

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE
LA RECHERCHE SCIENTIFIQUE
UNIVERSITE SAAD DAHLAB - BLIDA 1



Faculté des Sciences de la Nature et de la Vie

Département de Biologie

Mémoire de fin d'études

En vue de l'obtention du diplôme de Master dans le domaine SNV
Filière Sciences Biologiques

Option : Génétique

Thème

**Approche In-Silico pour une conception d'expérimentation
CRISPR-Cas9 ciblant le gène rps1 chez E.coli**

Présenté par :

Date de soutenance :

***ELBEY Rayane**

***SEHBOUB Yasmine**

Devant le jury :

Nom	Grade / Lieu	Qualité
Mme BENHOUNA	Maitre de conférence B Blida1	Présidente
Mme ARKAM	Maitre de conférence B Blida1	Examinatrice
Mr MOHAMED SAID	Maitre de conférence A Blida1	Promoteur
Mr SENINET	Assistant	Co-promoteur

Promotion : 2021-2022

Remerciements

En premier lieu, nous remercions le tout puissant de nous avoir prêté la force et le courage de mener à bien ce travail.

Nous remercions tout particulièrement, nos professeurs qui sans leurs aides nous n'avons pas pu voir ce jour-là. Nous avons nommé : Notre promoteur Mr MOHAMED SAID pour son aide précieuse tout au long de notre mémoire également tout au long de notre spécialité. Notre Co-Promoteur Mr SENINET pour son aide, sa totale implication.

Nous remercions les membres du jury qui ont accepté d'examiner et d'évaluer notre travail : la présidente du jury Mme BENHOUNA, et l'examinatrice Mme ARKAM.

Nous tenons à remercier également tous nos professeurs de notre parcours universitaire

Sur le plan personnel, nous remercions nos parents pour leurs soutiens, leurs courages, leurs positivités pendant les durs moments. Merci également à toute personne qui nous a aidé de près ou de loin.

Dédicace

Nous dédions ce travail à nos parents, source de motivation. Qui nous a doté d'une éducation digne, son amour a fait de nous ce que nous sommes Aujourd'hui.

À le père de chacune de nous deux qui a été toujours là pour nous et nous encouragé d'aller vers l'avant ainsi à la maman de chacune de nous deux pour son soutien inconditionnel. Et à tous ceux qui interviennent de loin ou de près de notre réussite.

Nous dédions ce travail ainsi à nos sœurs Asma, Rania, Khaoula. Également à nos frères Mohamed Hacen et Abdelmoumen que nous ne trouverons jamais assez de mots pour leurs exprimer notre amour, en leurs souhaitant pleins de succès.

À nos cousins, cousines, toutes nos familles, amis, et toute la section M2 génétique.

Et à tous ceux qui interviennent de loin ou de près de notre réussite.

Résumé

Le système CRISPR-Cas9, est un outil qui a été développé pour supprimer ou insérer de l'ADN dans des gènes cibles dans des cellules bactériennes ou des cellules de mammifères afin d'obtenir des cellules présentant les traits souhaités.

Dans cette étude in-silico, nous avons utilisé le système CRISPR-Cas9 pour Gene Editing d'une souche bactérienne non pathogène. *Escherichia coli*. Nous avons introduit une mutation dans le gène de la bactérie qui a converti son phénotype de sensibilité à l'antibiotique streptomycine en résistance à la streptomycine ; permettant ainsi aux bactéries de survivre sur un milieu contenant de la streptomycine. Il s'agit d'une mutation du génome (K43T) vers le gène rpsL par un seul changement de base nucléotidique une adénine (A) à une cytosine (C). Ce changement amène l'ADN à coder pour une Thréonine (T) au lieu de Lysine (K). Cette étude nous a permis de connaître la découverte et la caractérisation du système immunitaire procaryote CRISPR-Cas ; Ainsi ces applications comme outil d'édition du génome révolutionnaire dans les cellules vivantes de diverses espèces pour modifier, visualiser et identifier des séquences d'ADN et d'ARN. La modification du génome médiée par la protéine associée à CRISPR (Cas) 9 nous permet d'éditer rapidement et efficacement les génomes d'une variété d'organismes. Les avantages du système CRISPR/Cas9 c'est d'en faire un outil de génie génétique de plus en plus populaire pour des applications biologiques et thérapeutiques.

Mots clés : CRISPR-Cas9, K43T, ADN, Cas9, rpsL, E.coli.

Abstract

The CRISPR-Cas9 system, is a tool that was developed to delete or insert DNA into target genes in bacterial cells or mammalian cells to obtain cells with desired traits.

In this in-silico study, we used the CRISPR-Cas9 system for Gene Editing of a non-pathogenic bacterial strain, *Escherichia coli*. We introduced a mutation in the gene of the bacterium which converted its phenotype of sensitivity to the antibiotic streptomycin into resistance to streptomycin; this allowing the bacteria to survive on a medium containing streptomycin. It is a mutation of the genome (K43T) towards the rpsL gene by a single change of nucleotide base from adenine (A) to cytosine (C). This change causes DNA to encode a Threonine (T) instead of Lysine (K).

This study will allow us to know the discovery and characterization of the prokaryotic immune system CRISPR-Cas, this its applications as a revolutionary genome editing tool in living cells of various species to modify, visualize and identify DNA sequences and RNA. CRISPR-associated protein (Cas) 9-mediated genome editing allow us to quickly and efficiently edit the genomes of a variety of organisms. The advantages of the CRISPR/Cas9 system have made it a tool increasingly popular genetic engineering for biological applications and therapies.

Keys words : CRISPR-Cas9,K43T,DNA,Cas9,rpsL,E.coli.

ملخص

نظام CRISPR-Cas9 ، هو أداة تم تطويرها لحذف أو إدخال الحمض النووي في الجينات المستهدفة في الخلايا البكتيرية أو خلايا الثدييات للحصول على الخلايا ذات السمات المرغوبة في هذه الدراسة داخل السليكو ، استخدمنا نظام CRISPR-Cas9 لتحرير الجينات لسلالة بكتيرية غير مُعرضة ، *Escherichia coli*. أدخلنا طفرة في جين البكتيريا التي حولت النمط الظاهري لحساسية المضاد الحيوي الستربتومايسين إلى مقاومة الستربتومايسين. وبالتالي السماح للبكتيريا بالعيش على وسط يحتوي على الستربتومايسين. إنها طفرة في الجينوم (K43T) تجاه جين *rpsL* عن طريق تغيير واحد لقاعدة النوكليوتيدات من الأدينين (A) إلى السيتوزين (C). يتسبب هذا التغيير في قيام الحمض النووي بتشفير *Threonine (T)* بدلاً من *Lysine (K)*. ستسمح لنا هذه الدراسة بمعرفة اكتشاف وتوصيف نظام المناعة بدائية النواة CRISPR-Cas ، وبالتالي تطبيقاته كأداة ثورية لتحرير الجينوم في الخلايا الحية من مختلف الأنواع لتعديل وتصوير وتحديد تسلسل الحمض النووي و *dRNA*. يسمح لنا تحرير الجينوم 9 بواسطة البروتين المرتبط بـ CRISPR (*Cas*) بتعديل جينومات مجموعة متنوعة من الكائنات الحية بسرعة وكفاءة. جعلت مزايا نظام CRISPR / Cas9 منه أداة تزداد شعبية الهندسة الوراثية للتطبيقات البيولوجية و العلاجات.

الكلمات الرئيسية : CRISPR-Cas9, K43T, DNA, Cas9, rpsL, E.coli

Abréviation

ADN : Acide DesoxyriboNucléique

ARN: Acide RiboNucléique

ARNg: Acide RiboNucléique guide

ARNm: ARN messenger

CRISPR: Clustered Regulary Interspaced Short Palindrimic Repeats

crRNA : CRISPR RiboNucleic Acid

tracrRNA : Trans-activating crRNA

ARNc : ARN associé à CRISPR

pré-ARNc : précurseur-ARNcr

Cas9 : CRISPR associated protein 9

PAM : Protospacer Adjacent Motif

FRET: Fluorescence Resonance Energy Transfer

Hel: Hélicoïdal

ORD: ORigin Duplication

NHEJ: Non Homologous End Joining

HDR: Homologous DNA Repair

DSB: Double Strand Break

CTD: C-Terminal Domain

Liste des figures

<u>Figure</u>	<u>Titre</u>	<u>Page</u>
Figure 1	Aperçu du mécanisme d'action CRISPR/Cas	<u>6</u>
<u>Figure 2</u>	Biologie du système CRISPR-Cas de type II-A	<u>10</u>
<u>Figure 3</u>	La protéine Cas9 est composée de six domaines	<u>14</u>
<u>Figure 4</u>	L'ARN guide	<u>14</u>
<u>Figure 5</u>	Mécanisme d'activation de la protéine Cas9	<u>15</u>
<u>Figure 6</u>	les cassures par CRISPR-Cas et la Recombinaison	<u>15</u>
<u>Figure 7</u>	Représentations schématiques des mécanismes proposés de reconnaissance et de clivage de l'ADN cible médiés par CRISPR-Cas9	<u>17</u>
<u>Figure 8</u>	Régulation de l'activité endonucléase par REC3.	<u>18</u>
<u>Figure 9</u>	Vue rapprochée du site de reconnaissance PAM et de la boucle de verrouillage du phosphate	<u>19</u>
<u>Figure 10</u>	Recherche du gène rpsL avec l'outil NCBI Gene.	<u>21</u>
<u>Figure 11</u>	Représentation de l'outil UniProt et son mode d'utilisation.	<u>23</u>
<u>Figure 12</u>	Représentation de l'outil Sequence Manipulation Suite « Reverse Translate » et son mode d'utilisation.	<u>25</u>
<u>Figure 13</u>	Représentation de l'outil CHOPCHOP.	<u>26</u>
<u>Figure 14</u>	Représentation de l'outil Primer3Plus.	<u>28</u>

<u>Figure 15</u>	Représentation de l’outil DrugBank.	<u>30</u>
<u>Figure 16</u>	Représentation de la séquence d’acide Aminé & Nucléotidique de sous unité ribosomal 30S codé par le gène rpsL.	<u>32</u>
<u>Figure 17</u>	Représentation des différents ARNg ciblant la sous unité ribosomal 30S du gène rpsL.	<u>32</u>
<u>Figure 18</u>	Représentation des ARNg ciblant la mutation K43/T43 de la sous unité ribosomal 30S du gène rpsL.	<u>33</u>
<u>Figure 19</u>	Représentation de la séquence d’acide Aminé & Nucléotidique de sous unité ribosomal 30S codé par le gène rpsL après substitution K43/T43.	<u>35</u>
<u>Figure 20</u>	Représentation des amorces PCR Primer3Plus de sous unité ribosomal 30S codé par le gène rpsL après substitution K43/T43.	<u>37</u>
<u>Figure 21</u>	Amorces PCR Primer3Plus de sous unité ribosomal 30S codé par le gène rpsL après substitution K43/T43	<u>Annexe I</u>

Liste des tableaux

Tableau	Titre	Page
<u>Tableau I</u>	Les séquences d’ARNg correspondant pour la mutation K43/T43.	<u>33</u>
<u>Tableau II</u>	Protocole PCR pour concevoir la séquence cible d'ADN Cas9 K43/T43.	<u>35</u>
<u>Tableau III</u>	Les antibiotiques qui ont différentes actions sur la bactérie E.coli K12.	<u>36</u>
<u>Tableau IV</u>	Les quatre primer qui sont en mesure de couvrir la région ou notre substitution est situé.	<u>38</u>
<u>Tableau V</u>	Liste des ARNg ciblant la séquence du gène rpsL	<u>Annexe I</u>
<u>Tableau VI</u>	Amorces PCR pour la validation de l’expérience de la substitution K43 à T43 du gène rpsL	<u>Annexe I</u>

Table des matières

-Remerciement	
-Dédicace	
-Résumé	
-Abréviation	
-Liste des figures	
-Liste des tableaux	
-Introduction.....	1
Chapitre I : Partie bibliographique	
I.1 GENERALITES :	ERROR! BOOKMARK NOT DEFINED.
I.1.1 l'immunisation	15
I.1.2 l'acquisition de l'immunité.....	15
<i>I.1 Caractéristiques :</i>	Error! Bookmark not defined.
I.3 LE SYSTEME CRISPR/CAS9 DE TYPE II, EXEMPLE DE STREPTOCOCCUS PYOGENES :	ERROR! BOOKMARK NOT DEFINED.
I.3.1 La protéine cas9 :.....	Error! Bookmark not defined.
I.3.2 La Structure bilobée de la Protéine Cas9 :.....	Error! Bookmark not defined.
I.3.3 L'ensemble de complexe effecteur CRISPR-Cas9 (un crARN-tracrARN /sgARN) : ..	Error! Bookmark not defined.
I.3.4 Mécanisme Cas9 :	Error! Bookmark not defined.
I.3.5 Les cassures par CRISPR-Cas :	Error! Bookmark not defined.
I.3.6 Régulation de l'activité endonucléase par REC3 :.....	Error! Bookmark not defined.
I.3.7 Rôle clé du brin d'ADN non cible dans le repositionnement HNH : ..	Error! Bookmark not defined.
defined.	
CHAPITRE II : MATERIELS ET METHODES	ERROR! BOOKMARK NOT DEFINED.
I.1 APPROCHE BIO-INFORMATIQUES :	ERROR! BOOKMARK NOT DEFINED.
II.1.1 NCBI gene :	Error! Bookmark not defined.
II.1.2 Uniprot :	Error! Bookmark not defined.
II.1.3 Sequence Manipulation Suite « Reverse Translate » :.....	Error! Bookmark not defined.
II.1.4 CHOPCHOP :	Error! Bookmark not defined.
II.1.5 Primer3Plus :	Error! Bookmark not defined.
II.1.6 DrugBank :	Error! Bookmark not defined.
CHAPITRE III : RESULTATS ET DISCUSSIONS	ERROR! BOOKMARK NOT DEFINED.
III.1 RESULTATS DES OUTILS NCBI GENE, UNIPROT ET REVERSE TRANSLATE :	ERROR! BOOKMARK NOT DEFINED.

III.2 Résultats des outils CHOPCHOP, Primer3Plus, DragBank et la validation de PCR :
Error! Bookmark not defined.

III.3 Discussions des résultats : **Error! Bookmark not defined.**

CONCLUSION..... ERROR! BOOKMARK NOT DEFINED.

REFERENCES BIBLIOGRAPHIQUES : ERROR! BOOKMARK NOT DEFINED.

Annexes

Introduction

Le système CRISPR-Cas9 (protéine associée CRISPR 9) qui fait référence à un ensemble de séquences d'ADN, que l'on trouve dans de nombreuses bactéries (~ 50%) et la plupart des archées (~ 90%) (KS et al, 2019). Ce système procaryote pouvait être appliqué comme outil d'édition du génome révolutionnaire dans les cellules vivantes de diverses espèces pour modifier, visualiser et identifier des séquences définies d'ADN et d'ARN. CRISPR-Cas pourrait éditer le génome par des mécanismes de réparation dirigés par homologie et de jonction d'extrémités non homologues. De plus, la modification du ciblage de l'ADN par la méthodologie CRISPR-Cas offre une opportunité pour le diagnostic, la thérapie et l'investigation des troubles génétiques (**Ghaemi et al., 2021**).

Le système CRISPR (Clustered Regular Interspaced Short Palindromic Repeats)-Cas9 est un système immunitaire adaptatif bactérien dont l'adaptation en tant que technologie d'édition du génome a révolutionné les sciences de la vie (**Jimin et al., 2022**). L'impact majeur de CRISPR-Cas9 en biomédecine et en pharmacie motive des études fondamentales du mécanisme sous-jacent, où la base moléculaire de la reconnaissance et du clivage de l'ADN reste floue. Cette technologie d'édition du génome a révolutionné la recherche biomédicale et la thérapie génique. De nombreuses recherches expérimentales ont montré que CRISPR-Cas9 peut éditer/ remplacer avec précision un gène individuel et peut également être utilisé pour cibler largement le génome de cibles à des fins thérapeutiques. Plus d'une douzaine d'essais cliniques ont été approuvés pour le traitement de diverses maladies génétiques, y compris les cancers, à l'aide de l'édition du génome basée sur CRISPR-Cas9 (**Baohong., 2020**).

Dans CRISPR-Cas9, l'endonucléase Cas9 s'associe aux ARN guides et clive les séquences d'ADN complémentaires (**Kesavan et al., 2021**). La reconnaissance et le clivage spécifiques au site ne nécessitent qu'une courte séquence de 2 à 5 nucléotides, connue sous le nom de motif adjacent protospacer (PAM), qui est trouvée immédiatement adjacente à la séquence d'ADN complémentaire (**Giuseppina La et al., 2021**).

Avec la découverte de la protéine associée à CRISPR (Cas9) et l'application de la technologie de recombinaison de l'ADN, les scientifiques commencent à le modifier pour être utilisé comme outil d'édition d'un gène spécifique dans un génome (**Makarova et al., 2020**). Il existe de nombreux systèmes d'administration différents de CRISPR/Cas9 dans les cellules, y compris des méthodes à médiation virale et sans médiation virale ; une méthode différente a des avantages et des inconvénients différents (**Doudna., 2020**).

La streptomycine est un antibiotique bactéricide de la classe des aminoglycosides agissant principalement sur les organismes gram-négatifs. C'est un inhibiteur de la synthèse des protéines. Il se lie au petit ARNr 16S de la sous-unité 30S du ribosome bactérien, interférant avec la vitesse et la précision de la traduction avec une éventuelle inhibition de la synthèse des protéines et la mort des cellules bactériennes (**Al-Wawi et al., 2019**).

Dans cette étude, en utilisant l'outil bio-informatique, nous avons fait grâce au système CRISPR-Cas9, une mutation dans le gène qui code la protéine de la sous-unité ribosomique « rpsl » en changeant le 43^{ème} acide aminé qui est une lysine (K) en une thréonine (T). Nous

avons identifié également le rôle de cette mutation dans la survie et le développement de la bactérie *E. coli* sur un support qui contient de la streptomycine.

Le système CRISPR-Cas9 offre des applications très importantes, parmi lesquelles : Il génère des modèles animaux et cellulaires de maladies plus rapidement et plus facilement (**H. Mou et al., 2015**), fournit de nouvelles informations sur le traitement des maladies (**AC Komor et al., 2017**). CRISPR à haut débit peut aider à sélectionner des candidats lors d'analyses fonctionnelles génétiques systématiques et à identifier de nouvelles cibles médicamenteuses (**O. Shalem et al., 2015**)(**T. Wang et al.,2014**). Il est ainsi précieux pour l'imagerie, la détection et le diagnostic basés sur les acides nucléiques (**GJ Knott et al.,2018**). Il y a encore une gamme applications de la technologie CRISPR-Cas9 dans l'agriculture, la science des matériaux et divers autres domaines (**PD Hsu et al., 2014**).

Chapitre I : Partie bibliographique

I.1 Généralités :

CRISPR est une technologie d'Édition, qui a été dérivée du système immunitaire des bactéries, utilise une nucléase Cas9 et un ARN guide complémentaire à une séquence de 20 nucléotides d'un gène pour induire des cassures double brin dans l'ADN. Ce qui résulte une modification spécifique de gène ciblé dans des cellules de plantes, d'animaux et d'humains (Jacques P., 2015). Ce système fonctionne dans les bactéries ainsi que chez les archées et se déroule en deux étapes majeures : la première étant

I.1.1I L'immunisation : au cours de laquelle on a la libération de l'élément génétique envahissant provenant de phage sous forme d'ADN viral après infections (ou d'une autre bactérie sous forme d'un plasmide après conjugaison). Ensuite cette acquisition de matériel génétique provoque l'expression des protéines Cas, qui interagissent et forment un complexe. Ce dernier se lie à l'ADN invasif au voisinage d'une séquence PAM (Proto-spacer Adjacent Motif) et le clive en fragment (Spacer). Le nouvel espaceur dérivé des éléments génétiques invasifs est incorporé dans le réseau espaceurs répétés CRISPR par la machinerie d'acquisition (cas1, cas2 et csn2). Fournissant ainsi un enregistrement génétique d'une infection antérieure qui permet à l'hôte d'empêcher l'invasion future du même envahisseur. Alors un locus CRISPR dans un système CRISPR-Cas comprend un ensemble de séquence répétitives entrecoupées de courtes séquence non répétitives (espaceurs) pour former le réseau espaceurs répétés CRISPR où s'intègre l'ADN viral. Ainsi qu'un ensemble de CRISPR associés (Cas), précédant le Cas l'opéron est la transactivation de l'ARN CRISPR (tracé ARN) qui code pour un ARN non codant unique présentant une homologie avec la séquence répétée (**figure 1-A**). La seconde étape c'est :

I.1.2 L'acquisition de l'immunité : dans laquelle on a la transcription ultérieure du réseau CRISPR donc le nouvel espaceur est co-transcrit avec tous les autres espaceurs en long ARN précurseur CRISPR (Pré-ARNcr) contenant des répétitions et des espaceurs (entretoises), sous le contrôle de la séquence Leader (L) qui agit comme promoteur. Le tracé ARN est transcrit séparément puis s'annele aux répétitions de pré-ARNcr pour la maturation de l'ARNcr par clivage de la RNase III puis rognage du 5' fin l'ARNc par des nucléases

(maturation : traitement de l'ARNase III et l'ajustement ARNc 5'). Pour réduire la séquence guide à 20nt, le Pré-CRISPR (pré-crRNA) est clivé en élément « partial repeat -Spacer-partial repeat » par un complexe multiprotéique de Cas. Au 5' de ARNcr contient l'espaceur, un court segment d'ARN qui complète une séquence (PAM) d'un élément génétique étranger et le 3' contient un morceau de la séquence de répétitions CRISPR. Donc les crRNAs liée sur le complexe Multi-Cas activent la machinerie CRISPR-Cas en s'hybridant par leur séquence Spacer à l'ADN cible dont ils ont été dérivés lors d'une précédente infection. Suite à toutes ces étapes le complexe Cas dégrade l'ADN (ou l'ARN) invasif, conduisant à son inactivation, élimination ainsi l'agent infectieux (déclenchant sa destruction spécifique de la séquence de l'ADN invasif par nucléase Cas d'une seconde infection) (**Jiang F et al., 2017**). L'analyse des séquences de phage adjacentes aux proto-Spacer (qui vont donner les Spacer) révèle la présence en 3' de séquences conservées appelées CRISPR motifs ou proto-Spacer adjacent motifs (ou PAM). Pour qu'une séquence de phage soit ciblée par CRISPR, il faut donc qu'elle soit au voisinage d'un motif PAM. Cependant, ce motif PAM ne fait pas partie du Spacer, il n'est pas présent dans le locus CRISPR. Ceci permettrait d'éviter l'action du système CRISPR sur le génome de l'hôte (**Horvath et al., 2010**) (**figure 1-B**).

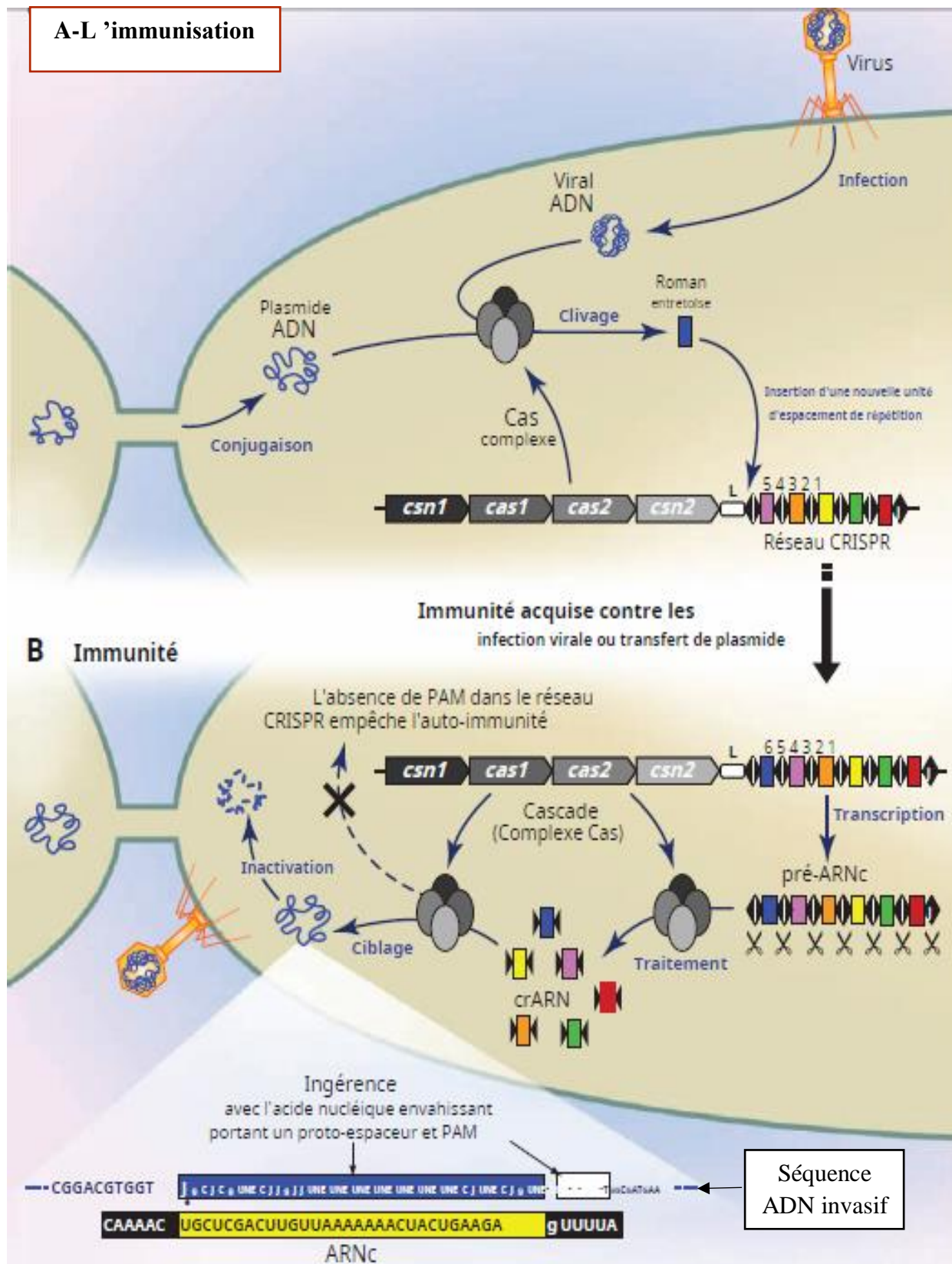


Figure 1 : Aperçu du mécanisme d'action CRISPR/Cas (Horvath et al., 2010).

I.1 Caractéristiques :

L'abréviation CRISPR désigne une famille de séquences répétées dans l'ADN, se caractérisant par des séries de répétitions directes et courtes de 20 à 37 nucléotides et régulièrement espacées par des séquences appelées « spacer », souvent uniques, d'une vingtaine de nucléotides. Les Cas (CRISPR-associated) sont des protéines associées à ce locus CRISPR. Ces répétitions ont été découvertes par hasard en 1987 chez *E.coli*, Lors du séquençage du gène *iap*, codant une enzyme protéolytique, il a été observé une séquence répétée dans la région en aval du gène (**Ishino Y et al., 1987**). Sont des systèmes de défenses bactériens adaptatifs protégeant contre toutes sortes d'agressions biologiques, incluant les phages, virus, plasmides et d'autres congénères. Les bactéries et archées ont mis en place au cours de leurs évolution le système de défense qui leur permette de détecter, de reconnaître et éliminer des ADN étrangers différents de leur propre ADN afin de survivre. Le système CRISPR-Cas9 est présent dans le génome de 90% des archées et de plus de 40% des bactéries et sont absents du génome des eucaryotes et de génome virus (**Horvath et al., 2010**).

Une courte séquence conservée, le motif situé à proximité immédiate de la séquence ciblée par l'ARNcr sur l'ADN envahisseur appelé PAM. Elle est essentielle dans la sélection et la dégradation de l'ADN cible dans la plupart des systèmes CRISPER-CAS. La fonction de ces répétitions été élucidée seulement en 2007, lorsqu'il a été montré que *Streptococcus thermophilus* pouvait acquérir une résistance contre les bactériophages en intégrant un morceau de génome du virus infectieux dans le locus CRISPR, permettant ainsi à la bactérie d'être protégée lors d'une réinfection, Il a alors été considéré que les bactéries possédaient un système semblable à l'immunité adaptative (**Barrangou R et al., 2007**). À l'heure actuelle, six différents types de locus CRISPR (I à VI) ont été mis en évidence, selon la classification actuelle de CRISPR-cas locus, et chacun emploie un ensemble unique de protéines Cas avec l'ARNc pour l'interférence CRISPR. Les systèmes CRISPER de type I et de type III qui utilisent un grand complexe protéique multi-Cas pour la liaison de l'ARNc et la dégradation de la séquence cible, contrairement aux systèmes CRISPR de type II utilisent une seule endonucléase d'ADN (Cas9), pour reconnaître les substrats d'ADNdb et cliver chaque brin avec un domaine nucléase distinct (HNH ou RuvC) et c'est le plus étudié et le plus utilisé (**Jiang F et al., 2017**). L'ADN étranger envahissant est coupé en petits fragments de 20 pb correspondant au « Spacer », incorporés dans le locus CRISPR entre les courtes séquences répétées. Les locus sont transcrits et génèrent des petits ARN (crRNA), qui vont servir de guide aux endonucléases, qui ciblent l'ADN étranger en se basant sur la complémentarité de séquences (**Jinek M et al., 2012**) (Ce

guide à double ARN ordonne à Cas9 de cliver tout ADN contenant une séquence cible complémentaire de 20 nucléotides (nt) et une PAM adjacente). Présence de multiples séquences “repeats” de 23 à 47 paires de bases, identiques avec peu ou pas de variation. Ces séquences forment le plus souvent des palindromes (haute stabilité). Leur nombre est généralement inférieur à 50. Présence de “spacer” de même taille mais de séquences très différentes entre les demi-“repeats” (longueur de l’ordre de 21 à 72 paires de bases). Ces séquences présentent un haut degré de similarité avec des séquences trouvées dans le génome de phages ou dans des plasmides. Chaque “spacer” est encadré par 2 demi-“repeats”. Présence de séquences “leader” de quelques centaines de paires de bases et d’un “terminal repeat” qui encadrent le locus CRISPR en présence de gènes divers Cas, Cas1 à Cas4, codant pour des protéines portant des fonctions typiques de nucléase, hélicase, polymérase et de liaison aux polynucléotides. Les locus CRISPR sont le plus souvent retrouvés sur le chromosome (plus rarement sur les plasmides). Et plusieurs locus CRISPR peuvent être trouvés dans un même microorganisme (bactérie ou archée). La fonction majeure du système CRISPR/Cas des microorganismes est de détruire le génome d’agents invasifs (phages, virus, plasmides) (**Horvath et al., 2010**).

Les différents systèmes présentent des points communs, au niveau de la première étape, l’acquisition de l’ADN étranger est essentielle afin de produire de nouveaux spacers qui sont intégrés à une des extrémités du locus CRISPR. Ce locus est transcrit en précurseur du crRNA (pre-crRNA) qui est ensuite maturé en crRNA et fixé au complexe CRISPR-interférence (complexe spécifique de chaque classe de systèmes CRISPR) (**Makarova KS et al., 2015**). Il convient de noter que le tracrRNA est requis pour la maturation du crRNA dans les systèmes de type II. Un sgRNA chimérique qui combine le crRNA et le tracrRNA en un seul transcrit d’ARN simplifie le système tout en conservant un clivage d’ADN spécifique à la séquence médié par Cas9 entièrement fonctionnel et donc les combinaisons des CRISPR et des gènes Cas forment les systèmes CRISPR/Cas (**Jiang F et al., 2017**). En modifiant la séquence d’ARN guide (espaceur) dans l’ARNc, ce système CRISPR-Cas9 simplifié à deux composants peut être programmé pour cibler pratiquement n’importe quelle séquence d’ADN d’intérêt dans le génome et générer davantage une rupture double brin à extrémité franche spécifique au site (ORD) (**Jinek M et al., 2012**). Pour assurer le clivage à un site spécifique, la Cas9 doit former un complexe avec le crRNA et le trans-activating crRNA (tracrRNA). Ce tracrRNA est requis pour la maturation du crRNA et fonctionne en présence de la RNase III. Le tracrRNA et crRNA forme ce que l’on appelle l’ARN guide (ou sgRNA). Le complexe CRISPR de type II est composé de l’endonucléase Cas9 et de ce sgRNA, qui va amener l’enzyme à la séquence cible

spécifique. Pour que ce complexe soit fonctionnel, la Cas9 requiert une séquence PAM (« protospacer adjacent motif ») conservée de 2 à 5 nt après le crRNA. Le clivage se fait spécifiquement 3 nt en amont de la séquence PAM, sur les deux brins d'ADN, et est effectuée par les domaines endonucléasiques RuvC et HNH de la Cas9. Ce mécanisme a été très bien caractérisé chez *Streptococcus pyogenes* (Jinek M et al., 2012).

I.3 Le système CRISPR/Cas9 de type II, exemple de *Streptococcus pyogenes* :

Il existe trois types différents de systèmes CRISPR-Cas, les types I, II et III, qui utilisent des mécanismes moléculaires différents et les machines impliquées varient d'un système à l'autre pour cibler l'ADN invasif et le dégrader. Bien que tous les systèmes CRISPR-Cas suivent les mêmes étapes et principes généraux. Le motif PAM, courte séquence adjacente à la séquence cible de l'ADN invasif, joue un rôle décisif dans les systèmes CRISPR de types I et II. Dans le système de type I utilisent des complexes de protéines Cas multi-sous-unités, dans les systèmes de le type II, une seule protéine effectrice Cas (Cas9) est nécessaire pour la reconnaissance de l'ADN invasif et sa dégradation, une caractéristique qui s'est révélée extrêmement utile pour les applications sur l'édition du génome (Jennifer A et al., 2014). Systèmes CRISPR-Cas présents dans *S. pyogènes* appartiennent au type IC (classe 1) et II-A (classe 2).et alors La bactérie *Streptococcus pyogenes* utilise un système CRISPR/Cas9 de type II, leur locus CRISPR en amont contient (tracrARN) codent pour un petit ARN appelé transactivating crARN (tracrARN) se lie à Cas9 et s'hybride aux "repeats" du pre-crRNA. Ces hybrides tracrARN:pre-crARN sont clivés par la ribonucléase III (RNase III, pour générer des hybrides tracrARN:crARN, ces derniers restent fixés sur le Cas9, ce qui provoque le recrutement de L'ADN invasif avec la séquence « Spacer » du crARN. La protéine Cas9 possède deux domaines catalytiques HNH et Ruv-C qui coupent respectivement le brin complémentaire et le brin codant de l'ADN invasif suite à ouverture du double brin de l'ADN du génome invasif au niveau de la région complémentaire au crARN (formation de la boucle), générant ainsi une coupure double brin (**figure2**). Cette coupure se fait trois nucléotides en amont (ou en aval selon le brin) de la séquence PAM, Pour le système CRISPR/Cas9 de *S. pyogènes* cette séquence PAM est NGG* sur le brin ciblé (CCN* sur le brin complémentaire), et la séquence ciblée par le crARN est GN20GG (Horvath et al., 2010).

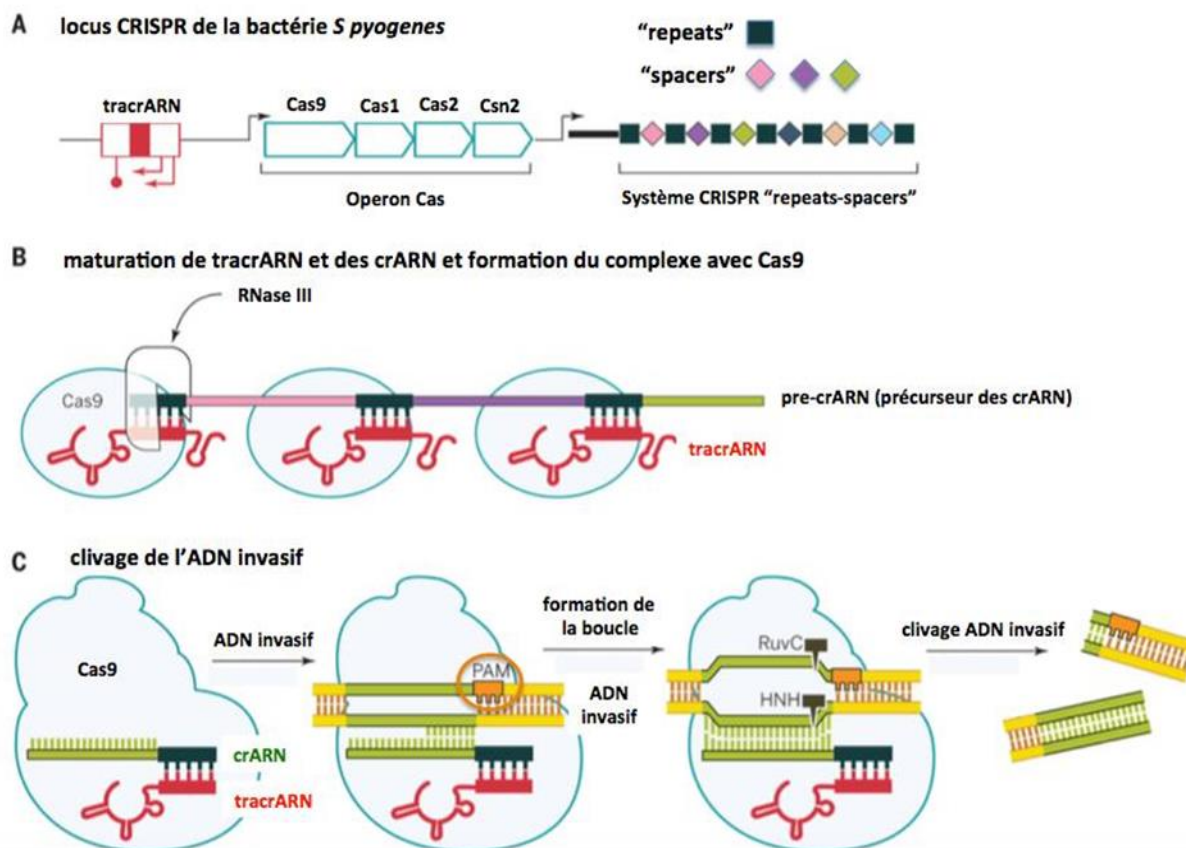


Figure 2 : Biologie du système CRISPR-Cas de type II-A (Doudna JA et al., 2014).

I.3.1 La protéine cas9 :

Cas9 (CRISPR associated protein 9) associée à des répétitions palindromiques courtes régulièrement espacées (CRISPR), est une protéine d'origine bactérienne aux propriétés antivirales, trouvée dans les systèmes CRISPR de type II (*S. pyogenes*Cas9), cette dernière est une grande endonucléase d'ADN multidomaine et multifonctionnelle (1 368 acides aminés). Elle coupe l'ADN de 3 pb en amont du PAM via ses deux domaines de nucléase distincts : un domaine de nucléase de type HNH qui clive le brin d'ADN complémentaire de la séquence d'ARN guide (brin cible) et un domaine de nucléase de type RuvC responsable du clivage de l'ADN, brin opposé au brin complémentaire (brin non cible) (Chen H et al., 2014). Cas9 fonctionne en conjonction avec les ARN CRISPR (ARNcr) et un ARNc transactivant (TracrARN) pour médier l'immunité spécifique à la séquence contre les bactériophages et d'autres éléments génétiques mobiles, donc la protéine cas9 est guidée dans son action par deux ARN, le tracrARN et le crARN adapté à l'ADN invasif ciblé, l'hybridation de ses deux tracrARN-crARN, construit un ARN chimérique unique, appelé ARN guide. Et le complexe ribonucléoprotéique résultant reconnaît et clive les molécules d'ADN contenant des séquences complémentaires d'un segment guide de 20 nucléotides dans le crRNA (figure 2). Cet ARN

guide qui peut être synthétisé en laboratoire permet de cibler tout ADN en adaptant la séquence du crARN . En raison de sa programmabilité et pour ce très important point les applications de CRISPR/Cas9 à l'édition du génome est révélé car La Cas9 coupe l'ADN à la position souhaitée, définie par la séquence de la partie crARN du guide. Et C'est la base de l'édition génomique (**Horvath et al., 2010**).

I.3.2 La Structure bilobée de la Protéine Cas9 :

C'est grâce à sa structure unique qu'elle va se lier à l'ARN guide (tracrARN:crARN ou sa forme fusionnée le sgARN). Ce duo protéine:ARN se liera ensuite à l'ADN. La protéine subit d'intenses réarrangements conformationnels durant sa liaison à l'ARN guide. La structure de Cas9 est formée par deux lobes distincts, le lobe de reconnaissance alpha-hélicoïdal (REC) qui joue un rôle clef dans l'association au guide ARN. Est composé de trois domaines alpha-hélicoïdaux (Hel-I, Hel-II et Hel-III) et ne partage pas de similitude structurelle avec d'autres protéines connues. Et le lobe de nucléase (NUC) contenant les domaines conservés de HNH et de nucléase RuvC, ainsi que le domaine C-terminal plus variable (CTD) (**Jinek M et al.,2014**). Des études de mutagenèse de Cas9 montrant que la mutation du domaine HNH (H840A) ou du domaine RuvC (D10A) convertit Cas9 en une nickase, tandis que la mutation des deux domaines de nucléase de Cas9 (appelée "Cas9 mort" ou dCas9) laisse son ARN -capacité de liaison à l'ADN guidée intacte tout en abolissant l'activité de l'endonucléase (**Jinek M et al.,2012**). Les deux lobes sont en outre reliés par deux segments de liaison, l'un formé par le pont riche en arginine hélice et l'autre par un lieu désordonné (résidus 712–717). Le CTD allongé affiche également un pli spécifique à Cas9 et contient des sites d'interaction PAM requis pour l'interrogation PAM. Cependant, cette région de reconnaissance PAM est largement désordonnée dans la structure apo-Cas9, indiquant que l'enzyme apo-Cas9 est maintenue dans une configuration inactive, incapable de reconnaître l'ADN cible avant de se lier à un ARN guide (**Sternberg SH et al., 2014**). Le domaine HNH de la Cas9 va permettre de cliver le brin complémentaire (le domaine de la nucléase HNH adopte le pli $\beta\beta\alpha$ -métal caractéristique partagé par d'autres endonucléases HNH et est très susceptible d'utiliser un mécanisme à un ion métallique pour le clivage de l'ADN du brin cible) (**Yang W., 2008**). Et le domaine RuvC de la Cas9 lui va cliver le brin opposé au brin complémentaire (le domaine de la nucléase Cas9 RuvC est susceptible d'utiliser un deux métaux. Mécanisme catalytique ionique pour le clivage du brin d'ADN non cible) (**Jinek M et al., 2014**). Le ciblage de l'ADN guidé par l'ARN CRISPR-Cas9 peut être découplé de l'activité de clivage en mutant les résidus catalytiques dans les domaines de la nucléase HNH et RuvC, ce qui en fait une plate-forme polyvalente pour de

nombreuses autres applications au-delà de l'édition du génome (**Horvath et al., 2010**). Les formes catalytiquement inactives de Cas9 peuvent également être utilisées pour la régulation programmable de la transcription et la visualisation des locus génomiques.

I.3.3 L'ensemble de complexe effecteur CRISPR-Cas9 (un crARN-tracrARN /sgARN) :

Le sgRNA comprend une séquence à double ARN dérivée de l'ARN CRISPR et un transcrit séparé (tracrRNA) qui lie et stabilise la protéine Cas9. Permet, comme son nom l'indique, de diriger une enzyme (appelée Cas9) vers la portion de la séquence génétique sur laquelle elle va agir. Ainsi la reconnaissance et le clivage de l'ADN spécifiques au site nécessitent que Cas9 soit assemblé avec un ARN guide (un crARN-tracrARN natif ou un sgARN) pour former un complexe de surveillance active de l'ADN. Celui-ci est complémentaire de l'ADN cible sur une longueur de 20 nucléotides, c'est cette séquence qui sera artificiellement modifiée pour permettre de désigner la cible choisie. La séquence d'espacement de 20 nt de l'ARNcr confère une spécificité de cible d'ADN, et l'ARNtracr joue un rôle crucial dans le recrutement de Cas9 (les crARN ciblent aussi les brins sens et anti sens de l'ADN invasif). S'il y a complémentarité, avec les 12 premières bases (la séquence appelée seed), l'invasion de l'ARN peut avoir lieu en même temps que l'ADN se déroule pour former une boucle R (c'est à dire une boucle composée d'ADN simple brin formé par hybridation avec un ARN). La liaison de l'ARN guide avec l'apocas9 conduit Cas9 à subir un réarrangement structurel substantiel d'une conformation inactive à une conformation compétente pour la reconnaissance de l'ADN, donc le chargement de l'ARN guide est un régulateur clé de la fonction enzymatique Cas9 (**Jinek M et al., 2014**) (**figure6**). Cas9 fait des interactions avec le sgRNA, Plus précisément, il établit de nombreux contacts directs avec le duplex répétition-antirépétition et la tige-boucle 1, ainsi que la région de liaison entre les tiges-boucles 1 et 2 via Hel-I, l'hélice de pont riche en arginine et le domaine CTD. Cas9 établit des contacts beaucoup moins étendus avec la tige-boucle 2 du sgRNA, principalement via ses domaines RuvC et CTD. Aucune interaction protéine-ARN n'a été observée pour la tige-boucle 3 dans la structure Cas9-ARNsg en raison de l'absence d'un 3'queue de tracrRNA dans la construction de sgRNA utilisée pour la cristallographie. Des études biochimiques démontrent que les sgARN dépourvus de la région de liaison et des boucles souches 2 et 3 sont toujours capables de déclencher le clivage de l'ADN médié par Cas9, bien qu'avec moins d'efficacité, alors que la suppression de la boucle souche 1 abolit complètement le clivage (**Jinek M et al., 2012**). Et des études fonctionnelles montrent que les boucles souches 2 et/ou 3 sont nécessaires pour une activité Cas9 robuste in vivo (**Nishimasu H et al., 2014**).

Ces observations indiquent que le duplex répété-anti-répétition et la tige-boucle 1 sont indispensables à la formation du complexe Cas9-sgRNA, alors que le lieur, la tige-boucle 2 et la tige-boucle 3 ne sont pas nécessaires à la fonction, mais peuvent stabiliser l'ARN guide. Se liant pour favoriser la formation de complexes actifs (**Jinek M et al., 2013**).

I.3.4 Mécanisme Cas9 :

La protéine Cas9 est une apoenzyme bilobée, inactive c'est-à-dire qu'elle ne peut pas remplir sa fonction de réalisation d'une certaine réaction biochimique, et elle est incomplète jusqu'à ce qu'elle rejoigne d'autres molécules appelée cofacteurs. Alors les deux macromolécules biologique (bilobé : le lobe de nucléase NUC et le lobe de nucléase RUVC) de la protéine Cas9 et l'ARN guide interagissent pour former un complexe capable d'identifier des séquences cibles à haute sélectivité. La sélection d'une séquence spécifique d'ADN est l'étape clé dans l'édition du génome d'un organisme. Cet enzyme endonucléase spécialisée pour couper l'ADN avec deux zones de coupe active, une pour chaque brin de la double hélice (**Philippe et al., 2010**). Donc la protéine Cas9 est responsable de la localisation et du clivage de l'ADN cible à la fois dans les systèmes naturels et artificiels de CRISPR-Cas9. Elle se compose par six domaine (REC I, REC II, Bridge Helix, PAM Interacting, HNH et RuvC) (**Jinek M et al., 2014**) (**Nishimasu H et al., 2014**). Le domaine Rec I est le plus grand et est responsable de la liaison de l'ARN guide. Le rôle du domaine REC II n'est pas encore bien compris. L'hélice de pont riche en arginine est cruciale pour initier l'activité de clivage lors de la liaison de l'ADN cible (**Nishimasu H et al., 2014**). Le domaine PAM-Interacting confère une spécificité PAM et est donc responsable de l'initiation de la liaison à l'ADN cible (**Anders C et al., 2014**) (**Jinek M et al., 2014**) (**Nishimasu H et al., 2014**) (**Sternberg SH et al., 2014**). Les domaines HNH et RuvC sont des domaines nucléases qui coupent l'ADN simple brin. Ils sont très homologues aux domaines HNH et RuvC présents dans d'autres **protéines** (**Jinek M et al., 2014**) (**Nishimasu H et al., 2014**) (figure 3).

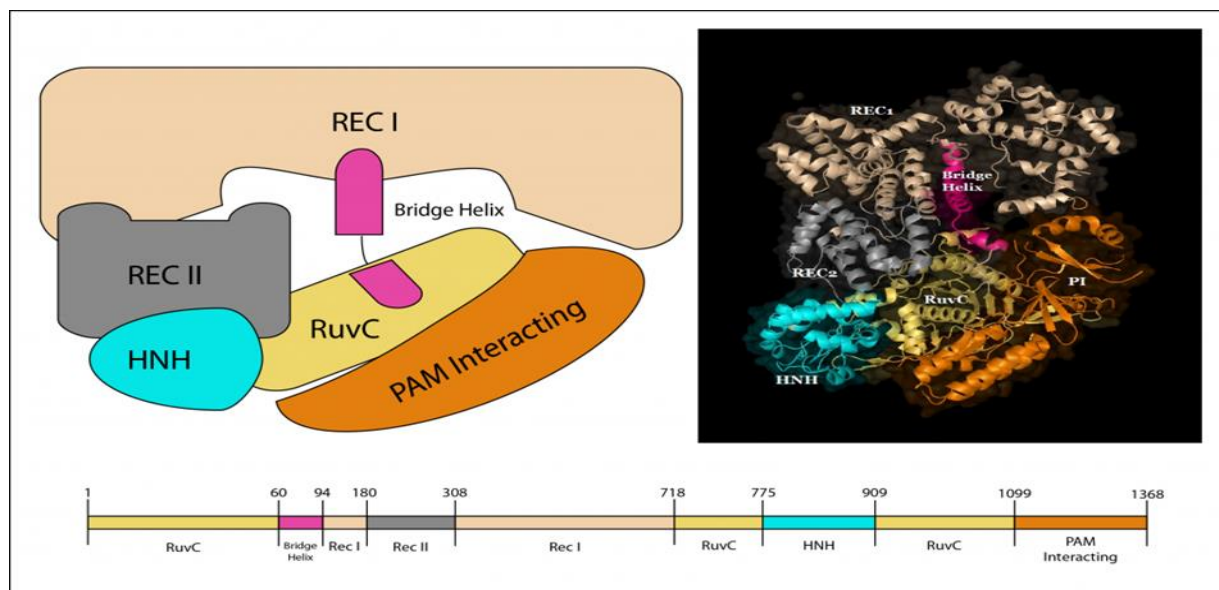


Figure 3 : La protéine Cas9 est composée de six domaines (Jinek M et al., 2014.).

L'ARN guide est conçu pour avoir une extrémité de 5' qui est complémentaire à la séquence d'ADN cible. Et la protéine Cas9 reste inactive en l'absence d'ARN guide. Dans les systèmes CRISPR modifiés, l'ARN guide est composé d'un seul brin d'ARN qui forme une forme en T composée d'une tétraboucle et de deux ou trois boucles de tige. Cet ARN guide artificiel se lie à la protéine Cas9 et, lors de sa liaison, induit un changement conformationnel dans la protéine. Le changement conformationnel convertit la protéine inactive en sa forme active (Jinek M et al., 2014)(figure4)(figure5).

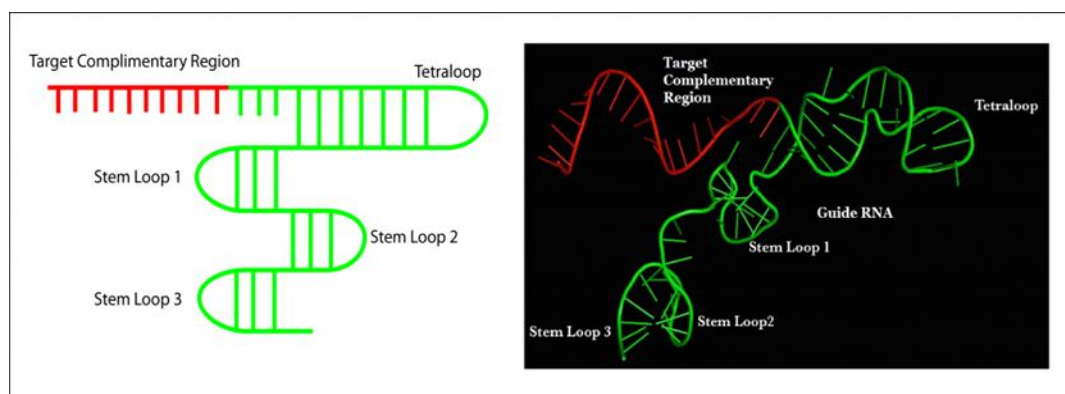


Figure 4 : L'ARN guide (Anders C et al., 2014.).

Donc Une fois la protéine Cas9 activée, elle recherche l'ADN cible en se liant à des séquences qui correspondent à sa séquence de motif adjacent (PAM) de protospacer (Sternberg SH et al., 2014) (figure6). Un PAM est une séquence à deux ou trois bases situées dans un nucléotide en aval de la région complémentaire à l'ARN guide. Les PAM ont été identifiés dans tous les systèmes CRISPR, et les nucléotides spécifiques qui définissent les PAM sont spécifiques à la

catégorie particulière du système CRISPR (Mojica FJM et al., 2009). Le PAM chez *Streptococcus pyogenes* est de 5'-NGG-3' (Jinek M et al., 2012). Lorsque la protéine Cas9 trouve une séquence cible potentielle avec le PAM approprié, la protéine fait fondre les bases immédiatement en amont du PAM et les associe à la région complémentaire de l'ARN guide (Sternberg SH et al., 2014). Si la région complémentaire et la région cible se couplent correctement, les domaines nucléases RuvC et HNH couperont l'ADN cible après la troisième base nucléotidique en amont du PAM (Anders C et al., 2014).

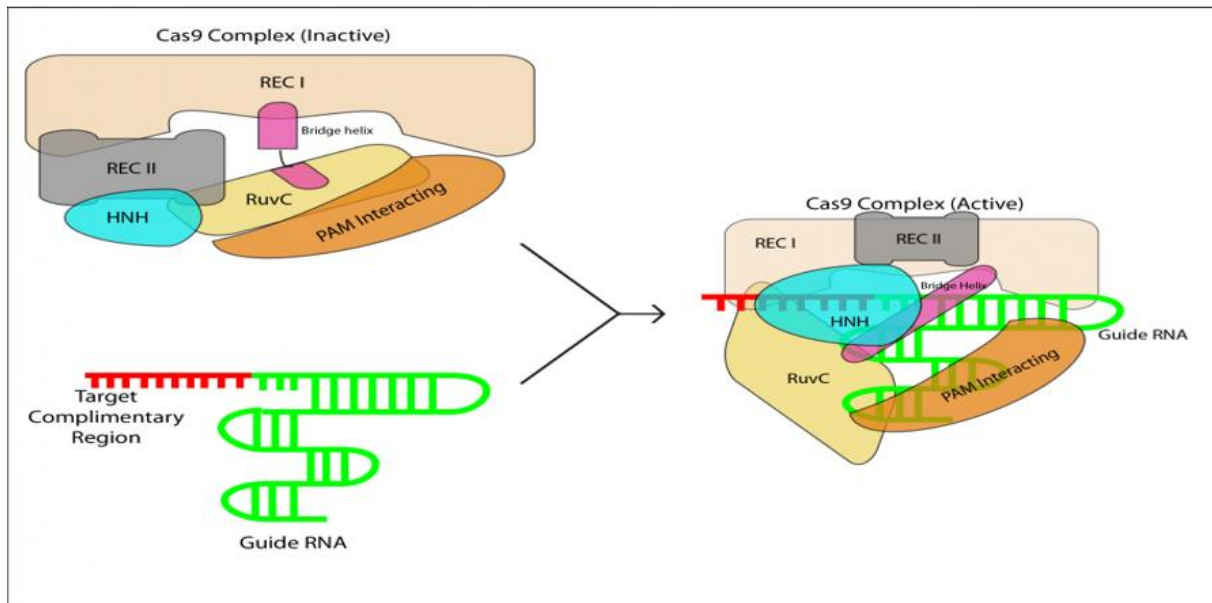


Figure 5 : Mécanisme d'activation de la protéine Cas9 (Anders C et al., 2014) et (Jinek M et al., 2014).

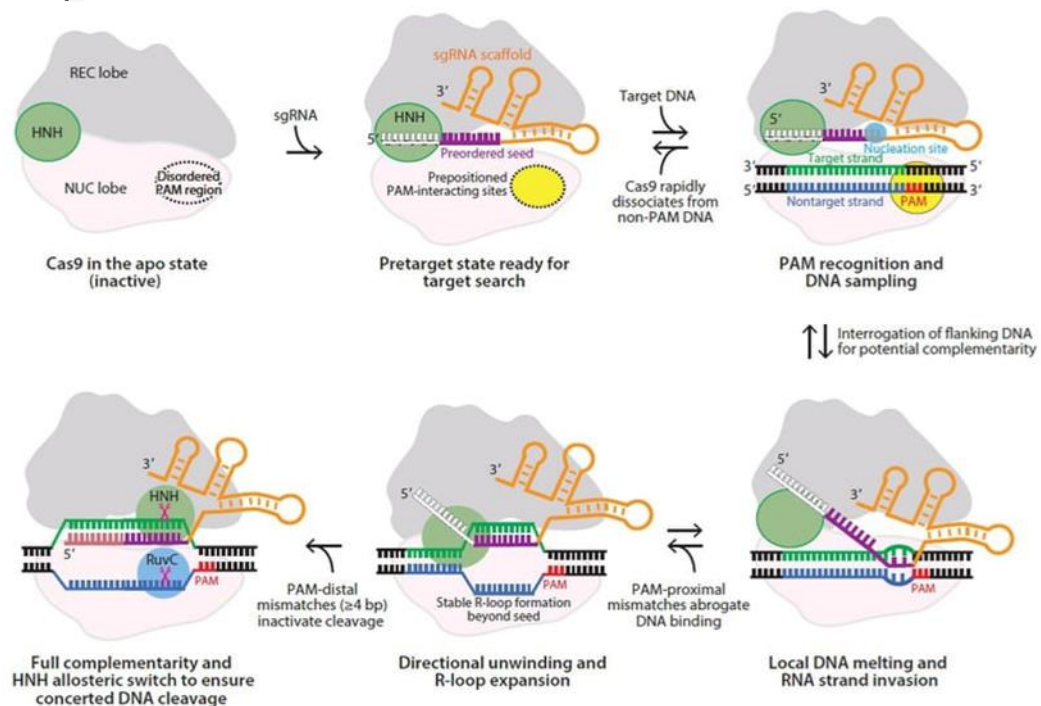


Figure 6 : Représentations schématiques des mécanismes proposés de reconnaissance et de clivage de l'ADN cible médiés par CRISPR-Cas9. (Jiang F et al., 2017).

I.3.5 Les cassures par CRISPR-Cas :

Le système CRISPR/Cas9 induit des cassures double brin de l'ADN et c'est là le point essentiel de son utilisation pour l'édition génomique. Cela permet de modifier spécifiquement le gène ciblé pour permettre de comprendre leur rôle dans des cellules de plantes, d'animaux et d'humains. Des variantes de la technique permettent également de réduire ou d'augmenter l'expression d'un gène choisi. Cette technologie peut donc être utilisée non seulement pour révéler le rôle d'un gène, mais aussi de développer des thérapies pour des maladies héréditaires et acquises. Ces cassures double brins produit par Cas9 peuvent être réparées spontanément par des mécanismes de réparation de l'ADN médiés par l'hôte. En l'absence d'un modèle de réparation, la voie de jonction des extrémités non homologues (NHEJ) prévalente et sujette aux erreurs est activée et provoque des insertions et des suppressions aléatoires ou même des substitutions au site DSB, entraînant fréquemment la perturbation de la fonction génique. En présence d'une matrice donneuse qui doit contenir, de chaque côté de la coupure de l'ADN ciblé, des séquences homologues allant de 50 à 2000 nucléotides, la voie de réparation dirigée par homologie (HDR) sans erreur peut être initiée pour créer les mutations souhaitées par recombinaison homologue, qui fournit la base pour effectuer une modification génétique précise, telle que l'inactivation, la suppression, la correction ou la mutagenèse du gène (**Jacques P., 2015**) (**figure 7**). La modification des gènes par recombinaison homologue est la plus souvent utilise par les expérimentateurs car la modification par recombinaison non homologue est difficile, et cela en raison que la réparation non homologue (NHEJ) prédomine dans la correction des cassures d'ADN, dans la plupart des situations, et cette préférence varie selon les types de cellules. La recombinaisons homologues conduisent à des modifications très précises, ce qui est très utile pour corriger des gènes impliqués dans des maladies héréditaires. Elles peuvent également permettre d'introduire au niveau d'un gène une séquence codant pour

un rapporteur afin de détecter la protéine qui en résulte (Jacques P., 2015).

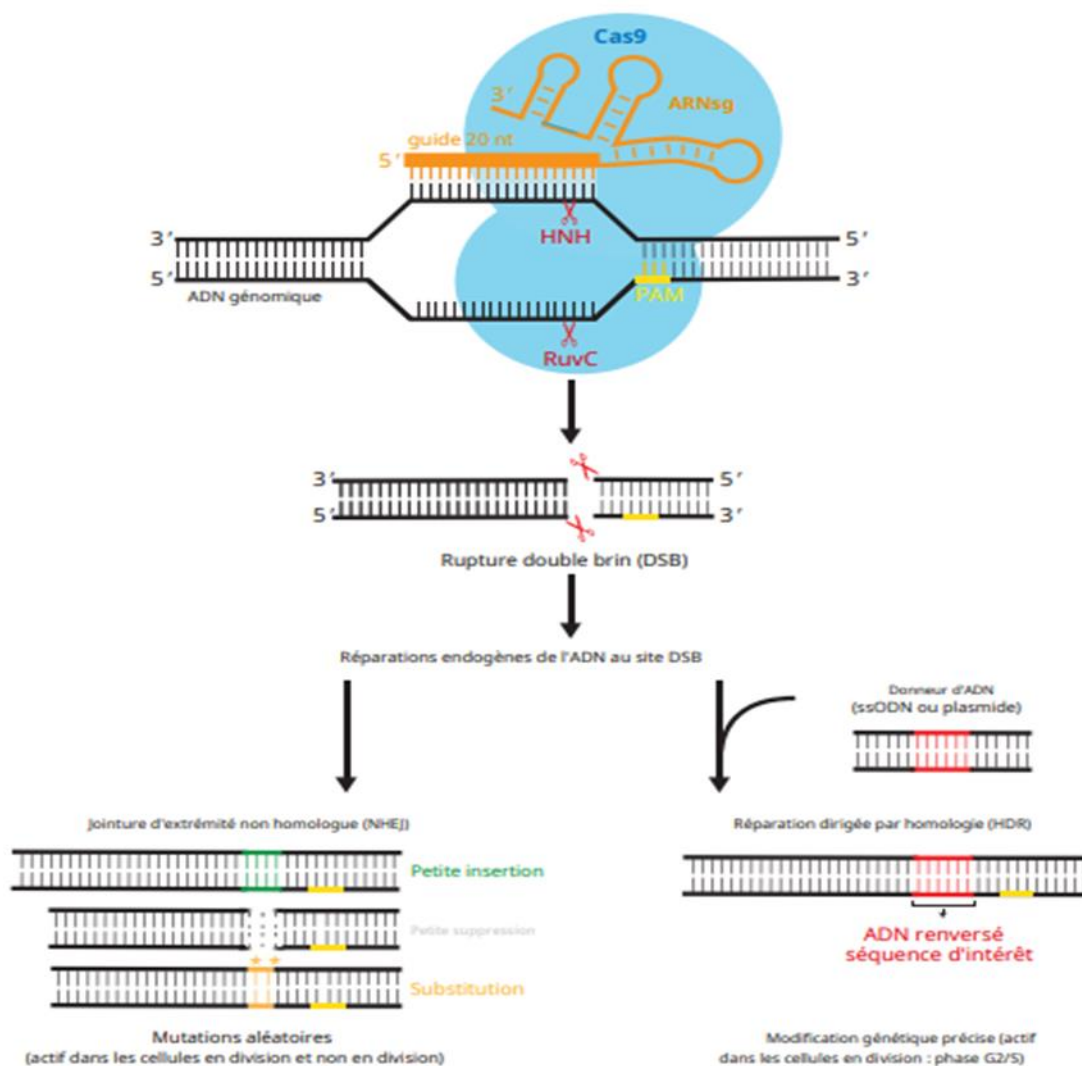


Figure 7 : les cassures par CRISPR-Cas et la Recombinaison (Jiang.F et al., 2017).

I.3.6 Régulation de l'activité endonucléase par REC3 :

REC3 agit comme un senseur de la qualité de l'hybridation entre l'ARNg et l'ADN, notamment dans la partie la plus distale au PAM (Chen JS et al., 2017). La présence de mésappariements inhibe le passage de REC3 vers sa forme active. Si le REC3 est activé, il agit comme un effecteur allostérique autorisant les changements conformationnels conjugués de REC2 et HNH. S'il est inactif, REC2 masque le groupement phosphate de la liaison phosphodiester pouvant être clivé par HNH. Le passage de REC3 à l'état actif déclenche un changement de conformation de REC2 puis de HNH. Une fois le domaine HNH repositionné, il stimule l'activité endonucléasique du domaine RuvC, permettant ainsi une coupure synchronisée des deux brins

d'ADN (Chen JS et al., 2017) (Figure 8).

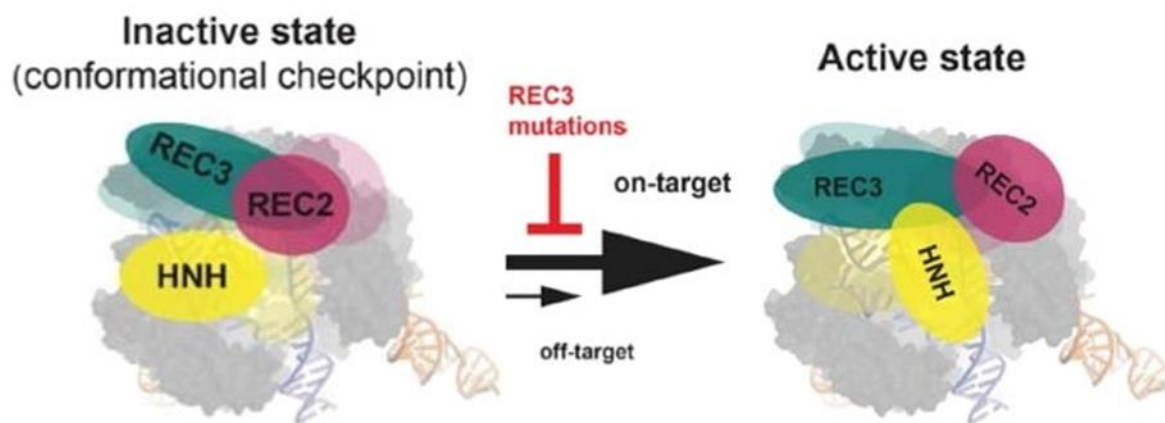


Figure 8 : Régulation de l'activité endonucléase par REC3. (Chen JS et al., 2017).

I.3.7 Rôle clé du brin d'ADN non cible dans le repositionnement HNH :

Dans la structure cristalline de Cas9 liée à un tel duplex partiel, le site actif HNH n'est pas positionné à proximité du phosphate scissile. Une raison possible de cette conformation inactive pourrait être due à l'absence d'un 3'fin du sgRNA utilisé pour la cristallisation, ce qui est essentiel pour stabiliser l'état actif fermé de Cas9. Selon les études FRET (Fluorescence Resonance Energy Transfer), la liaison de l'ADN non cible stabilise probablement le domaine intrinsèquement dynamique de la nucléase HNH dans une conformation fermée plutôt qu'en équilibre dynamique entre les conformations ouverte et fermée (Sternberg SH et al., 2015). Par conséquent, l'omission d'ADN non cible en amont du PAM dans la reconstitution complexe entraîne une conformation inactive, ce qui identifie le brin non cible comme un déterminant clé de l'activation conformationnelle de Cas9 (Palermo G et al., 2016). Le domaine HNH peut interagir avec Hel-II lors de la liaison de l'ADN db, ce qui joue un rôle important dans la conformation fermé et donc activé pour la scission ultérieure du brin d'ADN (Jiang et al., 2013) (Figure 9).

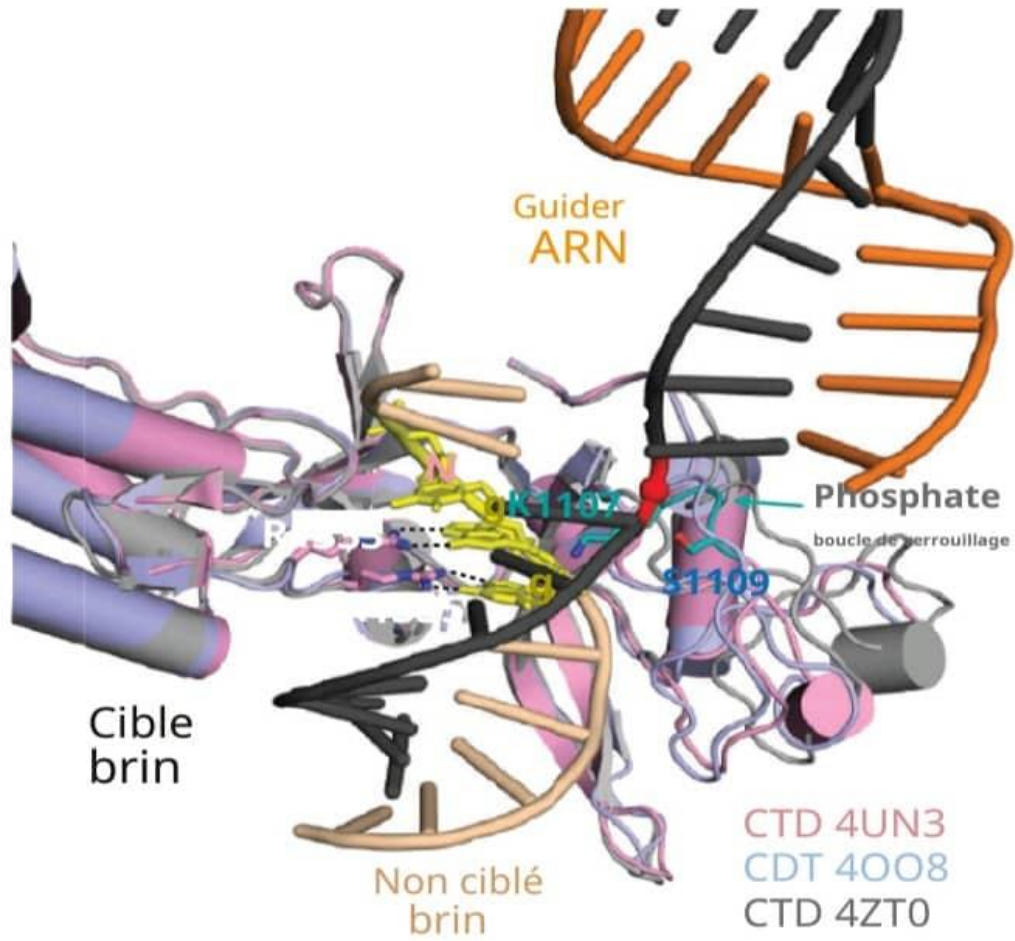


Figure 9 : Vue rapprochée du site de reconnaissance PAM et de la boucle de verrouillage du phosphate (Jiang F et al., 2017)

Chapitre II : Matériels et Méthodes

Il s'agit d'une étude Gene Editing In-Silico d'une mutation du génome (K43T) vers le gène rpsL par un seul changement de base nucléotidique une adénine ("A") à une cytosine ("C"). Ce changement amène l'ADN à coder pour une Thréonine « T » au lieu d'une Lysine « K » permettant ainsi à la bactérie de survivre sur un milieu qui empêcherait normalement sa croissance.

Dans la présente étude nous allons effectuer notre approche par différents outils Web et locaux à savoir :

- La recherche de L'ARNg par l'ensemble enzymatique Cas9 guider par un RNAGuide (RNAG). Ce complexe se lie au gène ciblé avec la protéine Cas9 fera une coupure double brin dans l'ADN, une fois que la protéine Cas9 a coupé le gène que nous allons effectuer une modification.
- La conception In-Silico des amorces PCR ciblant l'amplification de cette mutation
- La recherche des interactions pharmacogénomique causer par cette mutation chez E.Coli

I.1 Approche Bio-informatiques :

II.1.1 NCBI gene :

Le National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/gene/947845>), est l'institut national américain pour l'information biologique moléculaire. NCBI conduit des recherches dans la biologie informatique, développe des logiciels pour analyser des données de génome et fournir des informations biomédicales. Les ressources génétiques du NCBI comprennent des collections de séquences nucléotidiques organisées utilisées comme références, des grappes de séquences pour prédire et étudier les homologues, ainsi que diverses bases de données et outils pour l'étude de l'expression des gènes.

A l'aide de l'outil NCBI Gene nous allons rechercher le gène rpsL pour avoir la séquence nucléotidique (FASTA), (**voir la Figure-10**)

Search Page: The user searches for "MG1655" in the "Gene" field. The page title is "Rechercher le gène". A description states: "Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeq), maps, pathways, variants, phenotypes, and links to genome, phenotype, and locus-specific resources worldwide."

Search Results: The search results show "NC_009133" as the top result. A callout box says "Sélectionner le gène".

Gene Details: The details for "NC_009133 (3474178..3474552, complement)" are shown. A callout box says "Sélectionner le gène".

Genomic Context: The genomic context shows the "NC_009133" sequence. A callout box says "Cliquer sur FASTA".

FASTA Sequence: The FASTA sequence is displayed: "Escherichia coli str. K-12 substr. MG1655, complete genome". The sequence is: "NCBI Reference Sequence: NC_009133.3".

Figure 10 : Recherche du gène rpsL avec l’outil NCBI Gene.

II.1.2 Uniprot :

UniProt est la principale ressource mondiale de haute qualité, complète et librement accessible en matière de séquences protéiques et d’informations fonctionnelle (<http://www.uniprot.org>). Son nom dérive de la contraction de Universal Protein Resource (Base de Données Universelle des Protéines). C'est une base de données ouverte, stable et accessible en ligne. Uniprot est une base annotée, hiérarchisée où chaque séquence est accompagnée d'un ensemble riche de métadonnées et de liens vers de nombreuses autres bases de données.

A l’aide de l’outil UniProt, nous avons cherché la protéine codé par le gène rpsL, sur la liste des résultats nous avons sélectionné le gène rpsL (P0A753) exprimé par la bactérie E.Coli , pour avoir tous les informations sur cette protéine et télécharger la séquence FASTA d'acides aminés, (**voir la Figure-11**)

Align

1. Enter either protein sequences in FASTA format or UniProt identifiers into the form field, for example: TH1_HUMAN TH1_PIC

2. Click the Align button.

Rechercher le gène

Did you mean **rs1312**?

Entry	Gene names	Gene names	Organism	Length
<input type="checkbox"/> P0A7S3	30S ribosomal protein S12	rpsL, rplA, R1342, R13204	Escherichia coli (strain K12)	124
<input type="checkbox"/> P1A471	30S ribosomal protein S12	rpsL, rplA, rplA, R5161100	Bacillus subtilis (strain 168)	124
<input type="checkbox"/> Q55943	30S ribosomal protein S12	rpsL, TH13107	Thermus thermophilus (strain ATCC 29634 / DSM 175 / HB8)	132
<input type="checkbox"/> P12793	30S ribosomal protein S12	rpsL, rpl12	Thermus thermophilus	132
<input type="checkbox"/> Q9M712	30S ribosomal protein S12	rpsL, R1A2355	Rhodospirillum rubrum (strain ATCC BAA-98 / DSM200)	123

Sélectionner le gène

Display

Protein: **30S ribosomal protein S12**

Gene: **rpsL**

Organism: **Escherichia coli (strain K12)**

Status: **Reviewed** - Annotation score: ********* - Experimental evidence at protein level

FUNCTION

ribS4 and ribS5 plays an important role in translational accuracy. Interacts with nascent bases of the 16S rRNA that are involved in RNA decoding. Located at the interface of the 30S and 50S subunits. It traverses the body of the 30S subunit and probably holds the ribosomal structure together. The composed cluster of proteins S8, S12 and S17 appears to hold together the shoulder and platform of the 30S subunit. (By similarity).

Cryo-EM studies suggest that S12 contacts the E7 Tu bound P0ta in the A site during codon recognition. The contact is most likely broken as the aminoacyl-tRNA moves into the peptidyl transferase center in the 30S subunit. (By similarity)

Mass spectrometry

Molecular mass is 12651.3 Da. Determined by MALDI. (1 Publication)

Align

How to use this tool

Align two or more protein sequences with the Clustal Omega program (see also this FAQ) to view their characteristics alongside each other.

Rechercher le gène

1. Enter either protein sequences in FASTA format or UniProt identifiers into the form field, for example: TH1_HUMAN TH1_PIC

2. Click the Align button.

Display

Sequence status: **Complete**

Sequence processing: **The displayed sequence is further processed into a mature form.**

POA7S3: [UniProt] **FASTA** **ADD to basket**

Cliquer sur FASTA

```

10      20      30      40      50
1ATVQLVW PARAKAKS VVLEAQQK RQCTRVTT THQNSALA
60      70      80      90     100
KQKQKATK FQYFQDGE QRLQEVNT LTRGAVDL RQKRVNIG
110     120
ALGQDQK IYQAKQKQK IYPA
    
```

Mass spectrometry

Molecular mass is 12651.3 Da. Determined by MALDI. (1 Publication)

Figure 11 : Représentation de l'outil UniProt et son mode d'utilisation.

II.1.3 Sequence Manipulation Suite « Reverse Translate » :

Sequence Manipulation Suite est une collection de programmes JavaScript pour générer, formater et analyser de courtes séquences d'ADN et de protéines. Il est couramment utilisé par les biologistes moléculaires, pour l'enseignement et pour les tests de programmes et d'algorithmes. L'option « Reverse Translate » de l'outil SMS (https://www.bioinformatics.org/sms2/rev_trans.html) accepte une séquence protéique en entrée et utilise une table d'utilisation des codons pour générer une séquence d'ADN représentant la séquence codante non dégénérée la plus probable. Une séquence consensus dérivée de tous les codons possibles pour chaque acide aminé est également renvoyée. Cette option est utilisée lors de la conception d'amorces PCR pour s'hybrider à une séquence codante non séquencée d'une espèce apparentée.

A l'aide de l'outil Sequence Manipulation Suite « Reverse Translate » nous allons rechercher la séquence nucléotidique codé pour la protéine et avoir la séquence nucléotidique (FASTA), (voir la **Figure-12**)

SMS Sequence Manipulation Suite:
Reverse Translate

Reverse Translate accepts a protein sequence as input and uses a codon usage table to generate a DNA sequence representing the most likely non-degenerate coding sequence. A consensus sequence derived from all the possible codons for each amino acid is also returned. Use Reverse Translate when designing PCR primers to anneal to an unsequenced coding sequence from a related species.

Paste the raw sequence or one or more FASTA sequences into the text area below. Input limit is 20,000,000 characters.

>sample sequence
Collez la sequence ADN

Submit Clear Reset Enter the codon table you wish to use (in GCG format). The default codon usage table was generated using all the E. coli coding sequences in GenBank. It was obtained from the Codon Usage Database.

AmAcid	Codon	Number	/1000	Fraction	..
Gly	GGG	50527.00	11.12	0.15	
Gly	GGA	39036.00	8.59	0.12	
Gly	GGT	114185.00	25.14	0.34	
Gly	GGC	130043.00	28.63	0.39	

*This page requires JavaScript. See [browser compatibility](#).
*You can mirror this page or use it off-line.

new window | home | citation

>sample sequence
MATVNLVLRPRARKVAKSNVPALEACPKRGVCTRVYTTTPKYPNSALRKVCRVRLTNGF
EVTYSYIGEGEHLQEHVSVLTIRGGKDLPGVRYHTVRGALDCSGVKDRKQARSKYGVKRP
KA

Submit Clear Reset Enter the codon table you wish to use (in GCG format). The default codon usage table was generated using all the E. coli coding sequences in GenBank. It was obtained from the Codon Usage Database.

AmAcid	Codon	Number	/1000	Fraction	..
Gly	GGG	50527.00	11.12	0.15	
Gly	GGA	39036.00	8.59	0.12	
Gly	GGT	114185.00	25.14	0.34	
Gly	GGC	130043.00	28.63	0.39	

*This page requires JavaScript. See [browser compatibility](#).
*You can mirror this page or use it off-line.

Figure 12 : Représentation de l’outil Sequence Manipulation Suite « Reverse Translate » et son mode d’utilisation.

II.1.4 CHOPCHOP :

CHOPCHOP est un outil Web (<http://chopchop.cbu.uib.no/>) permettant d’identifier les cibles d’ARN guide unique (ARNg) CRISPR-Cas. Ce site web est un outil web permettant de sélectionner des sites cibles pour la mutagenèse dirigée par CRISPR/Cas9, CRISPR/Cpf1, CRISPR/Cas13 ou NICKASE/TALEN (Kornel Labun *et al.*, 2019). La rapidité et la facilité de mise en œuvre de ces technologies ont conduit à une explosion d’organismes mutants et transgéniques ainsi il accepte un large éventail d’entrées (identificateurs de gènes, régions génomiques ou séquences collées) et fournit un éventail d’options avancées pour la sélection de cibles (TessaG. Montague *et al.*, 2014).

Voir (figure 13)

Figure 13 : Représentation de l'outil CHOPCHOP.

II.1.5 Primer3Plus :

Primer3plus (<http://www.bioinformatics.nl/primer3plus>) est une nouvelle interface Web pour le populaire programme de conception d'amorces Primer3 en tant qu'alternative améliorée aux scripts. Il se compose d'un programme en ligne de commande et d'une interface Web. L'interface Web est une grande forme montrant toutes les options possibles. Cela rend l'interface puissante, mais en même temps déroutante pour les utilisateurs occasionnels. Primer3Plus fournit une interface utilisateur intuitive utilisant les technologies Web actuelles et a été développé en étroite collaboration avec des biologistes moléculaires et des techniciens qui conçoivent régulièrement des amorces. Il se concentre sur la tâche à accomplir et masque les paramètres détaillés de l'utilisateur jusqu'à ce qu'ils soient nécessaires. Les paramètres et les séquences d'amorces conçues peuvent être stockés localement pour une utilisation ultérieure. Primer3Plus prend en charge une gamme de formats de séquence courants, tels que FASTA. Enfin, les amorces sélectionnées par Primer3Plus peuvent être envoyées à un formulaire de commande, permettant une intégration étroite dans les systèmes de commande du laboratoire. De plus, l'architecture ouverte de Primer3Plus permet une extension ou une intégration facile de progiciels externes.

En utilisant l'outil Primer3Plus **Voir (figure 14)**, nous avons effectué la conception d'amorces rechercher pour le gène sélectionné le gène rpsL (K43/T43) avec la configuration suivante :

Primer Size	<u>Min:</u> <input type="text" value="18"/>	<u>Opt:</u> <input type="text" value="20"/>	<u>Max:</u> <input type="text" value="27"/>	
Primer Tm	<u>Min:</u> <input type="text" value="57.0"/>	<u>Opt:</u> <input type="text" value="60.0"/>	<u>Max:</u> <input type="text" value="63.0"/>	<u>Max Tm Difference:</u> <input type="text" value="100.0"/>
Primer Bound%	<u>Min:</u> <input type="text" value="-10.0"/>	<u>Opt:</u> <input type="text" value="97.0"/>	<u>Max:</u> <input type="text" value="110.0"/>	<u>Annealing Temp:</u> <input type="text" value="52.0"/>
Primer GC%	<u>Min:</u> <input type="text" value="20.0"/>	<u>Opt:</u> <input type="text" value="50.0"/>	<u>Max:</u> <input type="text" value="80.0"/>	
<u>Concentration of monovalent cations:</u>	<input type="text" value="50.0"/>	<u>ANNEALING Oligo Concentration:</u>	<input type="text" value="50.0"/>	Not the concentration of oligos in the reaction mix!
<u>Concentration of divalent cations:</u>	<input type="text" value="1.5"/>	<u>Concentration of dNTPs:</u>	<input type="text" value="0.6"/>	
<u>DMSO Concentration:</u>	<input type="text" value="0.0"/>	<u>Formamide Concentration:</u>	<input type="text" value="0.0"/>	

DMSO Factor:

0.6

Primer3Plus
pick primers from a DNA sequence

Load server settings: Default v Activate Settings v
Task: generic v

Main General Settings Advanced Settings Internal Oligo Penalties Advanced Seq. Save Sequence

Sequence Id:

Paste template sequence below or upload sequence file: Parcours... Aucun fichier sélectionné. Load Example

Excluded Regions: < > < Excluded > [Target] [Included] Clear Regions from Seq. Save Sequence

Introns: [] [] []

Included Region: [] [] []

Primer overlap positions: -

Pair O.K. Region List

Primer3Plus
pick primers from a DNA sequence

Load server settings: Default v Activate Settings v
Task: generic v

Main General Settings Advanced Settings Internal Oligo Penalties Advanced Seq. Save Sequence

Product Size Ranges: 501-600 601-700 401-500 701-850 851-1000 1001-1500 1501-3000 3001-5000 401

Primer Size: Min: 18 Opt: 20 Max: 27
Primer Tm: Min: 57.0 Opt: 60.0 Max: 63.0 Max Tm Difference: 100.0
Primer Bound%: Min: -10.0 Opt: 97.0 Max: 110.0 Annealing Temp: 52.0
Primer GC%: Min: 20.0 Opt: 50.0 Max: 80.0

Concentration of monovalent cations: 50.0 ANNEALING Oligo Concentration: 50.0 Not the concentration of oligos in the reaction mix!
Concentration of divalent cations: 1.5 Concentration of DNTPs: 0.6
DMSO Concentration: 0.0 Formamide Concentration: 0.0
DMSO Factor: 0.6

Mispriming/Repeat Library: NONE v

Load and Save
To upload or save a settings file from your local computer, choose here:
Load only Settings from file: Parcours... Aucun fichier sélectionné.
Save only Settings Save Everything Save as Primer3 Input Save Primer3Plus results

Mark selected region: < Excluded > [Target] [Included] Clear Regions from Seq. Save Sequence

Collez la sequence ADN

ATGGCGACGTGACCAAGCTGTGGCAAAACCGCGCGCAAAAGTGGCGAAAGCAAGCTGGCGGCGCTGGAAAGCTGCCCGAGAAA
CGCGGGGTGTGACCGGTGTATACCAACCCGGAAAAACCGAAACAGCGGCTGGCGAAAGTGTGGCGCTGGCAACAGGC
TTTGAGTGCACCAAGCTATGGCGCGAAAGGCAATACCTGCAAGGACATAGCGTATTCTTGTTCGCCGCGCGCGCGCGCAAGATCTG
CCGGAGTGTGCTATCTATACGTGGCGCGCGCGCTGGATTGCAGCGGGGTGAAAGATTCGCAACAGCGCGCGCAAGATATGGCGTAAA
CCCCGAAAGCG

Figure 14 : Représentation de l'outil Primer3Plus.

II.1.6 DrugBank :

DrugBank Online (<https://go.drugbank.com>) est une base de données en ligne complète et gratuite contenant des informations sur les médicaments et les cibles médicamenteuses. En tant que ressource bioinformatique et chimioinformatique, nous combinons des données détaillées sur les médicaments (c'est-à-dire chimiques, pharmacologiques et pharmaceutiques) avec des informations complètes sur les cibles médicamenteuses (c'est-à-dire la séquence, la structure et la voie). DrugBank Online est largement utilisé par l'industrie pharmaceutique, les chimistes médicaux, les pharmaciens, les médecins, les étudiants et le grand public. En raison de sa vaste portée, de ses références complètes et de ses descriptions détaillées des données, DrugBank permet des avancées majeures dans l'industrie de la médecine basée sur les données.

Nous avons recherché à l'aide de l'outil DrugBank l'interaction de la streptomycine avec notre gène rpsL modifié (K43/T43), avec la configuration suivante (**voir la Figure15**) :

Drug Types (default all):

- | | | | |
|---|---|---|---|
| <input checked="" type="checkbox"/> Approved | <input checked="" type="checkbox"/> Vet approved | <input checked="" type="checkbox"/> Nutraceutical | <input checked="" type="checkbox"/> Illicit |
| <input checked="" type="checkbox"/> Withdrawn | <input checked="" type="checkbox"/> Investigational | <input checked="" type="checkbox"/> Experimental | |

Protein Types (default all):

- | | | | |
|--|--|---|---|
| <input checked="" type="checkbox"/> Target | <input checked="" type="checkbox"/> Enzyme | <input checked="" type="checkbox"/> Carrier | <input checked="" type="checkbox"/> Transporter |
|--|--|---|---|

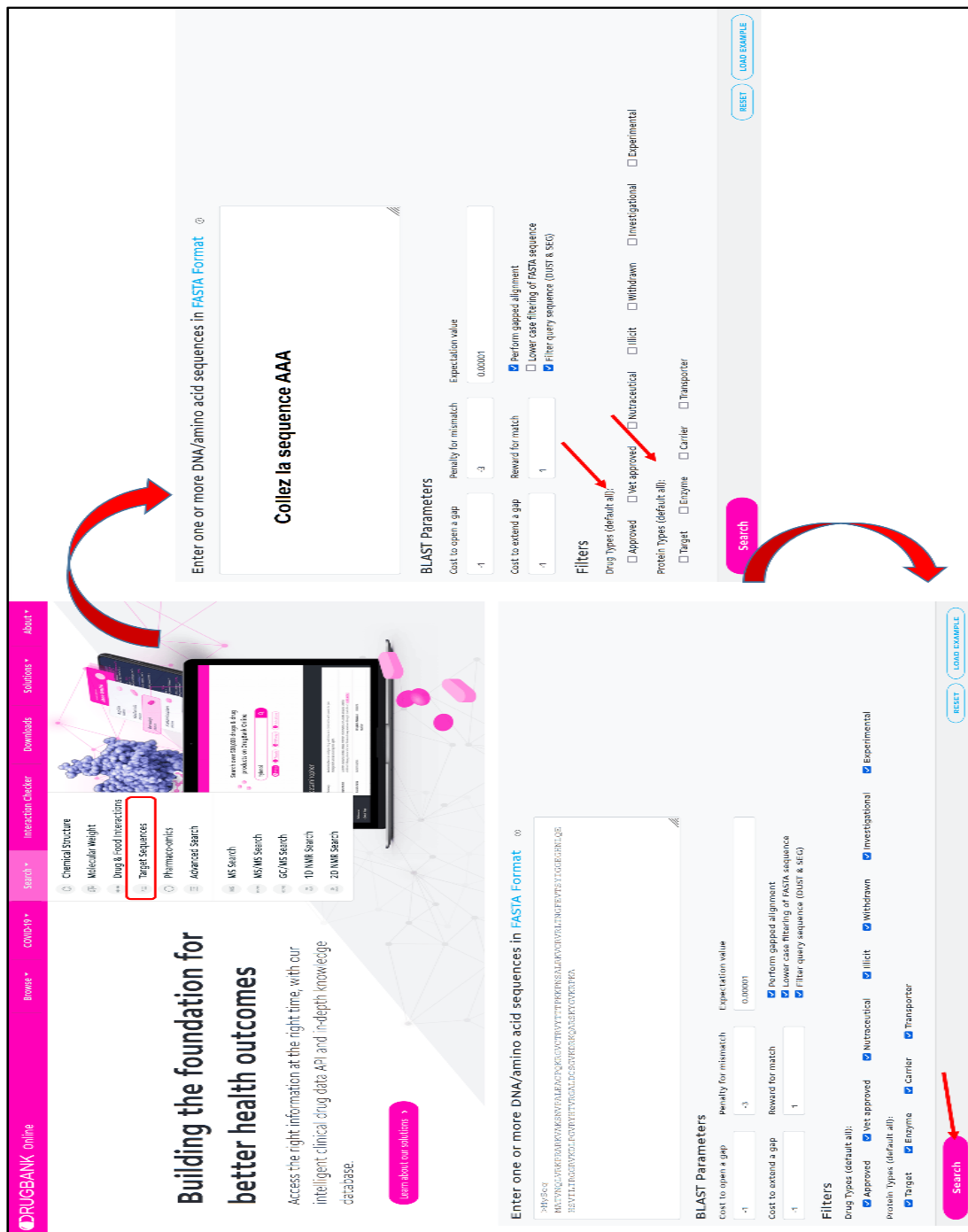


Figure 15 : Représentation de l'outil DrugBank.

Chapitre III : Résultats et Discussions

III.1 Résultats des outils NCBI Gene, UniProt et Reverse Translate :

D'après les résultats de l'outil NCBI Gene et Uniprot, le gène rpsL code pour la protéine RS12 de la sous unité ribosomale 30S avec une longueur de 124 Acides Aminés et un poids moléculaire = 13.73 Da

>sp|P0A7S3|RS12_ECOLI 30S ribosomal protein S12 OS=Escherichia coli (strain K12)

OX=83333 GN=rpsL PE=1 SV=2

MATVNQLVVRKPRARKVAKSNVPALEACPQKRGVCTRVTYTTTPKKPNSALRKVCRV
RLTNGFEVTSYIGGEGHNLQEHSVILIRGGRVKDLPGVRYHTVRGALDCSGVKDRKQ
ARSKYGVKRPKA

Celle-ci avec les sous unités ribosomales S4 et S5 joue un rôle important dans la précision translationnelle. Il interagit avec et stabilise les bases de l'ARNr 16S impliquées dans la sélection de l'ARNt dans le site A et avec le squelette de l'ARNm. Situé à l'interface des sous-unités 30S et 50S, il traverse le corps de la sous-unité 30S en contactant les protéines de l'autre côté et en maintenant probablement la structure de l'ARNr ensemble. Le cluster combiné de protéines S8, S12 et S17 semble maintenir ensemble l'épaule et la plate-forme de la sous-unité 30S (par similarité) (Urlaub H *et al.*, 1995).

La recherche de la séquence codante pour la protéine RS12 de la sous unité ribosomale 30S par l'outil SMS « Reverse Translate » a permis de retrouver une séquence d'une longueur de 372 nucléotides et d'un poids 11.53 kDa, avec le premier codant **ATG** et dernier **GCG** (voir **figure 16**)

Figure 17 : Représentation des différents ARNg ciblant la sous unité ribosomale 30S du gène rpsL.

L'ARNg ciblant la position K43 de la séquence du gène rpsL est le Rank 19 (Figure 18) (voir Annexe I), cette ARNg a cinq séquences que peuvent être utiliser pour la conception de la mutation K43/T43 sont représenté dans le *Tableau I*

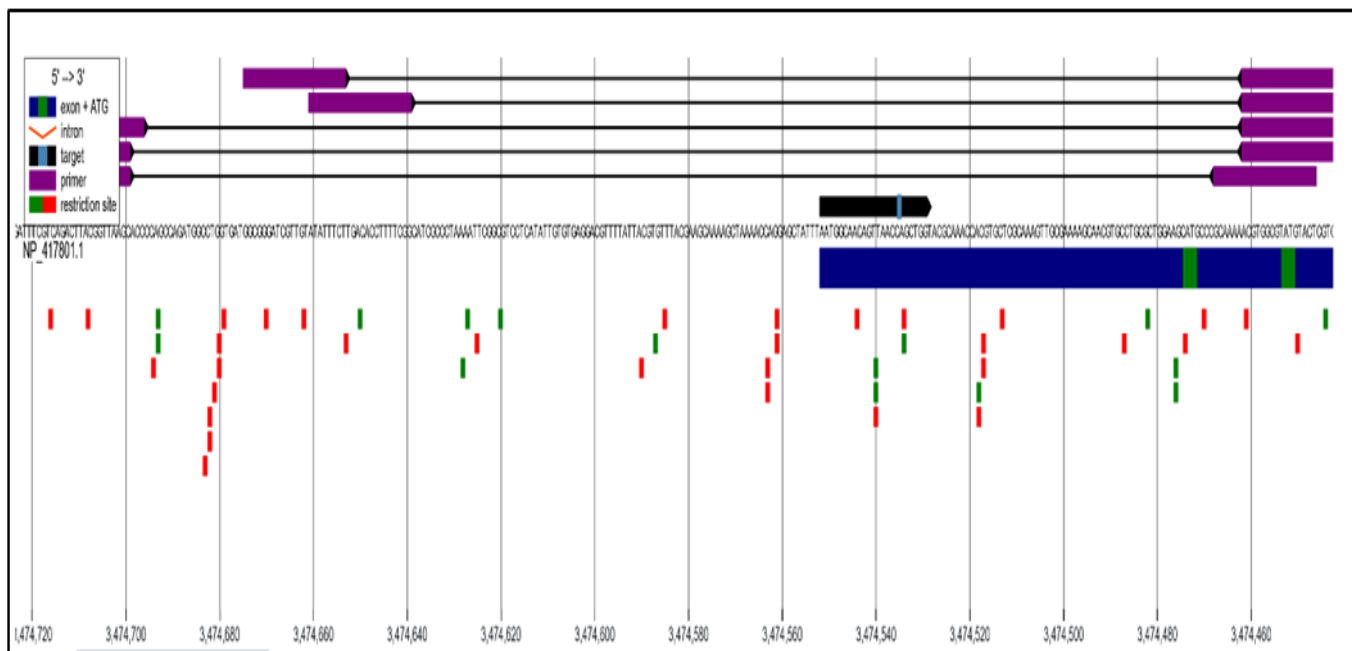


Figure 18 : Représentation des ARNg ciblant la mutation K43/T43 de la sous unité ribosomale 30S du gène rpsL.

Tableau I : Les séquences d'ARNg correspondant pour la mutation K43/T43

Pair	Left primer coordinates	Left primer	Right primer coordinates	Right primer	Product size
1	NC_000913.3 :3474700- 3474722	TCGTCAGACTTAC GGTTAAGCA	NC_000913.3 :3474447- 3474469	GAGTACATACGCC ACGTTTTTG	275
2	NC_000913.3 :3474700- 3474722	TCGTCAGACTTAC GGTTAAGCA	NC_000913.3 :3474441- 3474463	ATACACGAGTAC ATACGCCACG	281
3	NC_000913.3 :3474697- 3474719	TCAGACTTACGGT TAAGCACCC	NC_000913.3 :3474441- 3474463	ATACACGAGTAC ATACGCCACG	278

4	NC_000913.3 :3474640- 3474662	TATTTCTTGACAC CTTTTCGGC	NC_000913.3 :3474441- 3474463	ATACACGAGTAC ATACGCCACG	221
5	NC_000913.3 :3474654- 3474676	GCGGGATCGTTG TATATTTCTT	NC_000913.3 :3474441- 3474463	ATACACGAGTAC ATACGCCACG	235

Les prédictions du profil de réparation de l'ARNg sont calculées selon **Shen MW et al., 2018** qui spécule qu'après le clivage Cas9, la réparation de l'ADN sans matrice donneuse est généralement considérée comme stochastique, hétérogène et peu pratique au-delà de la perturbation génique, d'après les résultats décrits retrouver l'ARNg numéro 1 avec une taille de séquence 275 est la plus significative pour la conception de la mutation K43/T43 :

- Reference sequence :

TATTACGTGTTTACGAAGCAAAAGCTAAAACCAGGAGCTATTTAATGGCAACA
GTTAACCGAGCTGGTACGCAAACCACGTGCTCGCAAAGTTGCGAAAAGCAACGT
GCCTGCGCTGGAAG

- Precision score : 0.43
- Frame +0 frequency : 23.84
- Frame +1 frequency : 42.97
- Frame +2 frequency : 33.19
- 1-bp ins frequency : 17.35
- Highest Del frequency : 7.24
- Highest Ins frequency : 14.20
- Highest Outcome frequency : 14.20
- Microhomology deletion frequency : 58.26
- Microhomology-less deletion frequency : 24.39
- Left primer coordinates : **NC_000913.3:3474700-3474722**
Left primer: **TCGTCAGACTTACGGTTAAGCA**
- Right primer coordinates : **NC_000913.3:3474447-3474469**
Right primer: **GAGTACATACGCCACGTTTTTG**

La configuration et le Protocol PCR pour concevoir la séquence cible d'ADN Cas9 K43/T43 est décrite dans le **Tableau II.** comme suivant :

Tableau II : Protocol PCR pour concevoir la séquence cible d'ADN Cas9 K43/T43

Tableau III : Les quatre primer qui sont en mesure de couvrir la région ou notre substitution est situé.

Pair	Left primer	Start	Lenght	Product	Right rimer	Start	Lenght
1	CGAAAAGCAACGTGCC GG	50	18		CATATTTGCTGCGCGCCTG	352	19
2	ATGGCGACCGTGAACCA G	1	18		ATCTTTCACGCCGCTGCAA T	273	20
3	CGACCGTGAACCAGCTG G	5	18		GCCTGTTTGCGATCTTTCA CG	238	21
4	CCAGCTGGTGCGCAAAC C	15	18		TATTTGCTGCGCGCCTGTT T	350	20

Selon les résultats retrouver à partir de l'outil Primer3Plus le deuxième primer est le plus significatif et le plus long (**Figure 20**), ce qui permet une bonne couverture du gène rpsL lors la PCR:

▪ **Left Primer 2 :** CGAAAAGCAACGTGCCGG

Start: **Length:** **Tm:** **B:** **GC:** **Any:** **End:** **TB:** **HP:** **3' Stab:** **Penalty:**
 50 18 bp 60.4 C 94.810 % 61.1 % 0.0 0.0 7.0 46.5 6.1 2.427

▪ **Right Primer 2:** CATATTTGCTGCGCGCCTG

Start: **Length:** **Tm:** **B:** **GC:** **Any:** **End:** **TB:** **HP:** **3' Stab:** **Penalty:**
 352 19 bp 60.7 C 95.750 % 57.9 % 30.8 22.1 8.0 38.0 4.8 1.656

Left Primer 2:

Start: 50 **Length:** 18 bp **Tm:** 60.4 C **B:** 94.810 % **GC:** 61.1 % **Any:** 0.0 **End:** 0.0 **TB:** 7.0 **HP:** 46.5 **3' Stab:** 6.1 **Penalty:** 2.427

Hairpin Secondary Structure:

Tm: 46.5°C dG: -560 cal/mol dH: -18800 cal/mol dS: -59 cal/mol*K
 5' CGAAAAGCAA₇
 ||| C
 3' GGCCGTG₁

Right Primer 2:

Start: 352 **Length:** 19 bp **Tm:** 60.7 C **B:** 95.750 % **GC:** 57.9 % **Any:** 30.8 **End:** 22.1 **TB:** 8.0 **HP:** 38.0 **3' Stab:** 4.8 **Penalty:** 1.656

Self Any Secondary Structure:

Tm: 30.8°C dG: -9874 cal/mol dH: -65400 cal/mol dS: -179 cal/mol*K
 5' CATATTTGCTGCGCGCCTG 3'
 |||||
 3' GTCCGCGCTCGTTTATAC 5'

1	ATGGCGACCG	TGAACCAGCT	GGTGCGCAAA	CCGCGCGCGC	GCAAAGTGGC
51	GAAAAGCAAC	GTGCCGGCGC	TGGAAGCGTG	CCCGCAGAAA	CGCGGCGTGT
101	GCACCCGCGT	GTATACCACC	ACCCCGAAAA	CACCGAACAG	CGCGCTGCGC
151	AAAGTGTGCC	GCGTGCGCCT	GACCAACGGC	TTTGAAGTGA	CCAGCTATAT
201	TGGCGGCGAA	GGCCATAACC	TGCAGGAACA	TAGCGTGATT	CTGATTCGCG
251	GCGGCCGCGT	GAAAGATCTG	CCGGGCGTGC	GCTATCATAC	CGTGCGCGGC
301	GCGCTGGATT	GCAGCGGCGT	GAAAGATCGC	AAACAGGCGC	GCAGCAAATA
351	TGGCGTGAAA	CGCCCGAAAG	CG		

Figure 20: Représentation des amorces PCR Primer3Plus de sous unité ribosomal 30S codé par le gène rpsL après substitution K43/T43.

D'après la comparaison entre le sous-unité ribosomal 30S codé par le gène rpsL et celle édité par le système enzymatique Cas9 au niveau de l'acide Aminé « K 43 » vers l'acide Aminé « T

43 » par l’outil GeneBank, douze antibiotiques sont retrouver *Tableau IV* qui ont différentes actions sur la bactérie E.coli K12.

Tableau IV : Les antibiotiques qui ont différentes actions sur la bactérie E.coli K12.

N°	Drug	Drug Groupe	Pharmacological Action	Action
1	Tigecycline	approved	yes	binder
2	Streptomycin	approved, vet_approved	unknown	
3	Spectinomycin	approved, investigational, vet_approved	yes	inhibitor
4	Ribostamycin	approved, experimental	unknown	-
5	Netilmicin	approved, investigational	yes	inhibitor
6	Neomycin	approved, vet_approved	yes	inhibitor
7	Kanamycin	approved, investigational, vet_approved	yes	inhibitor
8	Gentamicin C1a	experimental	yes	adduct
9	Gentamicin	approved, vet_approved	yes	binder
10	Framycetin	approved	yes	inhibitor
11	Arbekacin	experimental, investigational	yes	inhibitor
12	Amikacin	approved, investigational, vet_approved	yes	inhibitor

III.3 Discussions des résultats :

Selon ces résultat nous avons pu constater que la bactérie E.coli.K12 est devenu résistante à la Streptomycine qui est dans le cas normal sensible car cette antibiotique à une action inhibitrice sur la sous-unité ribosomal 30S .

La streptomycine cible le ribosome pour la liaison et l'inactivation, entraînant une perte de synthèse des protéines. La mutation K43T dans la protéine rpsL rend le ribosome résistant à la streptomycine. Par conséquent, pour générer des E. coli résistants à la streptomycine, nous

devrons introduire la mutation de l'ADN correspondant à la mutation de la protéine K43T dans le génome d'E. coli. Comme indiqué ci-dessus, pour introduire des modifications spécifiques dans le génome, le DSB généré par Cas9 doit être réparé par Homology Directed Repair (HDR). Ce processus est très inefficace et nécessite que la mutation à introduire soit à moins de 30 pb du site de DSB.

Ainsi, dans ce cas, le choix du sgRNA dépend de l'emplacement et il pourrait ne pas être possible de contrôler la génération de DSB hors cible.

Pour réparer ce DSB et introduire la mutation <A> 128 <C> dans le génome (correspondant au « K 43 » / « T43 »), une matrice d'ADN simple brin portant la séquence mutante est introduite à des concentrations élevées afin qu'elle puisse être préférentiellement utilisée comme matrice donneuse pour le HDR.

La matrice donneuse porte la séquence de mutation A128C modifiée et l'action combinée de Cas9, sgRNA spécifique de rpsL, donne un DSB au niveau du gène rpsL d'E. coli qui peut être réparé par réparation dirigée par homologie à l'aide de la matrice donneuse. Après transformation avec les plasmides, il y a une longue période d'incubation (4-12 h) où le sgRNA et Cas9 sont exprimés et fonctionnels. L'incubation est effectuée à 30°C pour permettre le maintien du plasmide pCas qui a une origine de réplication sensible à la température. Pendant cette période, Cas9 dirige le clivage de l'ADN au niveau du gène rpsL.

Selon **Miskinyte M et al., 2013** les mutations ponctuelles du gène rpsL ont procuré un avantage de survie chez E. coli commensal dans la niche intracellulaire des macrophages, ce qui a augmenté le risque d'échec du traitement pendant un processus d'infection par la Streptomycine.

En inhibant la synthèse polypeptidique, la streptomycine déclenche des erreurs de lecture lors de la traduction génomique qui conduisent finalement à la mort des cellules bactériennes (**Ruusala T et al., 1984**).

Plusieurs mutations d'un seul nucléotide du gène rpsL, qui code pour la protéine ribosomique S12, ont été caractérisées comme conférant une résistance à la Streptomycine chez différentes bactéries (**Gregory et al., 2001**).

Selon **Jiang et al., 2013** cet éventail de modifications fait du marqueur antibiotique un outil robuste pour la mutagenèse génomique chez E. coli.

Compte tenu des taux d'occurrence négligeables de mutants spontanés observés chez E.coli, le choix de la résistance à la streptomycine conférée par la mutagenèse oligo-médiée au gène rpsL de E.Coli, l'expérience que nous avons menée in-Silico par la technologie Crisper Cas9

Résultats et discussions

dirigé par rpsL dont l'invasion de brin entraînerait la mutation faux-sens rpsL « K43 » / « T43 », nous avons d'abord noté que cette mutation est sensible pour Tigecycline et la Framycetin approuvée avec une action pharmacologique binding et inhibitrice respectivement, alors l'action de la Streptomycine reste inconnu sur la bactérie avec une mutation du gène rpsL « K43 » / « T43 ».

Conclusion

Conclusion

Le système CRISPR-Cas9 est un système d'édition des génomes. Il a été identifié comme un système d'immunité adaptative chez les bactéries contre les virus et les plasmides.

Il existe différents types de système CRISPR : Le type I, le type II, le type III.

Le système CRISPR de type II est composé de trois composants : le tracrRNA, le crRNA ces deux forment l'ARNsg, et l'endonucléase Cas9. L'exigence de fonctionnement de ce système au site cible est la présence d'une séquence PAM conservée de 2 à 5 nucléotides.

Dans notre étude, nous avons pu constater que la bactérie E.Coli K12, grâce à la mutation faite, est devenu une bactérie résistante à la Streptomycine, qui dans le cas normal, cet antibiotique se fixe sur la sous-unité ribosomale 30S et inhiber le fonctionnement de ribosome et donc empêcher la bactérie de se répliquer et de se développer.

Références bibliographiques

A

- Al-Wawi MZ, Hassan RM, Mohamed ME, Khan MF, Magaogao M et Hossain A : édition du génome médiée par CRISPR-Cas9 dans Escherichia coli. Int J Pharm Sci & Res 2019 ; 10(7): 3373-77. doi : 10.13040/IJPSR.0975-8232.10(7).3373-77.
- Anders C, Niewoehner O, Duerst A, Jinek M. 2014. Base structurelle de la reconnaissance de l'ADN cible dépendant de PAM par l'endonucléase Cas9.Nature513(7519):569–73.
- AC Komor , AH Badran , DR Liu : Technologies basées sur CRISPR pour la manipulation des génomes eucaryotes Cellule , 168 (2017) , p. 20 - 36

B

- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P.2007. CRISPR fournit une résistance acquise contre les virus chez les procaryotes.La science315 :1709–1712.
- Baohong Zhang (2020). Thérapie génique CRISPR/Cas , DOI : 10.1002/jcp.30064

C

- Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, Sternberg SH, Joung JK, Yildiz A, Doudna JA. Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. Nature. 2017;550(7676):407–410.
- Chen H, Choi J, Bailey S. 2014. Couper la sélection du site par les deux domaines de nucléase de l'endonucléase guidée par l'ARN Cas9.J. Biol. Chim.289(19):13284–94

D

- Doudna, JA (2020). La promesse et le défi du génome thérapeutique édition.Nature, 578,229–236.
- Doudna JA, Charpentier E. 2014. La nouvelle frontière de l'ingénierie du génome avec CRISPR-Cas9.La science 346(6213):1258096

G

- Ghaemi Asma, Elnaz Bagheri, Khalil Abnous, Seyed Mohammad Taghdisib, Mohamed Ramezani, Mona Alibolandi, Sciences de la vie 267 (2021) 118969.
- Gregory, S.T. , Cate, J.H.D. , and Dahlberg, A.E. (2001) Streptomycin-resistant and streptomycin-dependent mutants of the extreme thermophile Thermus thermophilus . J Mol Biol 309: 333–338.

- GJ Knott , JA Doudna : CRISPR-Cas guide l'avenir du génie génétique Sciences , 361 (2018) , p .866-869
- Giuseppina La Sala, Anders Gunnarsson, Karl Edman, Christian Tyrchan, Anders Hogner, Andrey I. Frolov . Démêler le dialogue allostérique entre le peptide coactivateur et le site de liaison au ligand dans le récepteur des glucocorticoïdes. *Journal of Chemical Information and Modeling* **2021**, 61 (7) , 3667-3680. <https://doi.org/10.1021/acs.jcim.1c00323>.

H

- Horvath P et Rodolphe Barrangou *La science*327, 167 (2010).
- H. Mou , Z. Kennedy , DG Anderson , H. Yin , W. Xue : Modèles de souris cancéreuses de précision grâce à l'édition du génome avec CRISPR-Cas9 *Génome Med .* , 7 (2015) , p. 53.

I

- Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A.1987. Séquence nucléotidique de laiapgène, responsable de la conversion des isozymes de la phosphatase alcaline dans *Escherichia coli*, et l'identification du produit génique. *Bactériol J*169 :5429–5433.

J

- Jiang F, Doudna JA. CRISPR–Cas9 Structures and Mechanisms. *Annual Review of Biophysics* 2017; 46: 505–529.
- Jiang, W. , Bikard, D. , Cox, D. , Zhang, F. , and Marraffini, L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31: 233–239.
- Jimin Wang, Erin Skeens, Pablo R. Arantes, Federica Maschietto, Brandon Allen, Gregory W. Kyro, George P. Lisi, Giulia Palermo, Victor S. Batista . Base structurale de la dynamique réduite de trois mutants Lys-à-Ala de l'endonucléase HNH pour l'enzyme CRISPR/Cas9 (CRISPR/Cas9) associée à de courtes répétitions palindromiques groupées régulièrement espacées. *Biochimie* **2022**, 61 (9) , 785-794. <https://doi.org/10.1021/acs.biochem.2c00127>
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. Une endonucléase d'ADN guidée par double ARN programmable dans l'immunité bactérienne adaptative. *La science*337(6096):816–21.
- Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, et al. 2014. Les structures des endonucléases Cas9 révèlent l'activation conformationnelle médiée par l'ARN. *La science*343(6176):1247997.
- Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. 2013. Édition du génome programmée par ARN dans les cellules humaines. *eVie*2:e00471.

- Jacques P. Tremblay (2015). CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. *médecine/sciences*, 31(11), 1014–1022. doi:10.1051/medsci/20153111016 .
- Jennifer A. Doudna et Emmanuelle Charpentier *La science*346, (2014); DOI : 10.1126 / science.1258096

L

- Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. *Nucleic Acids Res.* 2019 Jul 2;47(W1):W171-W174. doi: 10.1093/nar/gkz365.PMID: 31106371

M

- Miskinyte M, Gordo I. Increased survival of antibiotic-resistant *Escherichia coli* inside macrophages. *Antimicrob Agents Chemother.* 2013 Jan;57(1):189-95. doi: 10.1128/AAC.01632-12. Epub 2012 Oct 22. PMID: 23089747; PMCID: PMC3535928.
- Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, et al. 2015. Une classification évolutive mise à jour des systèmes CRISPR-Cas. *Nat. Rév. Microbiol.*13(11):722–36.
- Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. 2009. De courtes séquences de motifs déterminent les cibles du système de défense procaryote CRISPR. *Microbiologie*155(3):733–40.
- Makarova KS, Wolf YI, Iranzo, J., Shmakov, SA, Alkhnbashi, OS, Brouns, SJJ, Charpentier, E., Cheng, D., Haft, DH, Horvath, P., Moineau, S., Mojica, FJM, Scott, D., Shah, SA, Siksny, V., Terns, MP, Venclovas, Ä., White, MF, Yakunin, AF,... Koonin, EV (2020). Classification évolutive des systèmes CRISPR-Cas : une salve de classe 2 et de variantes dérivées. *Nature Reviews Microbiologie*, 18 ans,67–83.

N

- Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, et al. 2014. Structure cristalline de Cas9 en complexe avec ARN guide et ADN cible. *Cellule*156(5):935–49.

O

O. Shalem , NE Sanjana , F. Zhang : Génomique fonctionnelle à haut débit utilisant CRISPR-Cas9 *Nat. Rév. Genet.* , 16 (2015) , p. 299 – 311

P

- Palermo G, Miao Y, Walker RC, Jinek M, McCammon JA. 2016. Plasticité frappante de CRISPR-Cas9 et rôle clé de l'ADN non cible, révélé par des simulations moléculaires. *ACS Cent. Sci.*2(10):756–63
- PD Hsu , ES Lander , F. Zhang : Développement et applications de CRISPR-Cas9 pour l'ingénierie du génome *Cellule* , 157 (2014) , p. 1262 - 1278

R

- Ruusala, T., and Kurland, C.G. (1984) Streptomycin preferentially perturbs ribosomal proofreading. *Mol Gen Genet* 198: 100–104.

S

- Shen MW, Arbab M, Hsu JY, Worstell D, Culbertson SJ, Krabbe O, Cassa CA, Liu DR, Gifford DK, Sherwood RI. Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*. 2018 Nov;563(7733):646-651. doi: 10.1038/s41586-018-0686-x. Epub 2018 Nov 7. Erratum in: *Nature*. 2019 Mar;567(7746):E1-E2. PMID: 30405244; PMCID: PMC6517069.
- Sternberg SH, LaFrance B, Kaplan M, Doudna JA. 2015. Contrôle conformationnel du clivage cible de l'ADN par CRISPR–Cas9. *Nature* 527(7576):110–13.
- Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. 2014. Interrogation de l'ADN par l'endonucléase Cas9 guidée par l'ARN CRISPR. *Nature* 507(7490):62–67.

T

- T. Wang , JJ Wei , DM Sabatini , ES Lander : Screenings génétiques dans les cellules humaines à l'aide du système CRISPR-Cas9 *Sciences* , 343 (2014) , p. 80 - 84
- Tessa G. Montague, José M. Cruz, James A. Gagnon, George M. Church, Eivind Valen *Nucleic Acids Research*, Volume 42, Issue W1, 1 July 2014, Pages W401–W407, <https://doi.org/10.1093/nar/gku410>

U

- Urlaub H , V Krufft, O Bischof, E C Müller, B Wittmann-Liebold Protein-rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies; *EMBO J*. 1995 Sep 15;14(18):4578-88. doi: 10.1002/j.1460-2075.1995.tb00137.x.

Y

- Yang W. 2008. Un ion métallique équivalent dans la catalyse à un et deux ions métalliques. *Nat. Structure. Mol. Biol.* 15(11):1228–31.

Liste des Annexes

Amorces PCR Primer3Plus de sous unité ribosomal 30S codé par le gène rpsL après substitution K43/T43

Pair 1:

Left Primer 1:

Start: 3 Length: 20 bp Tm: 59.4 C B: 93.719 % GC: 55.0 % Any: 31.6 End: 31.6 TB: 9.0 HP: 37.3 3' Stab: 4.8 Penalty: 0.598

Self Any Secondary Structure:

Tm: 31.6°C dG: -9961 cal/mol dH: -70800 cal/mol dS: -196 cal/mol*K

```
5' GGCAACAGTTAACCAGCTGG 3'
      |||||
3' GGTCGACCAATTGACAACGG 5'
```

Self End Secondary Structure:

Tm: 31.6°C dG: -9961 cal/mol dH: -70800 cal/mol dS: -196 cal/mol*K

```
5' GGCAACAGTTAACCAGCTGG 3'
      |||||
3' GGTCGACCAATTGACAACGG 5'
```

Hairpin Secondary Structure:

Tm: 37.3°C dG: -27 cal/mol dH: -26600 cal/mol dS: -86 cal/mol*K

```
5' GGCAACAGTTAA}
      ||| |
3' GGTCGACC }
```

Right Primer 1:

Start: 354 Length: 20 bp Tm: 60.1 C B: 94.929 % GC: 55.0 % Any: 0.0 End: 0.0 TB: 7.0 HP: 46.9 3' Stab: 4.6 Penalty: 0.108

Hairpin Secondary Structure:

Tm: 46.9°C dG: -651 cal/mol dH: -21000 cal/mol dS: -66 cal/mol*K

```
5' GCCATA}
      |||
3' TCCGAGCAAGGTTC }
```

Pair: Product Size: 352 bp Tm: 88.9 C Any: 0.0 End: 0.0 TB: 14.0 Penalty: 0.706

First Primer Pair Previous Primer Pair **Next Primer Pair** Last Primer Pair

1	ATGGCAACAG	TTAACCAGCT	GGTACGCAAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCGCAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTT	GTTACCACAC	CGTACGTGGT
301	GCGCTTGA	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 2:

 Left Primer 2:

Start: 27 Length: 20 bp Tm: 60.9 C B: 96.695 % GC: 55.0 % Any: 10.9 End: 0.0 TB: 8.0 HP: 41.2 3' Stab: 2.8 Penalty: 0.930

[Self Any Secondary Structure:](#)

Tm: 10.9°C dG: -6310 cal/mol dH: -53400 cal/mol dS: -152 cal/mol*K
 5' CAAACCACGTGCTCGCAAAG 3'
 |||||
 3' GAAACGCTCGTCACCAAAC 5'

[Hairpin Secondary Structure:](#)

Tm: 41.2°C dG: -282 cal/mol dH: -21200 cal/mol dS: -67 cal/mol*K
 5' CAAACCACGTG
 ||
 3' GAAACGCTC

 Right Primer 2:

Start: 341 Length: 20 bp Tm: 60.4 C B: 95.605 % GC: 55.0 % Any: 0.0 End: 0.0 TB: 7.0 HP: 0.0 3' Stab: 2.5 Penalty: 0.390

Pair: [Product Size:](#) 315 bp [Tm:](#) 88.8 C [Any:](#) 10.3 [End:](#) 16.5 [TB:](#) 14.0 [Penalty:](#) 1.320

[Compl Any Secondary Structure:](#)

Tm: 10.3°C dG: -5235 cal/mol dH: -63400 cal/mol dS: -188 cal/mol*K
 5' CAAACCACGT-----GCTCGCAAAG 3'
 ||| |||||
 3' TTTCTGGCATTGTCGCGAGC 5'

[Compl End Secondary Structure:](#)

Tm: 16.5°C dG: -6000 cal/mol dH: -73600 cal/mol dS: -218 cal/mol*K
 5' AAAGACCGTAAGCAG--GCTCG 3'
 || | |||||
 3' GTTTGGTGACGAGCGTTTC 5'

1	ATGGCAACAG	TTAACCAGCT	GGTACGCAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCGCAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTC
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTT	GTTACCACAC	CGTACGTGGT
301	GCGCTTGACT	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 3:
Left Primer 3:

Start: 72 Length: 19 bp Tm: 60.0 C B: 94.189 % GC: 52.6 % Any: 15.7 End: 0.0 TB: 6.0 HP: 0.0 3' Stab: 1.9 Penalty: 1.005

Self Any Secondary Structure:

Tm: 15.7°C dG: -7063 cal/mol dH: -56200 cal/mol dS: -158 cal/mol*K
 5' GGAAGCATGCCCGCAAAAA 3'
 |||||
 3' AAAAAAGCCCGTACGAAGG 5'

Right Primer 3:

Start: 373 Length: 20 bp Tm: 60.4 C B: 95.605 % GC: 55.0 % Any: 0.7 End: 0.0 TB: 6.0 HP: 41.4 3' Stab: 4.3 Penalty: 0.390

Self Any Secondary Structure:

Tm: 0.7°C dG: -5196 cal/mol dH: -45400 cal/mol dS: -130 cal/mol*K
 5' AAGCCTTAGGACGCTTCACG 3'
 ||| |||
 3' GCACTTCGCAGGATTCGGAA 5'

Hairpin Secondary Structure:

Tm: 41.4°C dG: -295 cal/mol dH: -21000 cal/mol dS: -67 cal/mol*K
 5' AAGCCTTA
 |||||
 3' GCACTTCGCAGG

Pair: Product Size: 302 bp Tm: 88.5 C Any: 6.7 End: 0.0 TB: 11.0 Penalty: 1.395

Compl Any Secondary Structure:

Tm: 6.7°C dG: -5448 cal/mol dH: -53200 cal/mol dS: -154 cal/mol*K
 5' GGAAGCATGCCCGCAAAAA 3'
 |||||
 3' GCACTTCGCAGGATTCGGAA 5'


```

1  ATGGCAACAG  TTAACCAGCT  GGTACGCAAA  CCACGTGCTC  GCAAAGTTGC
51  GAAAAGCAAC  GTGCCTGCGC  TGAAGCATG  CCCGCAAAAA  CGTGGCGTAT
101  GTACTCGTGT  ATATACTACC  ACTCCTAAAA  AACCGAACTC  CGCGCTGCGT
151  AAAGTATGCC  GTGTTCGTCT  GACTAACGGT  TTCGAAGTGA  CTTCCTACAT
201  CGGTGGTGAA  GGTACAACC  TGCAGGAGCA  CTCCGTGATC  CTGATCCGTG
251  GCGGTCGTGT  TAAAGACCTC  CCGGGTGTTT  GTTACCACAC  CGTACGTGGT
301  GCGCTTGACT  GCTCCGGCGT  TAAAGACCGT  AAGCAGGCTC  GTTCCAAGTA
351  TGGCGTGAAG  CGTCCTAAGG  CTTAA

```


Pair 4:

 Left Primer 4:

[Start](#): 22 [Length](#): 19 bp [Tm](#): 59.8 C [B](#): 94.326 % [GC](#): 57.9 % [Any](#): 10.9 [End](#): 0.0 [TB](#): 7.0 [HP](#): 44.8 [3' Stab](#): 4.3 [Penalty](#): 1.203

[Self Any Secondary Structure:](#)

Tm: 10.9°C dG: -6310 cal/mol dH: -53400 cal/mol dS: -152 cal/mol*K
 5' GTACGCAAACCCACGTGCTC 3'
 |||||
 3' CTCGTGCACCAAACGCATG 5'

[Hairpin Secondary Structure:](#)

Tm: 44.8°C dG: -586 cal/mol dH: -23900 cal/mol dS: -75 cal/mol*K
 5' GTACGCAA
 ||| A
 3' CTCGTGCACC

 Right Primer 4:

[Start](#): 324 [Length](#): 20 bp [Tm](#): 60.3 C [B](#): 95.270 % [GC](#): 50.0 % [Any](#): 5.0 [End](#): 0.0 [TB](#): 6.0 [HP](#): 39.8 [3' Stab](#): 3.2 [Penalty](#): 0.250

[Self Any Secondary Structure:](#)

Tm: 5.0°C dG: -6529 cal/mol dH: -40800 cal/mol dS: -110 cal/mol*K
 5' TTTAACGCCGGAGCAGTCAA 3'
 ||||
 3' AACTGACGAGGCCGCAATTT 5'

[Hairpin Secondary Structure:](#)

Tm: 39.8°C dG: -137 cal/mol dH: -15100 cal/mol dS: -48 cal/mol*K
 5' TTTAACGCCG
 ||
 3' AACTGACGAG

Pair: [Product Size](#): 303 bp [Tm](#): 88.5 C [Any](#): 0.0 [End](#): 0.0 [TB](#): 13.0 [Penalty](#): 1.453

1	ATGGCAACAG	TTAACCAGCT	GGTACGAAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCGCAAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCTGCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTT	GTTACCACAC	CGTACGTGGT
301	GCGCTTGACT	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 5:

 Left Primer 5:

Start: 31 Length: 19 bp Tm: 60.4 C B: 95.329 % GC: 57.9 % Any: 10.9 End: 0.0 TB: 8.0 HP: 41.2 3' Stab: 3.2 Penalty: 1.369

Self Any Secondary Structure:

Tm: 10.9°C dG: -6310 cal/mol dH: -53400 cal/mol dS: -152 cal/mol*K
 5' CCACGTGCTCGCAAAGTTG 3'
 |||||
 3' GTTGAACGCTCGTGACC 5'

Hairpin Secondary Structure:

Tm: 41.2°C dG: -282 cal/mol dH: -21200 cal/mol dS: -67 cal/mol*K
 5' CCACGTG
 ||
 3' GTTGAACGCTC

 Right Primer 5:

Start: 358 Length: 20 bp Tm: 59.4 C B: 93.583 % GC: 50.0 % Any: 0.0 End: 0.0 TB: 7.0 HP: 46.9 3' Stab: 3.9 Penalty: 0.604

Hairpin Secondary Structure:

Tm: 46.9°C dG: -651 cal/mol dH: -21000 cal/mol dS: -66 cal/mol*K
 5' TCACGCCATA
 |||
 3' AGCAAGGTTTC

Pair: Product Size: 328 bp Tm: 88.9 C Any: 0.0 End: 0.0 TB: 15.0 Penalty: 1.973

1	ATGGCAACAG	TTAACCAGCT	GGTACGCAAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCGCAAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTC	GTTACCACAC	CGTACGTGGT
301	GCGCTTGACT	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 6:

 Left Primer 6:

[Start](#): 19 [Length](#): 19 bp [Tm](#): 59.4 C [B](#): 93.441 % [GC](#): 57.9 % [Any](#): 9.1 [End](#): 9.1 [TB](#): 8.0 [HP](#): 44.7 [3' Stab](#): 4.5 [Penalty](#): 1.568

[Self Any Secondary Structure:](#)

Tm: 9.1°C dG: -5972 cal/mol dH: -53000 cal/mol dS: -152 cal/mol*K
 5' CTGGTACGCAAACCACGTG 3'
 |||||
 3' GTGCACCAAACGCATGGTC 5'

[Self End Secondary Structure:](#)

Tm: 9.1°C dG: -5972 cal/mol dH: -53000 cal/mol dS: -152 cal/mol*K
 5' CTGGTACGCAAACCACGTG 3'
 |||||
 3' GTGCACCAAACGCATGGTC 5'

[Hairpin Secondary Structure:](#)

Tm: 44.7°C dG: -609 cal/mol dH: -25200 cal/mol dS: -79 cal/mol*K
 5' CTGGTACG
 ||||
 3' GTGCACCAAC

 Right Primer 6:

[Start](#): 347 [Length](#): 20 bp [Tm](#): 60.7 C [B](#): 96.059 % [GC](#): 55.0 % [Any](#): 0.0 [End](#): 0.0 [TB](#): 6.0 [HP](#): 0.0 [3' Stab](#): 3.2 [Penalty](#): 0.669

Pair: [Product Size](#): 329 bp [Tm](#): 88.8 C [Any](#): 0.0 [End](#): 0.0 [TB](#): 14.0 [Penalty](#): 2.237

1	ATGGCAACAG	TTAACCAGCT	GGTACGCAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCGCACAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTC	GTTACCACAC	CGTACGTGGT
301	GCGCTTGACT	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 7:

 Left Primer 7:

Start: 16 Length: 19 bp Tm: 59.4 C B: 93.276 % GC: 57.9 % Any: 5.6 End: 2.0 TB: 6.0 HP: 44.7 3' Stab: 4.2 Penalty: 1.574

[Self Any Secondary Structure:](#)

Tm: 5.6°C dG: -5578 cal/mol dH: -50000 cal/mol dS: -143 cal/mol*K
 5' CAGCTGGTACGCAAACCAC 3'
 |||||
 3' CACCAAACGCATGGTCGAC 5'

[Self End Secondary Structure:](#)

Tm: 2.0°C dG: -1111 cal/mol dH: -79500 cal/mol dS: -253 cal/mol*K
 5' CAGCTGGTACGCAAACCA-C 3'
 |||| || |||| |
 3' CACCAAACGCATGGTCGAC 5'

[Hairpin Secondary Structure:](#)

Tm: 44.7°C dG: -609 cal/mol dH: -25200 cal/mol dS: -79 cal/mol*K
 5' CAGCTGGTACG }
 ||||
 3' CACCAAAC }

 Right Primer 7:

Start: 327 Length: 20 bp Tm: 60.7 C B: 96.052 % GC: 55.0 % Any: 5.0 End: 0.0 TB: 7.0 HP: 39.8 3' Stab: 4.4 Penalty: 0.670

[Self Any Secondary Structure:](#)

Tm: 5.0°C dG: -6529 cal/mol dH: -40800 cal/mol dS: -110 cal/mol*K
 5' GTCCTTAACGCCGGAGCAGT 3'
 ||||
 3' TGACGAGGCCGCAATTTCTG 5'

[Hairpin Secondary Structure:](#)

Tm: 39.8°C dG: -137 cal/mol dH: -15100 cal/mol dS: -48 cal/mol*K
 5' GTCCTTAACGCCG }
 ||
 3' TGACGAG }

 Pair: [Product Size:](#) 312 bp [Tm:](#) 88.7 C [Any:](#) 1.5 [End:](#) 0.0 [TB:](#) 13.0 [Penalty:](#) 2.244

[Compl Any Secondary Structure:](#)

Tm: 1.5°C dG: -4929 cal/mol dH: -48700 cal/mol dS: -141 cal/mol*K
 5' CAGCTGGTACGCAAACCAC 3'
 ||| ||
 3' TGACGAGGCCGCAATTTCTG 5'


```

1   ATGGCAACAG TTAACCAGCT GGTACGCAAA CCACGTGCTC GCAAAGTTGC
51  GAAAAGCAAC GTGCCTGCGC TGAAGCATG CCCGCAAAAA CGTGCGGTAT
101 GTACTCGTGT ATATACTACC ACTCCTAAAA AACGGAATC  CGCGCTGCGT
151 AAAGTATGCC GTGTTCGTCT GACTAACGGT TTCGAAGTGA CTTCTACAT
201 CGGTGGTGAA GGTACAACC TGCAGGAGCA CTCCGTGATC CTGATCCGTG
251 GCGTTCGTGT TAAAGACCTC CCGGGTGTTC GTTACCACAC CGTACGTGGT
301 GCGCTTGACT GCTCCGGCGT TAAAGACCGT AAGCAGGCTC GTTCCAAGTA
351 TGGCGTGAAG CGTCCTAAGG CTTAA

```

Pair 8:

 Left Primer 8:

Start: 50 Length: 19 bp Tm: 61.0 C B: 96.295 % GC: 57.9 % Any: 2.4 End: 2.4 TB: 9.0 HP: 46.5 3' Stab: 4.8 Penalty: 2.011

[Self Any Secondary Structure:](#)

Tm: 2.4°C dG: -3387 cal/mol dH: -62400 cal/mol dS: -190 cal/mol*K
 5' CGAAAAGCA---ACGTGCCTGC 3'
 ||| |||| |||
 3' CGTCCGTGCA---ACGAAAAGC 5'

[Self End Secondary Structure:](#)

Tm: 2.4°C dG: -3387 cal/mol dH: -62400 cal/mol dS: -190 cal/mol*K
 5' CGAAAAGCA---ACGTGCCTGC 3'
 ||| |||| |||
 3' CGTCCGTGCA---ACGAAAAGC 5'

[Hairpin Secondary Structure:](#)

Tm: 46.5°C dG: -560 cal/mol dH: -18800 cal/mol dS: -59 cal/mol*K
 5' CGAAAAGCAA₇
 ||| C
 3' CGTCCGTG₁

 Right Primer 8:

Start: 350 Length: 20 bp Tm: 59.3 C B: 93.217 % GC: 50.0 % Any: 0.0 End: 0.0 TB: 7.0 HP: 0.0 3' Stab: 3.9 Penalty: 0.685

 Pair: [Product Size:](#) 301 bp [Tm:](#) 88.5 C [Any:](#) 12.8 [End:](#) 12.8 [TB:](#) 16.0 [Penalty:](#) 2.696

[Compl Any Secondary Structure:](#)

Tm: 12.8°C dG: -4788 cal/mol dH: -75900 cal/mol dS: -229 cal/mol*K
 5' CGAAAAGCAA---CGTGCCTGC 3'
 |||| ||| ||
 3' TTCGTCCGAGCAAGGTTTCAT 5'

[Compl End Secondary Structure:](#)

Tm: 12.8°C dG: -4788 cal/mol dH: -75900 cal/mol dS: -229 cal/mol*K
 5' TACTTGGAACGAGCCTGCTT 3'
 || ||| |||||
 3' CGTCCGTGCA---ACGAAAAGC 5'

1	ATGGCAACAG	TTAACAGCT	GGTACGCAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCGCAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTT	GTTACCACAC	CGTACGTGGT
301	GCGCTTGACT	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 9:
Left Primer 9:

Start: 26 Length: 18 bp Tm: 61.9 C B: 96.713 % GC: 61.1 % Any: 10.9 End: 0.0 TB: 6.0 HP: 41.2 3' Stab: 5.1 Penalty: 3.948

Self Any Secondary Structure:

Tm: 10.9°C dG: -6310 cal/mol dH: -53400 cal/mol dS: -152 cal/mol*K
 5' GCAAACCCAGTGTCTCGCA 3'
 |||||
 3' ACGCTGTGCACCAAACG 5'

Hairpin Secondary Structure:

Tm: 41.2°C dG: -282 cal/mol dH: -21200 cal/mol dS: -67 cal/mol*K
 5' GCAAACCCAGTGT
 ||
 3' ACGCTC

Right Primer 9:

Start: 362 Length: 20 bp Tm: 60.7 C B: 96.154 % GC: 55.0 % Any: 0.0 End: 0.0 TB: 8.0 HP: 46.9 3' Stab: 3.5 Penalty: 0.739

Hairpin Secondary Structure:

Tm: 46.9°C dG: -651 cal/mol dH: -21000 cal/mol dS: -66 cal/mol*K
 5' CGCTTCAGCCATA
 |||
 3' AGGTTC

Pair: Product Size: 337 bp Tm: 88.9 C Any: 1.0 End: 2.3 TB: 13.0 Penalty: 4.687

Compl Any Secondary Structure:

Tm: 1.0°C dG: -4000 cal/mol dH: -54900 cal/mol dS: -164 cal/mol*K
 5' GCAAACCA-----CGTGTCTCGCA 3'
 ||| ||||
 3' AGGTTCATACCGCACTTCGC 5'

Compl End Secondary Structure:

Tm: 2.3°C dG: -3510 cal/mol dH: -61100 cal/mol dS: -186 cal/mol*K
 5' TCCAAGTATGGCGTGAAGCG 3'
 |||| ||
 3' CGTTTGGTGCACGA-GCGT 5'

1	ATGGCAACAG	TTAACCAGCT	GGTACGCAAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCACAAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTTC	GTTACCACAC	CGTACGTGGT
301	GCGCTTGACT	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GT TCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Pair 10:
Left Primer 10:

Start: 91 Length: 20 bp Tm: 60.2 C B: 95.486 % GC: 55.0 % Any: 0.0 End: 0.0 TB: 9.0 HP: 0.0 3' Stab: 4.5 Penalty: 0.179

Right Primer 10:

Start: 298 Length: 20 bp Tm: 60.0 C B: 95.221 % GC: 55.0 % Any: 21.1 End: 0.0 TB: 8.0 HP: 41.9 3' Stab: 3.9 Penalty: 0.041

Self Any Secondary Structure:

Tm: 21.1°C dG: -7672 cal/mol dH: -65400 cal/mol dS: -186 cal/mol*K
 5' CACGTACGGTGTGGTAACGA 3'
 |||||
 3' AGCAATGGTGTGGCATGCAC 5'

Hairpin Secondary Structure:

Tm: 41.9°C dG: -390 cal/mol dH: -25100 cal/mol dS: -80 cal/mol*K
 5' CACGT
 ||| A
 3' AGCAATGGTGTGGC^J

Pair: Product Size: 208 bp Tm: 87.3 C Any: 1.1 End: 0.0 TB: 16.0 Penalty: 0.220

Compl Any Secondary Structure:

Tm: 1.1°C dG: -4654 cal/mol dH: -50100 cal/mol dS: -147 cal/mol*K
 5' CGTGGCGTATGTACTCGTGT 3'
 |||| ||
 3' AGCAATGGTGTGGCATGCAC 5'

1	ATGGCAACAG	TTAACCAGCT	GGTACGCAA	CCACGTGCTC	GCAAAGTTGC
51	GAAAAGCAAC	GTGCCTGCGC	TGGAAGCATG	CCCGCAAAA	CGTGGCGTAT
101	GTACTCGTGT	ATATACTACC	ACTCCTAAAA	AACCGAACTC	CGCGCTGCGT
151	AAAGTATGCC	GTGTTCGTCT	GACTAACGGT	TTCGAAGTGA	CTCCTACAT
201	CGGTGGTGAA	GGTCACAACC	TGCAGGAGCA	CTCCGTGATC	CTGATCCGTG
251	GCGGTCGTGT	TAAAGACCTC	CCGGGTGTC	GTTACCACAC	CGTACGTGGT
301	GCGCTTGA	GCTCCGGCGT	TAAAGACCGT	AAGCAGGCTC	GTTCCAAGTA
351	TGGCGTGAAG	CGTCCTAAGG	CTTAA		

Tableau VI : Amorces PCR pour la validation de l'expérience de la substitution « K43 » à « T43 » du gène rpsL

	Left primer coordinates	Left primer	Left primer T _m	Right primer coordinates	Right primer	Right primer T _m	Product size
1	NC_011601.1:3728393-3728415	AGTGACTTCCTACATCGGTGGT	59.9	NC_011601.1:3728199-3728221	GTTTGGCCTTACTTAACGGAGA	59.7	216
2	NC_011601.1:3728393-3728415	AGTGACTTCCTACATCGGTGGT	59.9	NC_011601.1:3728201-3728223	TTGGCCTTACTTAACGGAGAAC	59.7	214
3	NC_011601.1:3728393-3728415	AGTGACTTCCTACATCGGTGGT	59.9	NC_011601.1:3728220-3728242	AACCATTAAGCCTTAGGACGCT	60.5	195
4	NC_011601.1:3728429-3728451	TAAAGTATGCCGTGTTTCGTCTG	60.2	NC_011601.1:3728199-3728221	GTTTGGCCTTACTTAACGGAGA	59.7	252
5	NC_011601.1:3728429-3728451	TAAAGTATGCCGTGTTTCGTCTG	60.2	NC_011601.1:3728220-3728242	AACCATTAAGCCTTAGGACGCT	60.5	231

Tableau VI : Liste des ARNg ciblant la séquence du gène rpsL

Rank	Target sequence	Genomic location	Strand	GC conten	Self-comp.	Efficiency
1	TGTGGTAACGAACACCCGGGAGG	NC_011601.1:3728311	+	60	0	78.01
2	ACGGCATACTTTACGCAGCGCGG	NC_011601.1:3728438	+	55	1	74.89
3	GGTCTTTAACACGACCACCACGG	NC_011601.1:3728332	+	50	1	72.35
4	CGGTGTGGTAACGAACACCCGGG	NC_011601.1:3728308	+	60	2	73.26
5	AACCGTTAGTCAGACGAACACGG	NC_011601.1:3728419	+	45	0	67.54
6	GAAGTGACTTCCTACATCGGTGG	NC_011601.1:3728394	-	50	0	65.59
7	GAAGCATGCCCGCAAAAACGTGG	NC_011601.1:3728505	-	55	0	63.63
8	GTTTCGTTACCACACCGTACGTGG	NC_011601.1:3728301	-	55	2	65.28
9	CAAGCGCACACGTACGGTGTGG	NC_011601.1:3728293	+	65	2	64.50
10	TCCGTGATCCTGATCCGTGGTGG	NC_011601.1:3728346	-	60	3	64.97
11	GTGACCTTCACCACCGATGTAGG	NC_011601.1:3728384	+	55	1	62.36
12	ACGGTGTGGTAACGAACACCCGG	NC_011601.1:3728307	+	55	3	63.60
13	CGGCGTTAAAGACCGTAAGCAGG	NC_011601.1:3728263	-	55	1	60.71
14	GGTCGTGTTAAAGACCTCCCGGG	NC_011601.1:3728325	-	55	0	59.12
15	CACTCCGTGATCCTGATCCGTGG	NC_011601.1:3728349	-	60	4	62.59
16	ACCACCACGGATCAGGATCACGG	NC_011601.1:3728345	+	55	3	61.14
17	ACTTCTACATCGGTGGTGAAGG	NC_011601.1:3728388	-	50	0	57.94
18	GCAGTCAAGCGCACACGTACGG	NC_011601.1:3728288	+	60	0	57.21
19	CTGCTTACGGTCTTTAACGCCGG	NC_011601.1:3728264	+	50	1	56.98
20	AATGGCAACAGTTAACAGCTGG	NC_011601.1:3728578	-	45	0	55.64
21	TGGTGAAGGTACAACCTGCAGG	NC_011601.1:3728374	-	55	1	56.31
22	GGAGTTCGGTTTTTTAGGAGTGG	NC_011601.1:3728459	+	45	0	55.30
23	CACGTGGTTTTGCGTACCAGCTGG	NC_011601.1:3728563	+	60	1	55.14
24	TTCGAAGTGACTTCCTACATCGG	NC_011601.1:3728397	-	40	1	55.06
25	GGATCACGGAGTGCTCCTGCAGG	NC_011601.1:3728359	+	65	3	56.80
26	TTCGCAACTTTGCGAGCACGTGG	NC_011601.1:3728547	+	55	3	56.75
27	CGTGGTGCCTGACTGCTCCGG	NC_011601.1:3728283	-	65	1	53.96
28	ACTTGGAAACGAGCCTGCTTACGG	NC_011601.1:3728251	+	50	2	54.70
29	TGCCGTGTTTCGTCTGACTAACGG	NC_011601.1:3728421	-	50	0	51.95
30	GAAAAGCAACGTGCCTGCGCTGG	NC_011601.1:3728527	-	60	0	51.24
31	GTACATACGCCACGTTTTTGCGG	NC_011601.1:3728496	+	45	0	49.44
32	TGGTCGTGTTAAAGACCTCCCGG	NC_011601.1:3728326	-	50	0	48.31
33	TAACACGACACCACGGATCAGG	NC_011601.1:3728338	+	55	2	49.76
34	GATGGCGTGAAGCGTCCTAAGG	NC_011601.1:3728230	-	55	1	48.23
35	ACTTTACGCAGCGCGGAGTTCGG	NC_011601.1:3728445	+	55	1	47.11
36	TAATGGTTCCTCCGTTAAGTAAGG	NC_011601.1:3728205	-	35	2	47.80
37	AAGCAGGCTCGTTCCAAGTATGG	NC_011601.1:3728247	-	50	2	45.94
38	ACGGAGAACCATTAAGCCTTAGG	NC_011601.1:3728214	+	45	4	46.27
39	TGTTGCCATTAATAGCTCCTGG	NC_011601.1:3728591	+	40	1	42.80
40	TACATACGCCACGTTTTTGCGGG	NC_011601.1:3728497	+	45	0	40.31
41	TGCGGGCATGCTTCCAGCGCAGG	NC_011601.1:3728514	+	70	0	40.06
42	TGAAGCGTCTAAGGCTTAATGG	NC_011601.1:3728222	-	45	3	41.83
43	AGGACGCTTCACGCCATACTTGG	NC_011601.1:3728234	+	55	1	37.22
44	AGCGCGGAGTTCGGTTTTTTAGG	NC_011601.1:3728454	+	50	1	32.03
45	TAAAACCAGGAGCTATTTAATGG	NC_011601.1:3728596	-	30	1	26.20

Pair	Left primer	Start	Lenght	Product	Right primer	Start	Lenght	Product
1	CGAAAAGCAACGTGCCGG	50	18	<pre> 5' CGAAAAGCAA C 3' GGCCGTG </pre>	CATATTTGCTGCGCGCCTG	352	19	<pre> 5' CATATTTGCTGCGCGCCTG 3' 3' GTCGCGCGTGTATTATAC 5' </pre>
2	ATGGCGACCGTGAACCAG	1	18	<pre> 5' ATGGCGACC 3' GACCAAGT </pre>	ATCTTTCACGCCGCTGCAAT	273	20	<pre> 5' ATCTTTCACGCCG 3' TAACGTC </pre>
3	CGACCGTGAACCAGCTGG	5	18	<pre> 5' CGACCGTGAACCAGCTGG 3' 3' GGTGACCAAGTGCCAGC 5' </pre>	GCCTGTTTGCATCTTTCAC G	238	21	<pre> 5' CGACCGTGAACCAG 3' GGTC </pre>
4	CCAGCTGGTGCGCAAACC	15	18	<pre> 5' CCAGCTGGTGCGCAAACC 3' CCAAACGCGTGTGCGACC 5' </pre>	TATTTGCTGCGCGCCTGTTT	350	20	<pre> 5' CCAGCTG 3' CCAAACGCGT </pre>