

---

# DMP du projet "RegOPepS"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe - DMP template (english)" fourni par Science Europe.

## Plan Details

<b>Plan title</b>	DMP du projet "RegOPepS"
<b>Version</b>	Mid term version
<b>Fields of science and technology (from OECD classification)</b>	Biological sciences (Natural sciences)
<b>Language</b>	fra
<b>Creation date</b>	2021-06-16
<b>Last modification date</b>	2021-06-16
<b>Identifiant</b>	RegOPepS-PGD-24 mois
<b>Identifiant type</b>	local identifiant

## Project Details

**Project title** RegOPepS

### Abstract

Le peptidoglycane (PG) est le constituant majeur de la paroi bactérienne et la cible d'un grand nombre d'antibiotiques, en particulier les  $\beta$ -lactamines. Ces antibiotiques inhibent la dernière étape de polymérisation du PG en se liant de manière irréversible aux protéines de liaison à la pénicilline (PLP). Jusqu'à récemment, les PLP étaient les seules enzymes connues permettant de catalyser la réticulation du PG. Cependant, mon équipe a montré que cette réaction est également catalysée par une deuxième famille d'enzymes, les L,D-transpeptidases (LDT). Les PLP sont potentiellement inhibées par toutes les classes de  $\beta$ -lactamines, tandis que les LDT sont uniquement inhibées par les carbapénèmes. La formation des ponts de type 3-3 assurant la réticulation du PG par les LDT est prépondérante chez les mycobactéries comme *Mycobacterium tuberculosis*. Chez les autres espèces, les LDT ne sont pas essentielles et leur contribution au cours de la phase exponentielle de croissance est négligeable. Cependant, la proportion de liaisons 3-3 augmente au cours de la phase stationnaire, suggérant un rôle des LDT dans la maintenance du PG. Chez *Escherichia coli*, j'ai récemment montré que les LDT confèrent une résistance à la plupart des  $\beta$ -lactamines en remplaçant les PLP chez des mutants sélectionnés *in vitro*. Ce mécanisme de résistance nécessite une production accrue de l'alarmone guanosine pentaphosphate et tétraphosphate [(p)ppGpp]. Cette dernière a initialement été identifiée par sa capacité à déclencher la réponse stringente. *E. coli* a donc la capacité de polymériser le PG par deux voies de réticulation alternatives, dont la contribution relative est contrôlée par le (p)ppGpp. Les objectifs généraux du projet sont l'identification des enzymes impliquées dans la voie des LDT et du rôle de l'alarmone dans la régulation de leur synthèse. Notre première

hypothèse de recherche est que la polymérisation du PG par les PLP et les LDT implique différents ensembles de protéines. En construisant des souches qui dépendent conditionnellement des PLP ou des LDT pour la synthèse du PG, nous étudierons trois types d'enzymes. (i) L'inactivation de gènes candidats et une approche de type Tn-Seq seront utilisées pour identifier les endopeptidases responsables du clivage des liaisons de type 3-3. Cette réaction permet l'insertion de nouvelles sous-unités dans le PG. (ii) Certaines PLP (enzymes de Classe B) contiennent à la fois un domaine transpeptidase pour la réticulation du PG et un domaine dit de morphogénèse qui reste mal caractérisé sur le plan fonctionnel. Nous caractériserons ces domaines et déterminerons s'ils coopèrent avec les LDT pour la polymérisation du PG. (iii) En plus des transpeptidases, la polymérisation du PG nécessite des glycosyltransférases pour l'élongation des chaînes glycanes. Parmi les six glycosyltransférases de E. coli, nous chercherons à identifier celles qui polymérisent les chaînes glycanes réticulées par les LDT. Notre deuxième hypothèse de recherche est que le (p)ppGpp est un médiateur clé contrôlant l'adaptation du métabolisme du PG aux carences nutritives et à d'autres stress. Ceci pourrait impliquer un contrôle négatif de la voie des PLP associée à une activation de la voie des LDT pour maintenir l'intégrité de la paroi cellulaire. Nous identifierons parmi les gènes régulés par le (p)ppGpp ceux qui sont directement impliqués dans la synthèse du PG ainsi que le mécanisme de cette régulation. L'impact scientifique du projet concerne la régulation de la synthèse du PG qui reste très mal connue. Pour la première fois, l'adaptation du métabolisme du PG aux carences nutritives et à d'autres stress sera évaluée globalement en mettant l'accent sur une synthèse accrue des liaisons de type 3-3 par les LDT et sur les facteurs essentiels à la fonction de ces transpeptidases. Les résultats attendus incluent l'identification de nouvelles cibles pour le développement d'antibiotiques actifs sur les souches résistantes.

## Funding

- ANR : ANR- 19-CE44-0007

## Produits de recherche :

1. Whole genome Sequencing data, HPLC spectra, MS data

## Contributeurs

Nom	Affiliation	Rôles
Jean-Emmanuel HUGONNET		<ul style="list-style-type: none"> <li>• Coordinateur du projet</li> <li>• DMP manager</li> <li>• Personne contact pour les données</li> </ul>

## Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre

soumission.

# DMP du projet "RegOPepS"

---

## 1. Data description and collection or re-use of existing data

Les profils HPLC générés par les analyses de peptidoglycane sont collectés sous forme de chromatogrammes à partir du logiciel associé au purificateur AKTA (Unikorn). Les données de spectrométrie de masse sont générées à partir des fichiers bruts du spectromètre de masse (Maxis II ETD, Bruker, France Bruker) et converties au format universel à l'aide du logiciel Mass Expert (<http://www.msxpertsuite.org>). Les données cinétiques obtenues par spectrophotométrie et spectrofluorimétrie seront stockées localement sous forme de fichiers Xcel

Les données génomiques brutes seront obtenues auprès d'une société de séquençage externalisée (Novogene).

L'identification des mutations impliquera l'accès au génome public annoté d'*Escherichia coli* K12 au NCBI.

---

Les protocoles expérimentaux sont conservés dans un cahier de laboratoire. Ces données sont également archivées sous forme électronique (fichiers .pdf) et papier (copies imprimées des pages du cahier de laboratoire). Dans la mesure du possible, les formats standards et ouopen seront privilégiés à des fins de partage et de réutilisation.

Les copies de sauvegarde de ces données sont réalisées régulièrement sur un disque dur externe qui n'est pas accessible par réseau. 100 chromatogrammes HPLC (environ 1 Go) logiciel associé au purificateur AKTA (Unikorn). Les spectrographes de masse (environ 100 fichiers bruts format Mass Expert) sont stockés électroniquement au laboratoire (5 To). Les autres données expérimentales (spectres UV-Vis, chromatogrammes) sont collectées à l'aide d'appareils dédiés et stockées sous forme électronique

(formats csv, xls, pdf) (moins de 1Go). Des données génomiques brutes ont jusqu'à présent été produites pour 70 souches (format fastq) pour une taille totale d'environ 210 Go (3 Go par génome).

Les analyses Tn-Seq sont sauvegardées localement sur un serveur dédié.

Les données brutes de séquençage du génome entier des mutants slY (M1.1 à M1.4) sont disponibles dans la base de données Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA748867>).

---

## 2. Documentation and data quality

Les métadonnées sont incluses dans : Le rapport Unikorn associé aux fichiers de chromatogramme (nom de l'échantillon, flux, type de colonne, composition du tampon, description du gradient, longueur d'onde de détection) ; dans le format Mass expert (numéro de balayage, date de plage de balayage de l'énergie de collision d'acquisition) pour les données de spectrométrie de masse.

Pour les génomes que nous avons rendus publics, dans la base de données Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA748867>).

---

Le contrôle de la qualité des données est assuré par notre société de sous-traitance de séquençage du génome (Novogene).

Les bonnes pratiques de laboratoire sont suivies pour le contrôle et la qualité des données.

La reproductibilité des données est assurée par le fait que les expériences sont réalisées plusieurs fois et que plusieurs méthodes de caractérisation (génétique et biochimique) sont utilisées.

La qualité et la conformité des données de spectrométrie de masse seront vérifiées par les différents contrôles qualité et des étalonnages des appareils effectués régulièrement par le responsable de la plateforme d'analyse de masse (plateforme de masse du MNHN).

Les temps de rétention observés et publiés pour la séparation des fragments de peptidoglycane sont comparés.

Les masses monoisotopiques observées sont comparées aux masses mono isotopiques calculées pour déterminer si les différences en ppm sont inférieures à celles attendues de la résolution du spectromètre de masse (Maxis II ETD, Bruker, France Bruker).

---

## 3. Storage and backup during the research process

Les données sont stockées sur un serveur local dans notre laboratoire. Les sauvegardes sont stockées sur des disques durs externes.

---

Aucune donnée sensible n'est collectée pour notre projet jusqu'à présent.

---

## 4. Legal and ethical requirements, codes of conduct

Non applicable pour ce projet.

---

Le porteur du projet respecte les règles des tutelles scientifiques en matière de propriété intellectuelle et de brevet. Tous les scientifiques impliqués dans le projet seront associés aux publications en fonction de leurs contributions respectives. Afin de promouvoir l'accessibilité, des publications en libre accès seront privilégiées, dans la mesure du possible. A l'issue du processus de publications et en respectant les politiques éditoriales des éditeurs, les publications issues du projet seront déposées dans une archive ouverte, libre d'accès (HAL, ...)

---

Non applicable

---

## 5. Data sharing and long-term preservation

Les résultats obtenus seront utilisés pour la rédaction de deux thèses de doctorat. Ils seront également publiés dans des revues à comité de lecture appropriées. A l'issue du processus de publications et en respectant les politiques éditoriales des éditeurs, les publications liées au projet seront déposées dans une archive ouverte, libre d'accès (HAL, ...)

---

L'ensemble des données générées sera unique et devra donc être conservé sans mention de délai. Aucune donnée ne sera détruite. Le stockage des données brutes de séquençage utilisées pour les publications uniquement sont accessibles sur le NCBI.

---

Pour les analyses de données, pipeline Breseq. <https://barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing> Pour les données MS Mass Expert Logiciel Unikorn pour les profils HPLC.

---

L'attribution d'un identifiant unique et pérenne (DOI) sera assurée par la publication des données dans des revues scientifiques à comité de lecture. Les autres données archivées localement ne se verront pas attribuer d'identifiant unique et seront identifiés par les informations contenues dans les métadonnées associées.

---

## 6. Data management responsibilities and resources

Le coordinateur du projet

---

Les personnes travaillant sur le projet dédieront une partie de leur temps de travail à assurer que les données soient "FAIR".