membrane a été incubée pendant une heure à température ambiante avec l'anticorps secondaire Goat F(ab')₂ Anti-Rabbit IgG - H&L (HRP), pré-adsorbé (ab6013) (Abcam) également dilué à une concentration de 1:5000 dans la même solution. Nous avons également effectué une expérience de contrôle en utilisant l'anticorps primaire ANTI-FLAG® produit chez le lapin (F7425, Sigma-Aldrich), appliqué dans les mêmes conditions que l'anticorps Anti-Acetylated Lysine. Après incubation avec l'anticorps secondaire et les lavages appropriés, la détection a été réalisée à l'aide d'un système de détection par chimiluminescence et les signaux ont été capturés par autoradiographie.

Spectroscopie de Masse pour la détection des PTM

À la suite de l'immunoprécipitation de notre protéine d'intérêt, RctB, nous avons appliqué la spectroscopie de masse pour détecter les modifications post-traductionnelles, en particulier l'acétylation de lysine. Les échantillons immunoprécipités ont d'abord été soumis à une digestion par la trypsine pour générer des peptides qui peuvent être analysés par spectrométrie de masse. Cette digestion a été réalisée à 37°C pendant la nuit en utilisant une

quantité standard de trypsine. Les peptides obtenus ont ensuite été et analysés par spectrométrie de masse à l'aide d'un spectromètre de masse à trappe ionique linéaire couplé à une chromatographie liquide. L'analyse a été réalisée en mode de détection en tandem (MS/MS) afin de permettre une identification précise des peptides et des modifications. Nous avons utilisé une base de données de protéines pour identifier les séquences de peptides et leurs modifications post-traductionnelles. Les spectres de masse obtenus ont été recherchés dans la base de données en utilisant un algorithme de recherche spécifique, qui inclut une tolérance de masse pour les peptides et les ions de fragmentation, ainsi que l'acétylation de lysine comme modification variable. Pour confirmer l'acétylation, nous avons recherché une modification de masse de +42.01 Da correspondant à l'acétylation sur les résidus de lysine. Les résultats obtenus ont été validés par une analyse manuelle des spectres de masse, en recherchant la présence de pics de fragmentation y et b spécifiques pour les peptides acétylés.

V. Résultats non publiés : outil de synchronisation de la réplication chez *Vibrio cholerae*

En 2016, une étude (Wiktor et al., 2016) montre que l'on peut bloquer de manière robuste la réplication chez *E.coli* en ciblant l'*oriC* avec dCas9. La protéine dCas9 est un allèle catalytiquement inactivé de Cas9 qui se fixe à l'ADN mais qui ne fait pas de coupure double brin, cet allèle est principalement utilisé pour faire de l'interférence c'est-à-dire inhiber l'expression d'un gène cible.

Dans cette étude, les auteurs suggèrent que ce système pourrait être utilisé pour synchroniser les cellules. Le principe théorique est dans un premier temps de bloquer spécifiquement l'initiation de la réplication à *oriC* en ciblant avec un guide ARN une des boîtes DnaA de l'origine et en induisant une dCas9 (Figure 61). Les cellules peuvent alors terminer leur réplication en cours mais ne peuvent pas réinitier, au bout de x temps on se retrouve avec une population de cellules ayant toutes terminées leur réplication. Puis en tirant parti du fait que dCas9 est naturellement thermosensible chez *E.coli* on passe les cellules à 42°C et elles sont censées toutes initier leur réplication de manière synchrone.

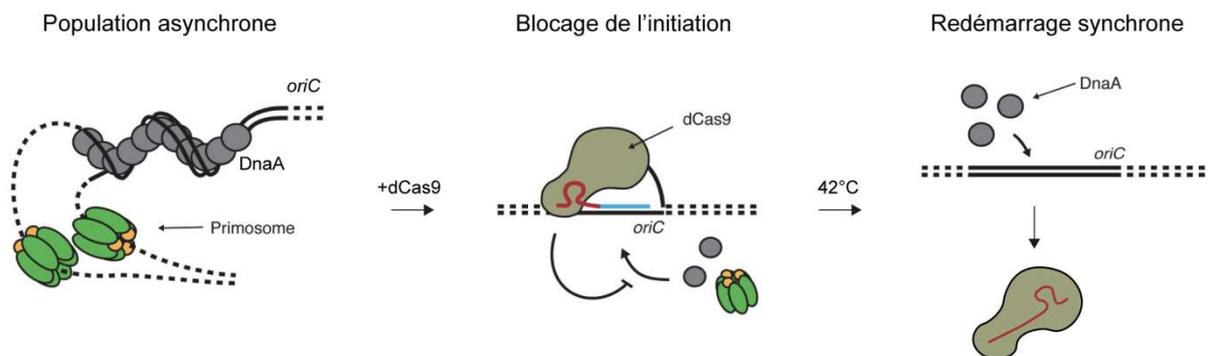


Figure 61. Principe de l'outil de synchronisation de la réplication via dCas9. Dans ce système développé chez *E.coli* une dCas9 ciblant l'*oriC* permet d'inhiber l'initiation de la réplication par encombrement stérique de l'origine et inhibition de la fixation de l'initiateur DnaA. Une fois l'origine bloquée et la réplication terminée, il est théoriquement possible de déstabiliser les dCas9 fixées à l'origine en augmentant la température à 42°C. Adapté de (Wiktor et al., 2016).

Contrairement à chez *E.coli* il n'y a pas de méthode robuste pour synchroniser la réplication chez *Vibrio cholerae*. Nous avons eu l'idée de réutiliser cet outil de blocage d'origine dans le but de créer un nouvel outil de synchronisation.

Nous partons d'un système donné par David Bikard qui regroupe sur un même plasmide : dCas9 sous le contrôle du promoteur *phf* inducible au DAPG (Vigouroux et al., 2018) - le guide (sgRNA) sous contrôle d'un promoteur constitutif.

La première étape a été de tester différents guides sgRNA ciblant l'*ori1* pour en sélectionner un qui bloque fortement l'initiation de la réplication (Figure 62).

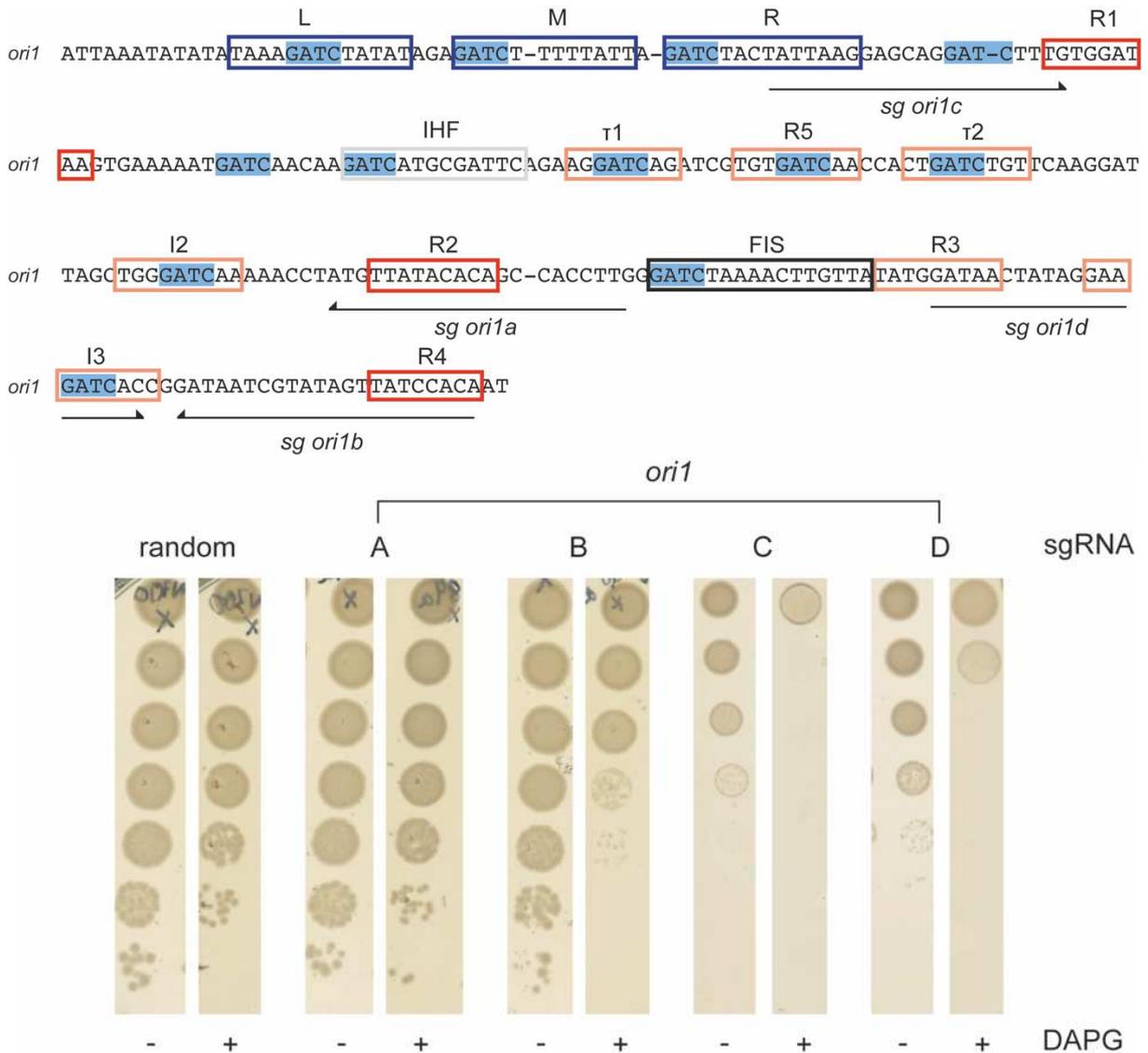


Figure 62. Test de l'efficacité de l'inhibition de croissance par dCas9 avec différents guides ciblant l'origine du Chr1 chez *Vibrio cholerae*. A. Séquence de l'*ori1* et détails des différents sites de fixation de DnaA/IHF/Fis. Les séquences cibles des différents guides sgRNA (sg *ori1a*, b, c ou d) sont représentées par des flèches noires. B. Test en goutte avec des dilutions successives d'une culture en phase stationnaire de *Vibrio cholerae* de 10^{-1} à 10^{-7} du haut vers le bas. Les gouttes sont déposées sur un milieu MH +/- inducteur (DAPG) et cultivées sur la nuit à 30°C. Le guide *random* ne cible aucune séquence du génome.

Nous avons ensuite sélectionné le sgRNA *ori1c* qui permet de bloquer de manière forte la réplication comme l'atteste le test en goutte (Figure 62). Cependant avec ce système

inductible au DAPG, on constate une fuite non négligeable du promoteur de la dCas9 en absence de l'inducteur. Ce phénomène de fuite est problématique car nous souhaitons créer un système propre ON/OFF sans perturber la réplication des cellules en conditions non induite. De plus un second problème est que la dCas9 est toujours active à 42°C chez *Vibrio cholerae* alors qu'elle ne l'est pas chez *E. coli*. Il est alors impossible d'utiliser ce système en l'état pour enlever l'inhibition de la réplication à 42°C (Figure 63).

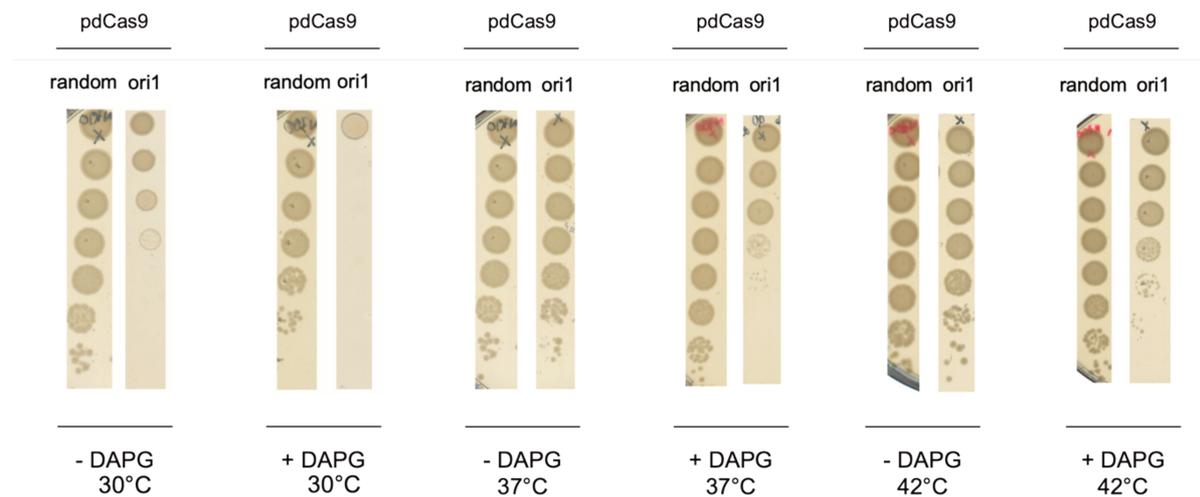


Figure 63. Test de la thermosensibilité de dCas9 chez *Vibrio cholerae*. A. Test en goutte d'efficacité de la dCas9 à différentes températures : 30, 37 et 42°C. Le guide utilisé ici est le guide *ori1c*.

Nous avons décidé d'adopter deux stratégies pour améliorer ce système. La première est de changer le promoteur pour un promoteur PBAD inductible à l'arabinose qui a l'avantage de présenter très peu de fuite. Puis nous allons également remplacer la dCas9 par un allèle thermosensible nommé tsRC9 isolé en 2016 dans le laboratoire d'Andreas Möglich (Richter et al., 2016). Grâce à ces deux éléments modifiés dans le plasmide, il semble que le système marche parfaitement en test à goutte (

Figure 64). C'est-à-dire une inhibition forte à 30°C, et pas d'inhibition à 42°C ainsi qu'une absence de fuite sans inducteur.

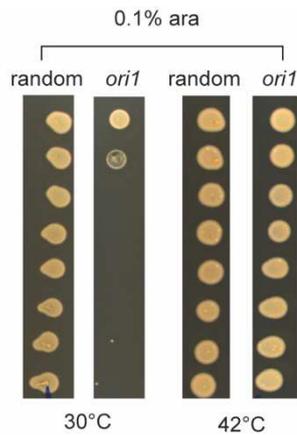


Figure 64. Inhibition robuste de la réplication de *V.cholerae* avec une dCas9 ciblant l'*ori1*. Test en goutte de l'efficacité du système modifié avec un allèle thermosensible de dCas9 (tsRC9) et un promoteur PBAD (inductible avec 0.1% arabinose) pour bloquer la croissance bactérienne

L'étape suivante était d'appliquer ce système à une culture en croissance exponentielle de *Vibrio cholerae* à 30°C pour tester le blocage de la réplication. Cependant en courbe de croissance nous n'observons aucune différence après induction entre les cultures avec un guide random et celles avec un guide ciblant *ori1*. Nous avons par la suite essayé de changer de milieu de culture (LB, M9) ou encore d'augmenter la concentration d'inducteur et pourtant nous étions toujours incapables d'obtenir une inhibition robuste de la croissance des cellules avec ce système (Figure 65).

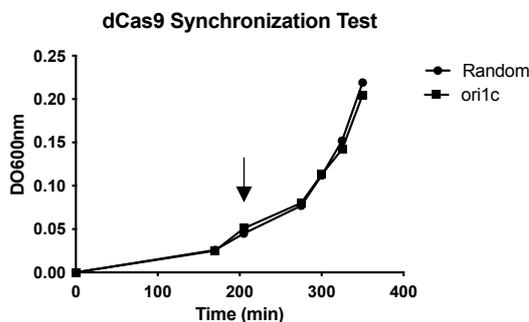


Figure 65. Courbe de croissance avec induction du système dCas9 anti *ori1*. Courbe de croissance d'une culture de *Vibrio cholerae* avec induction de dCas9 à T=200min (flèche noire).

Pour mieux comprendre ce qu'il se passe dans une culture en phase exponentielle après induction du système, nous avons observé les cellules par microscopie. Après deux heures d'incubation en présence d'inducteur, on observe une très grande hétérogénéité dans la population de cellules. La population est composée d'un mélange cellules filamenteuses et de cellules de taille normal. La formation de cellules filamenteuses attestant d'un blocage de la réplication de l'ADN (Figure 66).

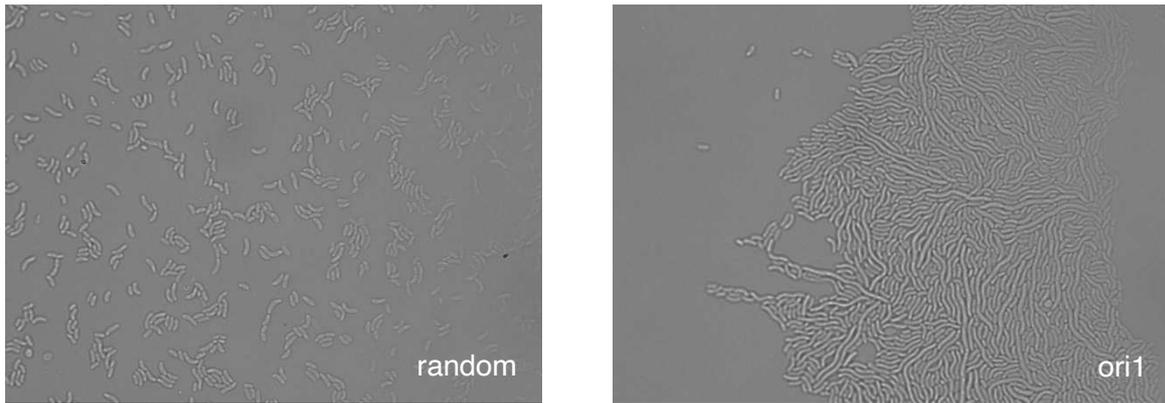


Figure 66. Différence majeure d'efficacité du blocage de l'initiation par dCas9 selon les conditions de croissance. Images de microscopie d'une culture de *Vibrio cholerae* après deux heures d'induction de dCas9 avec un guide random ou un guide *ori1*. La filamentation observée avec le guide *ori1* est synonyme de blocage de la réplication.

Cet outil de synchronisation de la réplication basé sur dCas9 semblait prometteur avec notamment les résultats en test de goutte en milieu solide qui étaient particulièrement convaincant. Cependant, en pratique ce système semble très compliqué à optimiser car il y a une forte hétérogénéité dans la population de cellules après l'induction du système. Nous avons finalement choisi d'adopter d'autres stratégies préexistantes pour nos expériences en de ChIPseq en population synchronisée : via SHX chez *V.cholerae* ou via un système hétérologue *E.coli dnaC2*.

CONCLUSIONS

Conclusions générales

Le but de ces travaux de thèse était de mieux comprendre le mécanisme de coordination de la réplication des chromosomes de *Vibrio cholerae*. En commençant nous savions que RctB, l'initiateur de la réplication du Chr2, présente différentes activités en fonction de son site de liaison : il agit comme un initiateur sur les itérons, un répresseur sur les sites 39m et un déclencheur sur le site *crtS*.

Notre première étude publiée (Fournes *et al.*, 2021) a permis de fournir de nouvelles informations sur la façon dont l'initiation de la réplication de Chr2 est régulée par *crtS* via RctB. Nous avons observé que lorsqu'on inactive un ou plusieurs sites 39mer présents sur l'*ori2*, le Chr2 sur-initie sa réplication et devient insensible à *crtS*. Nous avons également montré que la fixation de RctB aux sites 39m nécessite son domaine IV et que ce dernier est un domaine d'oligomérisation. Lorsque le domaine IV est inactivé, l'initiation de l'*ori2* devient également insensible à *crtS*. Ainsi nous proposons un modèle dans lequel *crtS* agit comme un site anti-inhibiteur, empêchant les sites 39m de l'*ori2* de réprimer l'initiation de la réplication.

Afin d'étudier en détail la liaison de RctB au niveau du génome, nous avons réalisé des expériences de ChIPseq sur RctB dans une seconde étude. Notre analyse a révélé que RctB se lie uniquement aux sites 39m sur l'*ori2*, qui contient pourtant à la fois des 39m et des itérons. Cela indique qu'*ori2* est inhibée pendant la majorité du temps. Pour examiner la liaison dynamique de RctB tout au long du cycle cellulaire, nous avons réalisé des expériences supplémentaires de ChIPseq sur une population synchronisée à différents temps. Nos résultats ont montré que le profil de liaison de RctB change pendant le cycle cellulaire, en effet après la réplication de *crtS* RctB passe des sites inhibiteurs aux sites activateurs (itérons). De plus, lorsque l'on retire le site *crtS* cela augmente spécifiquement le signal de RctB sur les sites inhibiteurs de l'*ori2*. Cela vient confirmer le modèle de fonctionnement d'anti-inhibiteur pour *crtS*.

Par ailleurs, nous avons résolu la structure du domaine IV de RctB et démontré son implication dans l'inhibition de la réplication du Chr2. Le domaine IV est un domaine d'oligomérisation qui semble contenir deux zones d'interaction, une forte et une plus faible. Ces différentes zones d'interaction sont cruciales à la fois pour la fixation stable aux 39mer et la formation du complexe d'inhibition. Le domaine IV n'est cependant pas important ni pour la

fixation à *crtS* et ni pour celle aux itérons. Lorsque ce domaine est inactivé, RctB peut se fixer librement sur les itérons de l'origine et initier la réplication de manière asynchrone.

Dans l'ensemble, ces résultats soutiennent un modèle où *crtS* déstabilise de manière directe un complexe d'inhibition qui verrouille l'origine de réplication du Chr2, permettant la coordination de la réplication de Chr1 et Chr2.

Nous avons ensuite porté le focus sur le site *crtS* et notamment le rôle crucial de Lrp. Nous avons démontré que Lrp se fixe à *crtS* et forme une structure nucléoprotéique ressemblant à celle des histones eucaryotes, dans laquelle l'ADN s'entoure autour de Lrp. Le site minimal de *crtS* de 62pb contient d'ailleurs deux sites de fixation pour Lrp. En étudiant une souche Δlrp , le phénotype est très proche d'une souche $\Delta crtS$, les cellules filamentent et sous-initient leur Chr2. Ceci s'explique car Lrp est essentiel pour la fixation de RctB à *crtS*, sans Lrp le site *crtS* devient non fonctionnel. Le changement de conformation de l'ADN induit par Lrp permet probablement à RctB de reconnaître *crtS*.

De plus, la fixation de Lrp sur *crtS* semble varier avec la phase de croissance. Une expérience faite en population synchronisée montre que la fixation de Lrp sur *crtS* fluctue avec le cycle cellulaire, Lrp semble être recruté après le passage de la fourche de réplication. Des résultats préliminaires semblent indiquer que c'est le niveau d'oligomérisation de Lrp qui permettait cette variation et notamment la formation d'hexadécamère.

Ainsi, ces résultats semblent indiquer que le rôle de Lrp serait double, Lrp permettrait à la fois le recrutement de RctB sur *crtS* et de sentir le passage de la fourche de réplication.

DISCUSSION

Discussion

Nous commençons à avoir une bonne vision de l'activation de l'*ori2* après la réplication de *crtS*. L'origine du Chr2 est inactivée par handcuffing pendant la majorité du cycle cellulaire car l'initiateur est fixé uniquement aux sites inactivateurs 39m. Puis lorsque *crtS* est répliquée, il y a changement du profil de fixation au niveau de l'origine avec RctB qui se fixe aux sites activateurs : les itérons ainsi qu'à la région DUE permettant l'ouverture de l'ADN et le recrutement du réplisome.

Malgré les efforts déployés, plusieurs questions demeurent toujours sans réponse à ce jour. Nous discuterons ici du rôle du contact *ori2-crtS*, du senseur de la fourche, de l'organisation du complexe d'inhibition. Nous prendrons également un peu de recul et nous jetterons un regard sur nos découvertes sous l'angle de l'évolution et de la phylogénie bactérienne.

***crtS-ori2* : une action directe et locale ?**

Plusieurs arguments permettent de soutenir le fait que l'action de *crtS* sur l'*ori2* se fait de manière directe. Par mesure de contact de chromosomes (HiC), les régions *ori2* et *crtS* sont en contact préférentiel dans la cellule en phase exponentielle de croissance (Val *et al.*, 2016). De plus, par microscopie à fluorescence, on observe également une colocalisation de *crtS* et d'*ori2* (Val *et al.*, 2016). Les deux sites sont donc proches physiquement dans la cellule.

D'un point de vue fonctionnel, en étudiant une souche avec deux copies de *crtS*, une proche de l'*ori1* et une à la position native sur le Chr1, nous observons une activation séquentielle d'*ori2*. En effet, dans cette souche les cellules naissent avec deux copies du Chr2, puis passent à trois puis quatre copies. Ces résultats indiquent que la réplication d'un site *crtS* est capable d'activer une seule *ori2*. Ce résultat a également été déduit mathématiquement à partir de l'étude par MFA de cette même souche à deux *crtS* par Ole Skovgaard.

Nous avons également observé directement par microscopie électronique (TEM) les complexes nucléoprotéiques formés par RctB en présence de deux sondes d'ADN contenant respectivement *crtS* et *ori2*. Certaines images suggèrent un contact direct entre les deux sondes d'ADN médiés par RctB. Cependant, lors de ces expérimentations nous n'avons pas ajouté Lrp, ce qui pourrait augmenter la fréquence de contact. D'autres approches plus

résolutives (AFM) ou quantitatives (FRET) permettraient de démontrer formellement ces contacts.

Nous pensons que pour la suite du projet il sera indispensable d'obtenir la structure de RctB sur ses différents sites de fixation : itérons, 39m et *crtS*. La présence de deux régions *linker* intrinsèquement désordonnées entre les domaines I et II ainsi qu'entre les domaines III et IV rendent particulièrement difficile l'obtention de cristaux pour faire de la diffraction. En collaboration avec l'équipe d'Abel Garcia Pino, nous allons réaliser l'observation du complexe nucléoprotéique RctB-Lrp-*ori2* par CryoEM (en cours).

Dans le cas des initiateurs de type Rep, c'est l'obtention des structures de RepE en dimère libres, en monomère fixé à un itéron, et en dimère sur l'opérateur qui a permis de comprendre les mécanismes de régulation de ces protéines. Le domaine NTD est une interface de dimérisation lorsque RepE est libre ou fixé sur l'opérateur. La monomérisation de RepE entraîne un changement conformationnel dans le domaine NTD le transformant en domaine de fixation à l'ADN.

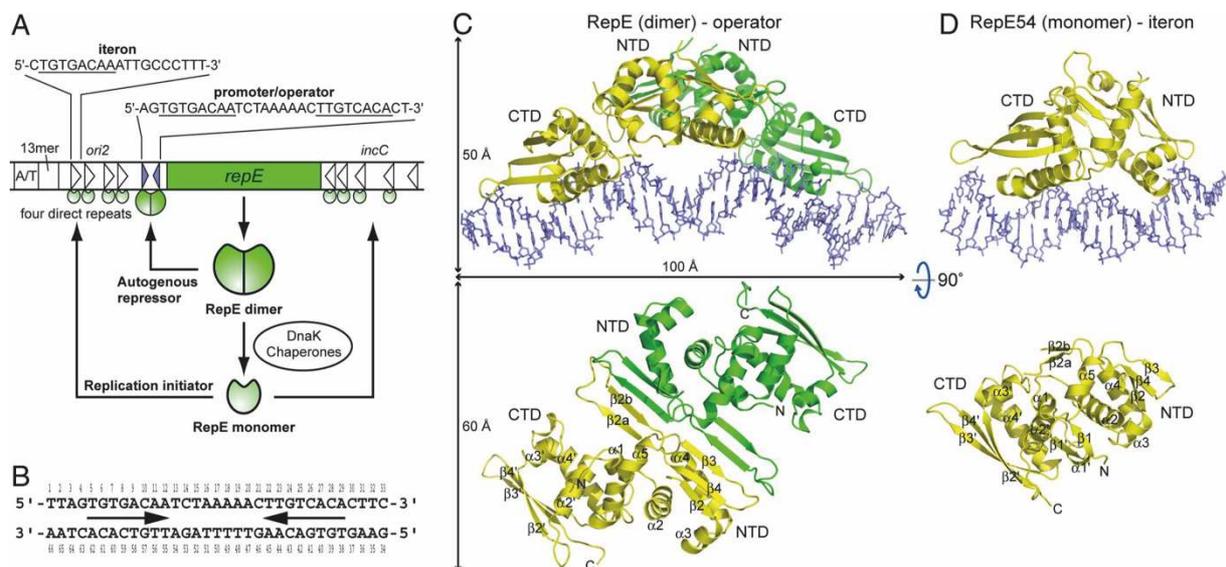


Figure 67. RepE change de conformation selon ses sites de fixation. (Nakamura *et al.*, 2007)

Sachant que RctB possède quatre domaines structuraux, que selon les sites de fixation il n'utilise pas les même HTH (HTH1, 2, 3 pour les itérons, HTH2 pour 39m, HTH2 pour *crtS*), qu'ils possèdent de longues régions linker, des changements importants de conformations entre chaque site de fixation sont très probables.

Quel rôle pour le passage de la fourche sur *crtS* ?

Nous savons que c'est la réplication du site *crtS* qui permet de déclencher l'initiation de l'*ori2*, en atteste l'expérience dans laquelle *crtS* est déplacé proche de l'*ori1* entraînant par conséquent une réplication précoce de l'*ori2* (Val *et al.*, 2016). Cependant le changement qu'entraîne le passage de la fourche à *crtS* permettant d'envoyer le signal d'initiation est encore mal compris.

La fourche de réplication peut jouer un rôle passif en provoquant : - le doublement du nombre de copies du site – le passage des motifs GATC d'un état totalement méthylé à un état hémiméthylé. Cependant, elle peut aussi jouer un rôle actif en induisant un changement dans la topologie de l'ADN, en effet, l'ouverture du duplex d'ADN entraîne l'apparition de super tours positifs en avant de la fourche de réplication au fur et à mesure de sa progression (Postow *et al.*, 2001). La fourche peut également retirer transitoirement des protéines attachées à l'ADN (Arias-Cartin *et al.*, 2017). Enfin, la réplication favorise la structuration de l'ADN simple brin en *hairpin* entre les fragments d'Okazaki (Loot *et al.*, 2010).

Certaines de ces conséquences du passage de la fourche de réplication ont été testées pour *crtS*. Premièrement, *crtS* est entouré de huit sites de méthylation GATC dont certains très conservés, tous en dehors du site minimal de 62pb. Nous savons que la liaison de RctB est sensible à l'état de méthylation des itérons, ceux-ci doivent être entièrement méthylés pour être reconnus. Le passage de la fourche de réplication sur *crtS* générerait des sites GATC hémiméthylés transitoirement qui pourraient affecter la liaison de RctB. Cependant, lorsqu'on inactive ces sites GATC il n'y a aucun impact sur l'activité de *crtS* (Val *et al.*, 2016), (de Lemos Martins *et al.*, 2018). Le rôle de ces sites GATC reste inexploité pour le moment.

Grâce à nos expériences de ChIPseq nous savons que RctB est toujours présent sur le site *crtS*, que ce soit en phase exponentielle/stationnaire ou tout le long du cycle cellulaire. Ainsi la réplication du site *crtS* n'a pas pour rôle de recruter RctB au moment opportun. Cependant, on observe clairement une variation dans la fixation de Lrp à *crtS*. En phase stationnaire la fixation de Lrp à *crtS* est réduite, et dans une population synchronisée il semble que la fixation de Lrp augmente après le passage de la fourche de réplication.

Même si Lrp est toujours fixé sur *crtS* dans toutes les conditions testées, son niveau de fixation semble pouvoir varier avec la réplication. Le passage de la fourche pourrait

permettre un changement dans le complexe nucléoprotéique Lrp-*crtS*-RctB d'une forme inactive à une forme active, probablement en influant sur le niveau d'oligomérisation de la protéine Lrp et possiblement via la formation d'hexadécamère (interaction entre deux octamères).

La protéine Lrp a été beaucoup étudiée mais très peu dans un contexte de réplication. Il est possible qu'un hexadécamère de Lrp se forme via une interaction entre deux chromatides sœurs après le passage de la fourche. Ce contact entre chromatides sœurs se ferait via une interaction entre deux octamères de Lrp fixés à l'ADN, cette piste n'a encore jamais été explorée pour Lrp. Nous pourrions tester ce contact avec des outils permettant de quantifier le contact entre chromatide sœurs (HiSC2 ou autre). Nous savons par ailleurs que d'autres NAPS dont H-NS permettent la cohésion entre chromatides sœurs.

Organisation du complexe d'inhibition de l'*ori2*

Contrairement aux plasmides à itérons R6K, P1 ou RK2 il n'y a pas de *handcuffing* intermoléculaire sur l'*ori2*. En effet, juste après l'initiation du Chr2, ParAB2 ségrège les deux copies naissantes de chaque côté de la cellule. Les origines ne sont donc pas en contact dans la cellule. Ainsi, nous pensons que les modèles de *handcuffing* dans lequel deux *ori2* sont couplés par RctB (notamment décrit dans (Venkova-Canova and Chattoraj, 2011)) sont faux.

Le *handcuffing* sur le Chr2 est intramoléculaire entre la région *ori2_{min}* et la région *inc*. Dans le cas du plasmide F, le *handcuffing* intramoléculaire se fait entre les itérons de la région *incF* et ceux de la région *oriV*. Dans le modèle du complexe d'inhibition, il est formé à partir de deux monomères de RepE fixés chacun à un itéron et d'un dimère faisant le pont. Le dimère est formé via une interface forte et il interagit avec le monomère fixé à l'ADN via une interface faible.

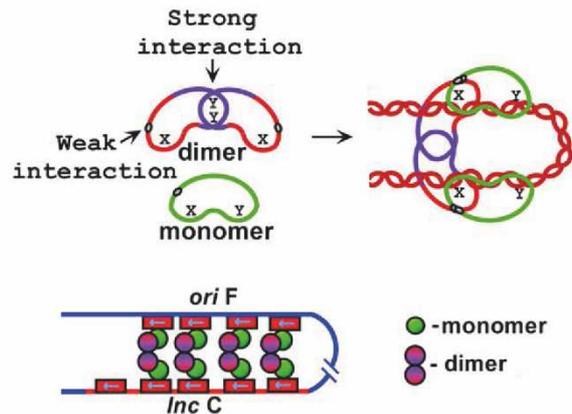


Figure 68 : Modèle du complexe d'inhibition d'oriV formé par RepE. (Zzaman and Bastia, 2005)

Dans le cas de l'*ori2*, nous proposons un nouveau modèle dans lequel le *handcuffing* de l'*ori2* se fait entre les sites 39m de la région *inc* et le 29m dans le promoteur de RctB. Ce modèle s'appuie sur les résultats du ChIPseq de RctB où l'on observe une fixation uniquement sur les trois 39m de l'*ori2* pendant la majorité du cycle cellulaire. Il s'appuie également sur la microscopie électronique où RctB est capable de former des boucles sur une molécule d'ADN avec uniquement les trois sites 39m. Cependant, des données structurales sur ce *handcuffing* de l'*ori2* sont indispensables pour comprendre le mécanisme complet.

Le domaine IV est au cœur de la formation de ce complexe d'inhibition, sans lui RctB est incapable de fixer les sites 39m. Nous possédons à présent la structure du domaine IV seul et il serait un domaine d'oligomérisation. En effet, ce dernier se divise en deux sous domaines contenant chacun une interface de dimérisation. Les deux interfaces semblent être de forces d'interaction différentes, avec une zone de dimérisation forte en Nter et une zone de dimérisation faible en Cter. La mutation L651P qui abolit la fixation de RctB aux 39m est situé en plein milieu de la zone de dimérisation faible. Nous pensons que RctB se fixe aux 39m uniquement sous la forme d'un dimère en interagissant via cette interface de dimérisation faible. L'interface de dimérisation forte permettrait de bridger l'ADN en interagissant avec d'autres molécules de RctB.

On ne peut pas exclure que le dimère de RctB interagissant via son domaine II soit impliqué dans le complexe d'inhibition. D'ailleurs, une étude récente a montré que des mutants monomériques dans l'interface de dimérisation du domaine 2 (D292R, D314P, A320R et S313W/D314W) rendent le nombre de copie d'un p*ORI2* insensible à la présence de *crtS* (Kothapalli et al., 2022). Une hypothèse permettant d'expliquer cette observation est que le *handcuffing* intramoléculaire médié par les 39m implique aussi la forme dimérique de RctB

(interagissant via le domaine II) comme c'est le cas pour RepE dans le plasmide F. Ainsi, en l'absence d'une forme dimérique de RctB, les 39mer seraient incapables de former le complexe d'inhibition et rendrait par conséquent le p*ORI2* insensible à *crtS*.

Comment dissocier le complexe d'inhibition?

Nous savons maintenant grâce à des expériences de ChIPseq que le domaine d'oligomérisation de RctB (le domaine IV) n'est pas nécessaire pour la fixation au site *crtS*. Nous pensons que le domaine IV de RctB fixé à *crtS* est libre pour interagir avec d'autres RctB et plus particulièrement ceux fixés sur les sites 39m de l'*ori2*.

Nous observons également dans une souche Δ *crtS* une augmentation spécifique du signal de RctB sur les sites 39m de l'origine (figure 3, 2nd article). Ce résultat suggère que *crtS* vient décrocher les RctB fixé sur les sites 39m. Cependant dans l'expérience de ChIPseq de RctB en population synchronisée on observe après la réplication de *crtS* à la fois des pics sur les 39mer et sur les itérons/DUE. Cette seconde expérience suggère plutôt un modèle dans lequel RctB fixé sur *crtS* interagit avec RctB fixé sur 39m empêchant l'handcuffing.

Le *handcuffing* est le mécanisme principal permettant l'inhibition de la réplication des plasmides à itérons. Pour les plasmides à itérons présent en plusieurs copies dans la cellule, c'est un mécanisme qui permet de bloquer l'initiation via des interactions inter ou intramoléculaires entre les protéines Rep fixées sur des itérons.

Dans le cas du plasmide à itérons RK2, il apparaît que le complexe d'inhibition formé par l'initiateur TrfA (de la famille Rep) entre deux molécules d'ADN est hautement stable, maintenant ainsi fermement l'origine du plasmide dans un état verrouillé. Pour défaire ce complexe d'inhibition et re-initier la réplication, des mécanismes actifs de dissociation nommés « uncuffing » sont nécessaires.

Au moins deux mécanismes participent au *uncuffing*, le premier requiert la présence de protéases (ClpAP et Lon) qui vont venir spécifiquement dégrader les initiateurs formant le complexe d'inhibition. A noter que le plasmide RK2 est instable dans une souche déficiente pour Lon, probablement en raison d'une origine trop inhibée. Le second mécanisme, est la présence de TrfA monomérique dans la cellule qui vont venir interagir avec les TrfA impliqués dans le complexe d'inhibition et les décrochés de celui-ci (Bury *et al.*, 2017).

Dans le cas du mécanisme *crtS-ori2* chez *Vibrio cholerae*, nous pensons que les molécules de RctB fixées à *crtS* peuvent interagir avec celles sur les 39m permettant ainsi le *uncuffing*. Il serait intéressant de savoir si des protéases tel que Lon ou Clp participent à ce mécanisme.

A la recherche de l'ancêtre du Chr2 : mégaplasmites portant des initiateurs RctB-like

Tous les membres de la famille *Vibrionaceae* ont un chromosome secondaire initié par RctB, cependant il y a aussi des espèces bactériennes dans au moins deux groupes extérieurs aux *Vibrionaceae* qui possèdent un mégaplasme initié par des protéines RctB-like (contenant notamment le domaine DUF3346). C'est le cas chez les *Enterobacterales* avec *Plesiomonas shigelloides* ou chez les *Alteromonadales* avec *Ferrimonas marina*. Ces protéines RctB-like ont approximativement la même taille que RctB de *V.cholerae*. Lorsque nous alignons la structure modèle de RctB et de ces protéines de type RctB, nous trouvons un alignement presque complet.

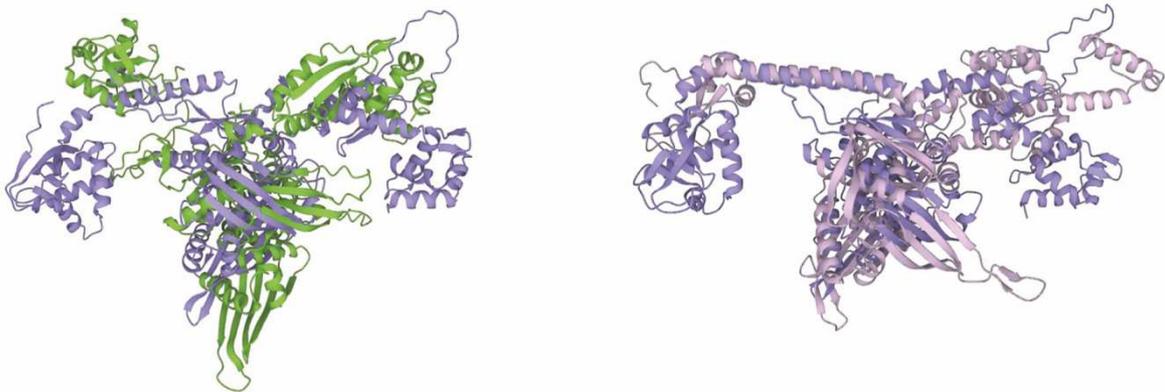


Figure. Alignement RctB et RctB-like. Alignement de structure de RctB (violet) avec des RctB-like de *P.shigelloides* (vert) et *F.marina* (rose).

Le modèle de réplication de *P. shigelloides* a été étudié récemment (Adam et al., 2022). La terminaison de la réplication du mégaplasme de 299kbp n'est pas synchrone avec le chromosome principal un, ce qui démontre qu'il n'y a pas de mécanisme de coordination. D'un point de vue évolutif, la structure actuelle de RctB_{Vch} avec quatre domaines semble précéder l'apparition de la synchronisation de la terminaison médiée par *crtS*.

A noter que la protéine RctB-like qui permet l'initiation du mégaplasme de *P.shigelloides* se divise également en quatre domaines structuraux. L'apparition du mécanisme de terminaison synchrone médié par *crtS* chez les *Vibrionaceae* ne semble donc pas avoir nécessité de changement majeur dans la protéine RctB.

Terminaison synchrone : un mécanisme commun aux chromides ?

La terminaison synchrone de la réplication du chromosome principal et du chromosome secondaire semble conservée chez tous les *Vibrionaceae* car il a été confirmé pour de nombreuses espèces de *Vibrio* (Kemter *et al.*, 2018). Par ailleurs, il semble que pour d'autres espèces éloignées phylogénétiquement des *Vibrionaceae* et contenant des chromides, on observe également une synchronisation de la terminaison de la réplication.

C'est notamment le cas pour *Agrobacterium tumefaciens* qui contient un chromosome secondaire avec une origine de réplication de type RepABC. En effet, dans une étude récente les auteurs ont étudié le profil de réplication des réplicons de *A.tumefaciens* par MFA et ont révélé que le Chr1 et le Chr2 terminent leur réplication de manière synchrone (Ren *et al.*, 2022). Nous pouvons nous demander comment est régulée cette terminaison synchrone et si elle est conservée pour les chromides de type RepABC. Si tel est le cas, nous serions face à exemple de convergence évolutive entre les deux grands types de chromides.

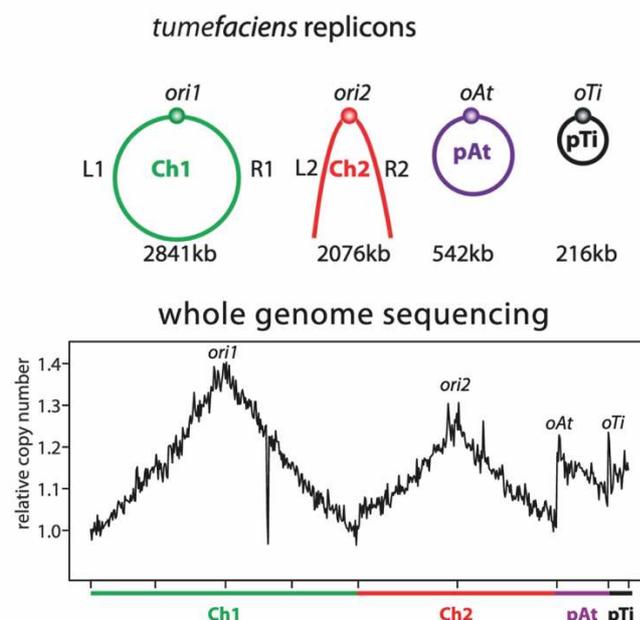


Figure. Étude du profil de réplication d'*Agrobacterium tumefaciens* par MFA. (Ren *et al.*, 2022).

En outre, la présence d'un *timer* sur le chromosome principal apparait comme une solution particulièrement robuste pour permettre une terminaison de la réplication synchrone avec le chromide. Il serait intéressant d'utiliser la méthode de ChIPseq sur les initiateurs de la réplication d'autres chromides pour voir s'ils se fixent sur le chromosome principal et ainsi découvrir des sites « *crtS-like* » chez ces organismes.

Références

- Abeles, A.L., and Austin, S.J. (1987). P1 plasmid replication requires methylated DNA. *EMBO J* 6, 3185-3189. 10.1002/j.1460-2075.1987.tb02630.x.
- Adam, Y., Brezellec, P., Espinosa, E., Besombes, A., Naquin, D., Paly, E., Possoz, C., van Dijk, E., Barre, F.X., and Ferat, J.L. (2022). *Plesiomonas shigelloides*, an Atypical Enterobacteriales with a Vibrio-Related Secondary Chromosome. *Genome Biol Evol* 14. 10.1093/gbe/evac011.
- Arias-Cartin, R., Dobihal, G.S., Campos, M., Surovtsev, I.V., Parry, B., and Jacobs-Wagner, C. (2017). Replication fork passage drives asymmetric dynamics of a critical nucleoid-associated protein in *Caulobacter*. *EMBO J* 36, 301-318. 10.15252/embj.201695513.
- Arias-Palomo, E., O'Shea, V.L., Hood, I.V., and Berger, J.M. (2013). The bacterial DnaC helicase loader is a DnaB ring breaker. *Cell* 153, 438-448. 10.1016/j.cell.2013.03.006.
- Baek, J.H., and Chatteraj, D.K. (2014). Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS Genet* 10, e1004184. 10.1371/journal.pgen.1004184.
- Baril, C., Richaud, C., Baranton, G., and Saint Girons, I.S. (1989). Linear chromosome of *Borrelia burgdorferi*. *Res Microbiol* 140, 507-516. 10.1016/0923-2508(89)90083-1.
- Beloin, C., Jeusset, J., Revet, B., Mirambeau, G., Le Hégarat, F., and Le Cam, E. (2003). Contribution of DNA conformation and topology in right-handed DNA wrapping by the *Bacillus subtilis* LrpC protein. *J Biol Chem* 278, 5333-5342. 10.1074/jbc.M207489200.
- Berger, M., and Wolde, P.R.T. (2022). Robust replication initiation from coupled homeostatic mechanisms. *Nat Commun* 13, 6556. 10.1038/s41467-022-33886-6.
- Bird, R.E., Louarn, J., Martuscelli, J., and Caro, L. (1972). Origin and sequence of chromosome replication in *Escherichia coli*. *J Mol Biol* 70, 549-566. 10.1016/0022-2836(72)90559-1.
- Bogan, J.A., and Helmstetter, C.E. (1996). *miuC* transcription, initiation of replication, and the eclipse in *Escherichia coli*. *J Bacteriol* 178, 3201-3206. 10.1128/jb.178.11.3201-3206.1996.
- Brendler, T., Abeles, A., and Austin, S. (1991). Critical sequences in the core of the P1 plasmid replication origin. *J Bacteriol* 173, 3935-3942. 10.1128/jb.173.13.3935-3942.1991.
- Brendler, T., Abeles, A., and Austin, S. (1995). A protein that binds to the P1 origin core and the *oriC* 13mer region in a methylation-specific fashion is the product of the host *seqA* gene. *EMBO J* 14, 4083-4089.
- Brezellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S., and Ferat, J.L. (2016). *DciA* is an ancestral replicative helicase operator essential for bacterial replication initiation. *Nat Commun* 7, 13271. 10.1038/ncomms13271.
- Brinkman, A.B., Ettema, T.J., de Vos, W.M., and van der Oost, J. (2003). The Lrp family of transcriptional regulators. *Mol Microbiol* 48, 287-294. 10.1046/j.1365-2958.2003.03442.x.
- Bruhn, M., Schindler, D., Kemter, F.S., Wiley, M.R., Chase, K., Koroleva, G.I., Palacios, G., Sozhamannan, S., and Waldminghaus, T. (2018). Functionality of Two Origins of Replication in *Vibrio cholerae* Strains With a Single Chromosome. *Front Microbiol* 9, 2932. 10.3389/fmicb.2018.02932.
- Bury, K., Wegrzyn, K., and Konieczny, I. (2017). Handcuffing reversal is facilitated by proteases and replication initiator monomers. *Nucleic Acids Res* 45, 3953-3966. 10.1093/nar/gkx166.
- Cairns, J. (1963). The bacterial chromosome and its manner of replication as seen by autoradiography. *J Mol Biol* 6, 208-213. 10.1016/s0022-2836(63)80070-4.
- Campbell, J.L., and Kleckner, N. (1990). *E. coli oriC* and the *dnaA* gene promoter are sequestered from *dam* methyltransferase following the passage of the chromosomal replication fork. *Cell* 62, 967-979. 10.1016/0092-8674(90)90271-f.
- Cargemel, C., Walbott, H., Durand, D., Legrand, P., Ouldali, M., Ferat, J.L., Marsin, S., and Quevillon-Cheruel, S. (2022). The apo-form of the *Vibrio cholerae* replicative helicase DnaB is

a labile and inactive planar trimer of dimers. *FEBS Lett* 596, 2031-2040. 10.1002/1873-3468.14403.

Caspi, R., Pacek, M., Consiglieri, G., Helinski, D.R., Toukdarian, A., and Konieczny, I. (2001). A broad host range replicon with different requirements for replication initiation in three bacterial species. *EMBO J* 20, 3262-3271. 10.1093/emboj/20.12.3262.

Cervantes-Rivera, R., Pedraza-Lopez, F., Perez-Segura, G., and Cevallos, M.A. (2011). The replication origin of a repABC plasmid. *BMC microbiology* 11, 158. 10.1186/1471-2180-11-158.

Chatterjee, S., Jha, J.K., Ciaccia, P., Venkova, T., and Chatteraj, D.K. (2020). Interactions of replication initiator RctB with single- and double-stranded DNA in origin opening of *Vibrio cholerae* chromosome 2. *Nucleic Acids Res.* 10.1093/nar/gkaa826.

Chatteraj, D.K. (2000). Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol* 37, 467-476.

Chen, S., Rosner, M.H., and Calvo, J.M. (2001). Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from *Escherichia coli*. *J Mol Biol* 312, 625-635. 10.1006/jmbi.2001.4955.

Ciaccia, P.N., Ramachandran, R., and Chatteraj, D.K. (2018). A Requirement for Global Transcription Factor Lrp in Licensing Replication of *Vibrio cholerae* Chromosome 2. *Front Microbiol* 9, 2103. 10.3389/fmicb.2018.02103.

Collier, J. (2012). Regulation of chromosomal replication in *Caulobacter crescentus*. *Plasmid* 67, 76-87. 10.1016/j.plasmid.2011.12.007.

Coskun-Ari, F.F., and Hill, T.M. (1997). Sequence-specific interactions in the Tus-Ter complex and the effect of base pair substitutions on arrest of DNA replication in *Escherichia coli*. *J Biol Chem* 272, 26448-26456. 10.1074/jbc.272.42.26448.

Couturier, E., and Rocha, E.P. (2006). Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Mol Microbiol* 59, 1506-1518. 10.1111/j.1365-2958.2006.05046.x.

Cui, Y., Wang, Q., Stormo, G.D., and Calvo, J.M. (1995). A consensus sequence for binding of Lrp to DNA. *J Bacteriol* 177, 4872-4880. 10.1128/jb.177.17.4872-4880.1995.

Das, N., and Chatteraj, D.K. (2004). Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. *Mol Microbiol* 54, 836-849. 10.1111/j.1365-2958.2004.04322.x.

Datta, H.J., Khatri, G.S., and Bastia, D. (1999). Mechanism of recruitment of DnaB helicase to the replication origin of the plasmid pSC101. *Proc Natl Acad Sci U S A* 96, 73-78. 10.1073/pnas.96.1.73.

de Lemos Martins, F., Fournes, F., Mazzuoli, M.V., Mazel, D., and Val, M.E. (2018). *Vibrio cholerae* chromosome 2 copy number is controlled by the methylation-independent binding of its monomeric initiator to the chromosome 1 *crtS* site. *Nucleic Acids Res* 46, 10145-10156. 10.1093/nar/gky790.

de los Rios, S., and Perona, J.J. (2007). Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J Mol Biol* 366, 1589-1602. 10.1016/j.jmb.2006.12.032.

Debaugny, R.E., Sanchez, A., Rech, J., Labourdette, D., Dorignac, J., Geniet, F., Palmeri, J., Parmeggiani, A., Boudsocq, F., Anton Leberre, V., et al. (2018). A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids. *Mol Syst Biol* 14, e8516. 10.15252/msb.20188516.

Demarre, G., and Chatteraj, D.K. (2010). DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle. *PLoS Genet* 6, e1000939. 10.1371/journal.pgen.1000939.

Diaz-Lopez, T., Davila-Fajardo, C., Blaesing, F., Lillo, M.P., and Giraldo, R. (2006). Early events in the binding of the pPS10 replication protein RepA to single iteron and operator DNA sequences. *J Mol Biol* 364, 909-920. 10.1016/j.jmb.2006.09.013.

Diaz-Lopez, T., Lages-Gonzalo, M., Serrano-Lopez, A., Alfonso, C., Rivas, G., Diaz-Orejas, R., and Giraldo, R. (2003). Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. *J Biol Chem* *278*, 18606-18616. 10.1074/jbc.M212024200.

diCenzo, G.C., and Finan, T.M. (2017). The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol Mol Biol Rev* *81*. 10.1128/MMBR.00019-17.

diCenzo, G.C., Mengoni, A., and Perrin, E. (2019). Chromids Aid Genome Expansion and Functional Diversification in the Family Burkholderiaceae. *Mol Biol Evol* *36*, 562-574. 10.1093/molbev/msy248.

Donachie, W.D. (1968). Relationship between cell size and time of initiation of DNA replication. *Nature* *219*, 1077-1079. 10.1038/2191077a0.

Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwinska, J., and Zawilak-Pawlik, A. (2012). *Helicobacter pylori* oriC--the first bipartite origin of chromosome replication in Gram-negative bacteria. *Nucleic Acids Res* *40*, 9647-9660. 10.1093/nar/gks742.

Dubarry, N., Willis, C.R., Ball, G., Lesterlin, C., and Armitage, J.P. (2019). In Vivo Imaging of the Segregation of the 2 Chromosomes and the Cell Division Proteins of *Rhodobacter sphaeroides* Reveals an Unexpected Role for MipZ. *mBio* *10*. 10.1128/mBio.02515-18.

Duderstadt, K.E., Chuang, K., and Berger, J.M. (2011). DNA stretching by bacterial initiators promotes replication origin opening. *Nature* *478*, 209-213. 10.1038/nature10455.

Duderstadt, K.E., Mott, M.L., Crisona, N.J., Chuang, K., Yang, H., and Berger, J.M. (2010). Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator. *J Biol Chem* *285*, 28229-28239. 10.1074/jbc.M110.147975.

Duigou, S., Knudsen, K.G., Skovgaard, O., Egan, E.S., Lobner-Olesen, A., and Waldor, M.K. (2006). Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J Bacteriol* *188*, 6419-6424. 10.1128/JB.00565-06.

Egan, E.S., Lobner-Olesen, A., and Waldor, M.K. (2004). Synchronous replication initiation of the two *Vibrio cholerae* chromosomes. *Curr Biol* *14*, R501-502. 10.1016/j.cub.2004.06.036.

Egan, E.S., and Waldor, M.K. (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* *114*, 521-530. 10.1016/S0092-8674(03)00611-1.

Ekundayo, B., and Bleichert, F. (2019). Origins of DNA replication. *PLoS Genet* *15*, e1008320. 10.1371/journal.pgen.1008320.

Espinosa, E., Barre, F.X., and Galli, E. (2017). Coordination between replication, segregation and cell division in multi-chromosomal bacteria: lessons from *Vibrio cholerae*. *International microbiology : the official journal of the Spanish Society for Microbiology* *20*, 121-129. 10.2436/20.1501.01.293.

Espinosa, E., Paly, E., and Barre, F.X. (2020). High-Resolution Whole-Genome Analysis of Sister-Chromatid Contacts. *Mol Cell* *79*, 857-869 e853. 10.1016/j.molcel.2020.06.033.

Ferullo, D.J., Cooper, D.L., Moore, H.R., and Lovett, S.T. (2009). Cell cycle synchronization of *Escherichia coli* using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods* *48*, 8-13. 10.1016/j.ymeth.2009.02.010.

Filutowicz, M., Davis, G., Greener, A., and Helinski, D.R. (1985). Autorepressor properties of the pi-initiation protein encoded by plasmid R6K. *Nucleic Acids Res* *13*, 103-114. 10.1093/nar/13.1.103.

Fournes, F., Niault, T., Czarnecki, J., Tissier-Visconti, A., Mazel, D., and Val, M.E. (2021). The coordinated replication of *Vibrio cholerae*'s two chromosomes required the acquisition of a unique domain by the RctB initiator. *Nucleic Acids Res* *49*, 11119-11133. 10.1093/nar/gkab903.

Fournes, F., Val, M.E., Skovgaard, O., and Mazel, D. (2018). Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes. *Front Microbiol* *9*, 1833. 10.3389/fmicb.2018.01833.

Frimodt-Moller, J., Charbon, G., Krogfelt, K.A., and Lobner-Olesen, A. (2016). DNA Replication Control Is Linked to Genomic Positioning of Control Regions in *Escherichia coli*. *PLoS Genet* *12*, e1006286. 10.1371/journal.pgen.1006286.

Frimodt-Moller, J., Charbon, G., and Lobner-Olesen, A. (2017). Control of bacterial chromosome replication by non-coding regions outside the origin. *Curr Genet* 63, 607-611. 10.1007/s00294-016-0671-6.

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., and Yokoyama, S. (2003). Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res* 31, 2077-2086. 10.1093/nar/gkg309.

Fujimitsu, K., and Katayama, T. (2004). Reactivation of DnaA by DNA sequence-specific nucleotide exchange in vitro. *Biochemical and biophysical research communications* 322, 411-419. 10.1016/j.bbrc.2004.07.141.

Fujimitsu, K., Senriuchi, T., and Katayama, T. (2009). Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev* 23, 1221-1233. 10.1101/gad.1775809.

Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P., et al. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293, 668-672. 10.1126/science.1060966.

Galli, E., Ferat, J.L., Desfontaines, J.M., Val, M.E., Skovgaard, O., Barre, F.X., and Possoz, C. (2019). Replication termination without a replication fork trap. *Sci Rep* 9, 8315. 10.1038/s41598-019-43795-2.

Gasset-Rosa, F., Diaz-Lopez, T., Lurz, R., Prieto, A., Fernandez-Tresguerres, M.E., and Giraldo, R. (2008). Negative regulation of pPS10 plasmid replication: origin pairing by zipping-up DNA-bound RepA monomers. *Mol Microbiol* 68, 560-572. 10.1111/j.1365-2958.2008.06166.x.

Gerding, M.A., Chao, M.C., Davis, B.M., and Waldor, M.K. (2015). Molecular Dissection of the Essential Features of the Origin of Replication of the Second *Vibrio cholerae* Chromosome. *mBio* 6, e00973. 10.1128/mBio.00973-15.

Giraldo, R., Andreu, J.M., and Diaz-Orejas, R. (1998). Protein domains and conformational changes in the activation of RepA, a DNA replication initiator. *EMBO J* 17, 4511-4526. 10.1093/emboj/17.15.4511.

Gross, M.H., and Konieczny, I. (2020). Polyphosphate induces the proteolysis of ADP-bound fraction of initiator to inhibit DNA replication initiation upon stress in *Escherichia coli*. *Nucleic Acids Res.* 10.1093/nar/gkaa217.

Guo, X., Flores, M., Mavingui, P., Fuentes, S.I., Hernandez, G., Davila, G., and Palacios, R. (2003). Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res* 13, 1810-1817. 10.1101/gr.1260903.

Hall, J.P.J., Botelho, J., Cazares, A., and Baltrus, D.A. (2022). What makes a megaplasmid? *Philos Trans R Soc Lond B Biol Sci* 377, 20200472. 10.1098/rstb.2020.0472.

Han, K., Li, Z.F., Peng, R., Zhu, L.P., Zhou, T., Wang, L.G., Li, S.G., Zhang, X.B., Hu, W., Wu, Z.H., et al. (2013). Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci Rep* 3, 2101. 10.1038/srep02101.

Harrison, P.W., Lower, R.P., Kim, N.K., and Young, J.P. (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol* 18, 141-148. 10.1016/j.tim.2009.12.010.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477-483. 10.1038/35020000.

Helmstetter, C.E. (1968). Origin and sequence of chromosome replication in *Escherichia coli* B-r. *J Bacteriol* 95, 1634-1641. 10.1128/jb.95.5.1634-1641.1968.

Helmstetter, C.E., and Cummings, D.J. (1963). Bacterial Synchronization by Selection of Cells at Division. *Proc Natl Acad Sci U S A* 50, 767-774. 10.1073/pnas.50.4.767.

Hiraga, S. (1976). Novel F prime factors able to replicate in *Escherichia coli* Hfr strains. *Proc Natl Acad Sci U S A* 73, 198-202. 10.1073/pnas.73.1.198.

Hirota, Y., Ryter, A., and Jacob, F. (1968). Thermosensitive mutants of *E. coli* affected in the processes of DNA synthesis and cellular division. *Cold Spring Harb Symp Quant Biol* 33, 677-693. 10.1101/sqb.1968.033.01.077.

Ingmer, H., Fong, E.L., and Cohen, S.N. (1995). Monomer-dimer equilibrium of the pSC101 RepA protein. *J Mol Biol* 250, 309-314. 10.1006/jmbi.1995.0378.

Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K., and Katayama, T. (2004). DiaA, a novel DnaA-binding protein, ensures the timely initiation of *Escherichia coli* chromosome replication. *J Biol Chem* 279, 45546-45555. 10.1074/jbc.M402762200.

Jaskolska, M., Adams, D.W., and Blokesch, M. (2022). Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. *Nature*. 10.1038/s41586-022-04546-y.

Jers, C., Ravikumar, V., Lezyk, M., Sultan, A., Sjolting, A., Wai, S.N., and Mijakovic, I. (2017). The Global Acetylome of the Human Pathogen *Vibrio cholerae* V52 Reveals Lysine Acetylation of Major Transcriptional Regulators. *Front Cell Infect Microbiol* 7, 537. 10.3389/fcimb.2017.00537.

Jha, J.K., and Chatteraj, D.K. (2016). Inactivation of Individual SeqA Binding Sites of the *E. coli* Origin Reveals Robustness of Replication Initiation Synchrony. *PloS one* 11, e0166722. 10.1371/journal.pone.0166722.

Jha, J.K., Demarre, G., Venkova-Canova, T., and Chatteraj, D.K. (2012). Replication regulation of *Vibrio cholerae* chromosome II involves initiator binding to the origin both as monomer and as dimer. *Nucleic Acids Res* 40, 6026-6038. 10.1093/nar/gks260.

Jha, J.K., Ghirlando, R., and Chatteraj, D.K. (2014). Initiator protein dimerization plays a key role in replication control of *Vibrio cholerae* chromosome 2. *Nucleic Acids Res* 42, 10538-10549. 10.1093/nar/gku771.

Jha, J.K., Li, M., Ghirlando, R., Miller Jenkins, L.M., Wlodawer, A., and Chatteraj, D. (2017). The DnaK Chaperone Uses Different Mechanisms To Promote and Inhibit Replication of *Vibrio cholerae* Chromosome 2. *mBio* 8. 10.1128/mBio.00427-17.

Jiang, Y., Pacek, M., Helinski, D.R., Konieczny, I., and Toukdarian, A. (2003). A multifunctional plasmid-encoded replication initiation protein both recruits and positions an active helicase at the replication origin. *Proc Natl Acad Sci U S A* 100, 8692-8697. 10.1073/pnas.1532393100.

Kaplan, D.L. (2006). Replication termination: mechanism of polar arrest revealed. *Curr Biol* 16, R684-686. 10.1016/j.cub.2006.08.013.

Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T., and Katayama, T. (2014). Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production and replication initiation. *Nucleic Acids Res* 42, 13134-13149. 10.1093/nar/gku1051.

Kasho, K., and Katayama, T. (2013). DnaA binding locus *datA* promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc Natl Acad Sci U S A* 110, 936-941. 10.1073/pnas.1212070110.

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA Cycle in *Escherichia coli*: Activation, Function and Inactivation of the Initiator Protein. *Front Microbiol* 8, 2496. 10.3389/fmicb.2017.02496.

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. *Nat Rev Microbiol* 8, 163-170. 10.1038/nrmicro2314.

Kemter, F.S., Messerschmidt, S.J., Schallopp, N., Sobetzko, P., Lang, E., Bunk, B., Sproer, C., Teschler, J.K., Yildiz, F.H., Overmann, J., and Waldminghaus, T. (2018). Synchronous termination of replication of the two chromosomes is an evolutionary selected feature in *Vibrionaceae*. *PLoS Genet* 14, e1007251. 10.1371/journal.pgen.1007251.

Kemter, F.S., Schallopp, N., Sperlea, T., Serrania, J., Sobetzko, P., Fritz, G., and Waldminghaus, T. (2019). Stringent response leads to continued cell division and a temporal restart of DNA replication after initial shutdown in *Vibrio cholerae*. *Mol Microbiol* 111, 1617-1637. 10.1111/mmi.14241.

Keyamura, K., Fujikawa, N., Ishida, T., Ozaki, S., Su'etsugu, M., Fujimitsu, K., Kagawa, W., Yokoyama, S., Kurumizaka, H., and Katayama, T. (2007). The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP DnaA-specific initiation complexes. *Genes Dev* 21, 2083-2099. 10.1101/gad.1561207.

Kim, J.W., Bugata, V., Cortes-Cortes, G., Quevedo-Martinez, G., and Camps, M. (2020). Mechanisms of Theta Plasmid Replication in Enterobacteria and Implications for Adaptation to Its Host. *EcoSal Plus* 9. 10.1128/ecosalplus.ESP-0026-2019.

Koch, B., Ma, X., and Lobner-Olesen, A. (2010). Replication of *Vibrio cholerae* chromosome I in *Escherichia coli*: dependence on dam methylation. *J Bacteriol* 192, 3903-3914. 10.1128/JB.00311-10.

Komori, H., Matsunaga, F., Higuchi, Y., Ishiai, M., Wada, C., and Miki, K. (1999). Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 Å resolution. *EMBO J* 18, 4597-4607. 10.1093/emboj/18.17.4597.

Konieczny, I., Bury, K., Wawrzycka, A., and Wegrzyn, K. (2014). Iteron Plasmids. *Microbiol Spectr* 2. 10.1128/microbiolspec.PLAS-0026-2014.

Konieczny, I., and Liberek, K. (2002). Cooperative action of *Escherichia coli* ClpB protein and DnaK chaperone in the activation of a replication initiation protein. *J Biol Chem* 277, 18483-18488. 10.1074/jbc.M107580200.

Kothapalli, R., Ghirlando, R., Khan, Z.A., Chatterjee, S., Keddi, N., and Chatteraj, D.K. (2022). The dimerization interface of initiator RctB governs chaperone and enhancer dependence of *Vibrio cholerae* chromosome 2 replication. *Nucleic Acids Res* 50, 4529-4544. 10.1093/nar/gkac210.

Kroner, G.M., Wolfe, M.B., and Freddolino, P.L. (2019). *Escherichia coli* Lrp Regulates One-Third of the Genome via Direct, Cooperative, and Indirect Routes. *J Bacteriol* 201. 10.1128/JB.00411-18.

Kunnimalaiyaan, S., Inman, R.B., Rakowski, S.A., and Filutowicz, M. (2005). Role of pi dimers in coupling ("handcuffing") of plasmid R6K's gamma ori iterons. *J Bacteriol* 187, 3779-3785. 10.1128/JB.187.11.3779-3785.2005.

Lesterlin, C., Barre, F.X., and Cornet, F. (2004). Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol Microbiol* 54, 1151-1160. 10.1111/j.1365-2958.2004.04356.x.

Lewis, J.S., Jergic, S., and Dixon, N.E. (2016). The *E. coli* DNA Replication Fork. *Enzymes* 39, 31-88. 10.1016/bs.enz.2016.04.001.

Lobner-Olesen, A., Skovgaard, O., and Marinus, M.G. (2005). Dam methylation: coordinating cellular processes. *Curr Opin Microbiol* 8, 154-160. 10.1016/j.mib.2005.02.009.

Loot, C., Bikard, D., Rachlin, A., and Mazel, D. (2010). Cellular pathways controlling integron cassette site folding. *EMBO J* 29, 2623-2634. 10.1038/emboj.2010.151.

Maaloe, O., and Hanawalt, P.C. (1961). Thymine deficiency and the normal DNA replication cycle. I. *J Mol Biol* 3, 144-155. 10.1016/s0022-2836(61)80041-7.

Macek, B., Forchhammer, K., Hardouin, J., Weber-Ban, E., Grangeasse, C., and Mijakovic, I. (2019). Protein post-translational modifications in bacteria. *Nat Rev Microbiol* 17, 651-664. 10.1038/s41579-019-0243-0.

Mackiewicz, P., Zakrzewska-Czerwinska, J., Zawilak, A., Dudek, M.R., and Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Res* 32, 3781-3791. 10.1093/nar/gkh699.

Marinus, M.G., and Lobner-Olesen, A. (2014). DNA Methylation. *EcoSal Plus* 6. 10.1128/ecosalplus.ESP-0003-2013.

Marsin, S., Adam, Y., Cargemel, C., Andreani, J., Baconnais, S., Legrand, P., Li de la Sierra-Gallay, I., Humbert, A., Aumont-Nicaise, M., Velours, C., et al. (2021). Study of the DnaB:DciA interplay reveals insights into the primary mode of loading of the bacterial replicative helicase. *Nucleic Acids Res* 49, 6569-6586. 10.1093/nar/gkab463.

Martinez-Antonio, A., and Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* *6*, 482-489. 10.1016/j.mib.2003.09.002.

Mazel, D., Dychinco, B., Webb, V.A., and Davies, J. (1998). A distinctive class of integron in the *Vibrio cholerae* genome. *Science* *280*, 605-608.

McGarry, K.C., Ryan, V.T., Grimwade, J.E., and Leonard, A.C. (2004). Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc Natl Acad Sci U S A* *101*, 2811-2816. 10.1073/pnas.0400340101.

Mitkova, A.V., Khopde, S.M., and Biswas, S.B. (2003). Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of *Escherichia coli* in RNA primer synthesis. *J Biol Chem* *278*, 52253-52261. 10.1074/jbc.M308956200.

Miyoshi, K., Tatsumoto, Y., Ozaki, S., and Katayama, T. (2021). Negative feedback for DARS2-Fis complex by ATP-DnaA supports the cell cycle-coordinated regulation for chromosome replication. *Nucleic Acids Res* *49*, 12820-12835. 10.1093/nar/gkab1171.

Molina-Garcia, L., Gasset-Rosa, F., Moreno-Del Alamo, M., Fernandez-Tresguerres, M.E., Moreno-Diaz de la Espina, S., Lurz, R., and Giraldo, R. (2016). Functional amyloids as inhibitors of plasmid DNA replication. *Sci Rep* *6*, 25425. 10.1038/srep25425.

Mori, J.F., and Kanaly, R.A. (2022). Natural Chromosome-Chromid Fusion across rRNA Operons in a Burkholderiaceae Bacterium. *Microbiol Spectr*, e0222521. 10.1128/spectrum.02225-21.

Nakamura, A., Wada, C., and Miki, K. (2007). Structural basis for regulation of bifunctional roles in replication initiator protein. *Proc Natl Acad Sci U S A* *104*, 18484-18489. 10.1073/pnas.0705623104.

Nielsen, O., and Lobner-Olesen, A. (2008). Once in a lifetime: strategies for preventing re-replication in prokaryotic and eukaryotic cells. *EMBO Rep* *9*, 151-156. 10.1038/sj.embor.2008.2.

Nordstrom, K. (2006). Plasmid R1--replication and its control. *Plasmid* *55*, 1-26. 10.1016/j.plasmid.2005.07.002.

Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005). *Vibrios* commonly possess two chromosomes. *J Bacteriol* *187*, 752-757. 10.1128/JB.187.2.752-757.2005.

Olliver, A., Saggiaro, C., Herrick, J., and Sclavi, B. (2010). DnaA-ATP acts as a molecular switch to control levels of ribonucleotide reductase expression in *Escherichia coli*. *Mol Microbiol* *76*, 1555-1571. 10.1111/j.1365-2958.2010.07185.x.

Orlova, N., Gerding, M., Ivashkiv, O., Olinares, P.D.B., Chait, B.T., Waldor, M.K., and Jeruzalmi, D. (2017). The replication initiator of the cholera pathogen's second chromosome shows structural similarity to plasmid initiators. *Nucleic Acids Res* *45*, 3724-3737. 10.1093/nar/gkw1288.

Ozaki, S., and Katayama, T. (2012). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res* *40*, 1648-1665. 10.1093/nar/gkr832.

Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S.Y., Yokoyama, S., Kurumizaka, H., and Katayama, T. (2008). A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J Biol Chem* *283*, 8351-8362. 10.1074/jbc.M708684200.

Petersen, J., Brinkmann, H., and Pradella, S. (2009). Diversity and evolution of repABC type plasmids in Rhodobacterales. *Environ Microbiol* *11*, 2627-2638. 10.1111/j.1462-2920.2009.01987.x.

Pinto, U.M., Pappas, K.M., and Winans, S.C. (2012). The ABCs of plasmid replication and segregation. *Nat Rev Microbiol* *10*, 755-765. 10.1038/nrmicro2882.

Planchenault, C., Pons, M.C., Schiavon, C., Siguier, P., Rech, J., Guynet, C., Dauverd-Girault, J., Cury, J., Rocha, E.P.C., Junier, I., et al. (2020). Intracellular Positioning Systems Limit the Entropic Eviction of Secondary Replicons Toward the Nucleoid Edges in Bacterial Cells. *J Mol Biol* *432*, 745-761. 10.1016/j.jmb.2019.11.027.

Platko, J.V., and Calvo, J.M. (1993). Mutations affecting the ability of *Escherichia coli* Lrp to bind DNA, activate transcription, or respond to leucine. *J Bacteriol* *175*, 1110-1117. 10.1128/jb.175.4.1110-1117.1993.

Postow, L., Crisona, N.J., Peter, B.J., Hardy, C.D., and Cozzarelli, N.R. (2001). Topological challenges to DNA replication: conformations at the fork. *Proc Natl Acad Sci U S A* *98*, 8219-8226. 10.1073/pnas.111006998.

Ramachandran, R., Jha, J., Paulsson, J., and Chattoraj, D. (2017). Random versus Cell Cycle-Regulated Replication Initiation in Bacteria: Insights from Studying *Vibrio cholerae* Chromosome 2. *Microbiol Mol Biol Rev* *81*. 10.1128/MMBR.00033-16.

Rasmussen, T., Jensen, R.B., and Skovgaard, O. (2007). The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J* *26*, 3124-3131. 10.1038/sj.emboj.7601747.

Ree, R., Varland, S., and Arnesen, T. (2018). Spotlight on protein N-terminal acetylation. *Exp Mol Med* *50*, 1-13. 10.1038/s12276-018-0116-z.

Ren, Z., Liao, Q., Karaboja, X., Barton, I.S., Schantz, E.G., Mejia-Santana, A., Fuqua, C., and Wang, X. (2022). Conformation and dynamic interactions of the multipartite genome in *Agrobacterium tumefaciens*. *Proc Natl Acad Sci U S A* *119*. 10.1073/pnas.2115854119.

Reyes-Lamothe, R., and Sherratt, D.J. (2019). The bacterial cell cycle, chromosome inheritance and cell growth. *Nat Rev Microbiol* *17*, 467-478. 10.1038/s41579-019-0212-7.

Riber, L., Frimodt-Moller, J., Charbon, G., and Lobner-Olesen, A. (2016). Multiple DNA Binding Proteins Contribute to Timing of Chromosome Replication in *E. coli*. *Front Mol Biosci* *3*, 29. 10.3389/fmolb.2016.00029.

Richardson, T.T., Harran, O., and Murray, H. (2016). The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. *Nature* *534*, 412-416. 10.1038/nature17962.

Richter, F., Fonfara, I., Bouazza, B., Schumacher, C.H., Bratovic, M., Charpentier, E., and Moglich, A. (2016). Engineering of temperature- and light-switchable Cas9 variants. *Nucleic Acids Res* *44*, 10003-10014. 10.1093/nar/gkw930.

Rosenberg, C., Boistard, P., Denarie, J., and Casse-Delbart, F. (1981). Genes controlling early and late functions in symbiosis are located on a megaplasmid in *Rhizobium meliloti*. *Mol Gen Genet* *184*, 326-333. 10.1007/BF00272926.

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H., and Katayama, T. (2017). Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in *Escherichia coli*. *Nucleic Acids Res* *45*, 12354-12373. 10.1093/nar/gkx914.

Schaechter, M., Bentzon, M.W., and Maaloe, O. (1959). Synthesis of deoxyribonucleic acid during the division cycle of bacteria. *Nature* *183*, 1207-1208. 10.1038/1831207a0.

Schaechter, M., Maaloe, O., and Kjeldgaard, N.O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J Gen Microbiol* *19*, 592-606. 10.1099/00221287-19-3-592.

Seitz, H., Weigel, C., and Messer, W. (2000). The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. *Mol Microbiol* *37*, 1270-1279. 10.1046/j.1365-2958.2000.02096.x.

Sekimizu, K., Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* *50*, 259-265. 10.1016/0092-8674(87)90221-2.

Sharma, S., Sathyanarayana, B.K., Bird, J.G., Hoskins, J.R., Lee, B., and Wickner, S. (2004). Plasmid P1 RepA is homologous to the F plasmid RepE class of initiators. *J Biol Chem* *279*, 6027-6034. 10.1074/jbc.M310917200.

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harbor perspectives in biology* *5*, a012922. 10.1101/cshperspect.a012922.

Skarstad, K., and Lobner-Olesen, A. (2003). Stable co-existence of separate replicons in *Escherichia coli* is dependent on once-per-cell-cycle initiation. *EMBO J* *22*, 140-150. 10.1093/emboj/cdg003.

Smillie, C., Garcillan-Barcia, M.P., Francia, M.V., Rocha, E.P., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol Mol Biol Rev* 74, 434-452. 10.1128/MMBR.00020-10.

Smith, J.L., and Grossman, A.D. (2015). In Vitro Whole Genome DNA Binding Analysis of the Bacterial Replication Initiator and Transcription Factor DnaA. *PLoS Genet* 11, e1005258. 10.1371/journal.pgen.1005258.

Smits, W.K., Goranov, A.I., and Grossman, A.D. (2010). Ordered association of helicase loader proteins with the *Bacillus subtilis* origin of replication in vivo. *Mol Microbiol* 75, 452-461. 10.1111/j.1365-2958.2009.06999.x.

Stokke, C., Waldminghaus, T., and Skarstad, K. (2011). Replication patterns and organization of replication forks in *Vibrio cholerae*. *Microbiology (Reading)* 157, 695-708. 10.1099/mic.0.045112-0.

Suwanto, A., and Kaplan, S. (1989). Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. *J Bacteriol* 171, 5850-5859. 10.1128/jb.171.11.5850-5859.1989.

Tabata, S., Oka, A., Sugimoto, K., Takanami, M., Yasuda, S., and Hirota, Y. (1983). The 245 base-pair *oriC* sequence of the *E. coli* chromosome directs bidirectional replication at an adjacent region. *Nucleic Acids Res* 11, 2617-2626. 10.1093/nar/11.9.2617.

Tani, T.H., Khodursky, A., Blumenthal, R.M., Brown, P.O., and Matthews, R.G. (2002). Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci U S A* 99, 13471-13476. 10.1073/pnas.212510999.

Toukdarian, A.E., Helinski, D.R., and Perri, S. (1996). The plasmid RK2 initiation protein binds to the origin of replication as a monomer. *J Biol Chem* 271, 7072-7078. 10.1074/jbc.271.12.7072.

Trucksis, M., Michalski, J., Deng, Y.K., and Kaper, J.B. (1998). The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc Natl Acad Sci U S A* 95, 14464-14469.

Uga, H., Matsunaga, F., and Wada, C. (1999). Regulation of DNA replication by iterons: an interaction between the *ori2* and *incC* regions mediated by RepE-bound iterons inhibits DNA replication of mini-F plasmid in *Escherichia coli*. *EMBO J* 18, 3856-3867. 10.1093/emboj/18.13.3856.

Val, M.E., Kennedy, S.P., Soler-Bistue, A.J., Barbe, V., Bouchier, C., Ducos-Galand, M., Skovgaard, O., and Mazel, D. (2014). Fuse or die: how to survive the loss of Dam in *Vibrio cholerae*. *Mol Microbiol* 91, 665-678. 10.1111/mmi.12483.

Val, M.E., Marbouty, M., de Lemos Martins, F., Kennedy, S.P., Kemble, H., Bland, M.J., Possoz, C., Koszul, R., Skovgaard, O., and Mazel, D. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci Adv* 2, e1501914. 10.1126/sciadv.1501914.

Val, M.E., Skovgaard, O., Ducos-Galand, M., Bland, M.J., and Mazel, D. (2012). Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues. *PLoS Genet* 8, e1002472. 10.1371/journal.pgen.1002472.

Venkova-Canova, T., and Chatteraj, D.K. (2011). Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proc Natl Acad Sci U S A* 108, 6199-6204. 10.1073/pnas.1013244108.

Venkova-Canova, T., Saha, A., and Chatteraj, D.K. (2012). A 29-mer site regulates transcription of the initiator gene as well as function of the replication origin of *Vibrio cholerae* chromosome II. *Plasmid* 67, 102-110. 10.1016/j.plasmid.2011.12.009.

Venkova-Canova, T., Srivastava, P., and Chatteraj, D.K. (2006). Transcriptional inactivation of a regulatory site for replication of *Vibrio cholerae* chromosome II. *Proc Natl Acad Sci U S A* 103, 12051-12056. 10.1073/pnas.0605120103.

Vigouroux, A., Oldewurtel, E., Cui, L., Bikard, D., and van Teeffelen, S. (2018). Tuning dCas9's ability to block transcription enables robust, noiseless knockdown of bacterial genes. *Mol Syst Biol* 14, e7899. 10.15252/msb.20177899.

Wegrzyn, K., Zabrocka, E., Bury, K., Tomiczek, B., Wieczor, M., Czub, J., Uciechowska, U., Moreno-Del Alamo, M., Walkow, U., Grochowina, I., et al. (2021). Defining a novel domain that

provides an essential contribution to site-specific interaction of Rep protein with DNA. *Nucleic Acids Res* 49, 3394-3408. 10.1093/nar/gkab113.

Wegrzyn, K.E., Gross, M., Uciechowska, U., and Konieczny, I. (2016). Replisome Assembly at Bacterial Chromosomes and Iteron Plasmids. *Front Mol Biosci* 3, 39. 10.3389/fmolb.2016.00039.

Weigel, C., Schmidt, A., Ruckert, B., Lurz, R., and Messer, W. (1997). DnaA protein binding to individual DnaA boxes in the Escherichia coli replication origin, oriC. *EMBO J* 16, 6574-6583. 10.1093/emboj/16.21.6574.

Wiktor, J., Lesterlin, C., Sherratt, D.J., and Dekker, C. (2016). CRISPR-mediated control of the bacterial initiation of replication. *Nucleic Acids Res* 44, 3801-3810. 10.1093/nar/gkw214.

Xie, G., Johnson, S.L., Davenport, K.W., Rajavel, M., Waldminghaus, T., Detter, J.C., Chain, P.S., and Sozhamannan, S. (2017). Exception to the Rule: Genomic Characterization of Naturally Occurring Unusual Vibrio cholerae Strains with a Single Chromosome. *International journal of genomics* 2017, 8724304. 10.1155/2017/8724304.

Xu, Z.Q., and Dixon, N.E. (2018). Bacterial replisomes. *Curr Opin Struct Biol* 53, 159-168. 10.1016/j.sbi.2018.09.006.

Yamamoto, S., Lee, K.I., Morita, M., Arakawa, E., Izumiya, H., and Ohnishi, M. (2018). Single Circular Chromosome Identified from the Genome Sequence of the Vibrio cholerae O1 bv. El Tor Ogawa Strain V060002. *Genome Announc* 6. 10.1128/genomeA.00564-18.

Zamora, M., Ziegler, C.A., Freddolino, P.L., and Wolfe, A.J. (2020). A Thermosensitive, Phase-Variable Epigenetic Switch: pap Revisited. *Microbiol Mol Biol Rev* 84. 10.1128/MMBR.00030-17.

Zhang, Q., Zhou, A., Li, S., Ni, J., Tao, J., Lu, J., Wan, B., Li, S., Zhang, J., Zhao, S., et al. (2016). Reversible lysine acetylation is involved in DNA replication initiation by regulating activities of initiator DnaA in Escherichia coli. *Sci Rep* 6, 30837. 10.1038/srep30837.

Ziegler, C.A., and Freddolino, P.L. (2021). The leucine-responsive regulatory proteins/feast-famine regulatory proteins: an ancient and complex class of transcriptional regulators in bacteria and archaea. *Crit Rev Biochem Mol Biol* 56, 373-400. 10.1080/10409238.2021.1925215.

Zorman, S., Seitz, H., Sclavi, B., and Strick, T.R. (2012). Topological characterization of the DnaA-oriC complex using single-molecule nanomanipulation. *Nucleic Acids Res* 40, 7375-7383. 10.1093/nar/gks371.

Zzaman, S., and Bastia, D. (2005). Oligomeric initiator protein-mediated DNA looping negatively regulates plasmid replication in vitro by preventing origin melting. *Mol Cell* 20, 833-843. 10.1016/j.molcel.2005.10.037.