

---

# DMP du projet "Deciphering the pathways of microbial sulfur disproportionation in hydrothermal taxa"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe : modèle structuré" fourni par Science Europe.

## Renseignements sur le plan

<b>Titre du plan</b>	DMP du projet "Deciphering the pathways of microbial sulfur disproportionation in hydrothermal taxa"
<b>Version</b>	Version initiale
<b>Domaines de recherche (selon classification de l'OCDE)</b>	Biological sciences (Natural sciences)
<b>Langue</b>	fra
<b>Date de création</b>	2022-12-08
<b>Date de dernière modification</b>	2022-12-09

## Renseignements sur le projet

<b>Titre du projet</b>	Deciphering the pathways of microbial sulfur disproportionation in hydrothermal taxa
<b>Acronyme</b>	MISD

## Résumé

Le soufre est un élément omniprésent et abondant dans les environnements hydrothermaux, et est crucial pour la production d'énergie. La dismutation microbienne des composés inorganiques soufrés (=MSD) est moins documentée que les autres réactions d'oxydo-réduction du cycle du soufre (sulfato-réduction, sulfo-réduction et sulfo-oxydation). Pourtant, la MSD pourrait contribuer de manière significative aux flux de soufre des écosystèmes naturels, notamment hydrothermaux. Les voies cataboliques de la MSD ne sont que partiellement élucidées, et impliquent chez certains taxons des enzymes de la voie de sulfato-réduction. La méconnaissance des mécanismes moléculaires qui sous-tendent ce processus empêche l'application de méthodes d'investigation génomique pour identifier et quantifier les taxons mettant en œuvre ce processus dans les habitats naturels. La part de la dismutation dans les flux d'espèces soufrées des habitats naturels par rapport à ceux de la sulfo-oxydation et de la sulfato-réduction n'est pas connue, car la MSD est confondue avec ces voies dans les bilans globaux, puisqu'elle conduit à la production de sulfates et de sulfures. La chronologie de l'apparition de la dismutation du soufre et de la sulfato-réduction n'est pas non plus établie, car les données sur les isotopes du soufre dans les archives géologiques anciennes n'ont pas permis de tirer une conclusion sans équivoque. Ce projet ambitieux vise à isoler de nouveaux modèles hydrothermaux dismutant les composés inorganiques soufrés, à étudier leur écophysiologie, à décrypter les chemins métaboliques empruntés et les intermédiaires chimiques de cette réaction et à rechercher des gènes marqueurs de cette réaction. Il propose également de tenter de déterminer si ce métabolisme microbien est antérieur ou postérieur à celui de la sulfato-réduction. Pour aborder ces questions, des approches culturelles, physiologiques, génomiques, protéomiques, analytiques et phylogénomiques/phylogénétiques seront mises en œuvre.

## Sources de financement

- Agence Nationale de la Recherche : ANR-22-CE02-0001

## Date de début

2022-10-01

## Date de fin

2026-09-30

## Partenaires

- BIOLOGIE ET ECOLOGIE DES ECOSYSTEMES MARINS PROFONDS (200412248A)
- Plateforme d'Analyse Protéomique de Paris Sud-Ouest, UMR MICALIS /INRAE, Jouy-en-Josas (201119643H)

## Produits de recherche :

1. Default research output (Jeu de données)

## Contributeurs

Nom	Affiliation	Rôles
Alain Karine - <a href="https://orcid.org/0000-0002-8812-9081">https://orcid.org/0000-0002-8812-9081</a>	BEEP CNRS	<ul style="list-style-type: none"> <li>• Coordinateur du projet</li> <li>• Personne contact pour les données</li> <li>• Responsable de la conservation à long terme des données</li> <li>• Responsable du plan de gestion de données</li> </ul>
Les responsables scientifiques du projet		<ul style="list-style-type: none"> <li>• Responsable des questions éthiques</li> <li>• Responsable du dépôt et de la diffusion des données</li> <li>• Responsable du stockage des 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 "Deciphering the pathways of microbial sulfur disproportionation in hydrothermal taxa"

---

## 1. Description des données et collecte ou réutilisation de données existantes

### 1.1 Description générale du produit de recherche

<b>Nom</b>	Default research output
<b>Description</b>	<p>Ce projet conduira à la production de données nouvelles sur la biologie d'espèces bactériennes issues de sources hydrothermales océaniques profondes. Ces données seront générées à partir de souches préexistantes disponibles dans des collections publiques de microorganismes (<a href="https://www.dsmz.de/">https://www.dsmz.de/</a>), et de souches microbiennes qui seront isolées dans le cadre de cette étude à partir d'échantillons hydrothermaux profonds.</p> <p>Ces nouvelles données seront de natures diverses: nouveaux isolats microbiens, génomes, protéomes, transcriptomes, données de physiologie, images.</p> <p>Concrètement, les nouveaux isolats microbiens seront obtenus en réalisant des cultures d'enrichissement et des isolements à partir d'échantillons de roches hydrothermales qui ont été prélevées dans le cadre des campagnes océanographiques MoMARSAT2019 (<a href="https://doi.org/10.17600/18001110">https://doi.org/10.17600/18001110</a>), MoMARSAT2020 (<a href="https://doi.org/10.17600/18000684">https://doi.org/10.17600/18000684</a>) et HERMINE2 (<a href="https://doi.org/10.17600/18001851">https://doi.org/10.17600/18001851</a>). Les roches ont été broyées, sous une atmosphère anaérobie (sous azote) et stérile, à bord du bateau, puis conditionnées dans des flacons Schott stockés à 4°C jusqu'à la mise en culture au laboratoire.</p>
<b>Type</b>	Jeu de données
<b>Mots clés (texte libre)</b>	Source hydrothermale océanique profonde, dismutation du soufre, physiologie, protéome, évolution
<b>Langue</b>	fra
<b>Date de publication</b>	2022-12-09
<b>Contient des données personnelles ?</b>	Non
<b>Contient des données sensibles ?</b>	Non
<b>Prend en compte des aspects éthiques ?</b>	Non

---

### 1.2 Est-ce que des données existantes seront réutilisées ?

<b>Justification</b>	Une ou des souche(s) disponible(s) dans des collections publiques de microorganismes feront l'objet d'études fondamentales dans le cadre de ce projet. Cela concerne notamment la bactérie <i>Dissulfuribacter thermophilus</i> S69T (DSM 25762T) dont le protéome sera étudié durant la première année de ce projet.
----------------------	---

---

### 1.3 Comment seront produites/collectées les nouvelles données ?

**Titre de la méthode**

Données expérimentales

**Description**

Les types et formats de données produites dans le cadre de ce projet sont listées dans le Tableau 1 ci-dessous.

Tableau 1. Type et formats des données brutes et données traitées

Méthode	Format des données brutes	Format des données après traitement	Objectif
Suivi de croissance bactérienne	Nombres	CSV ; Graphiques en TIFF, PNG ou SVG	Détermination des gammes de T°, pH et salinités de croissance
Suivi de croissance bactérienne	Nombres	CSV ; Graphiques en TIFF, PNG ou SVG	Détermination des donneurs et accepteurs d'électrons et des sources de carbone
Séquençage de génomes	Séquences d'ADN (FASTQ)	FASTA	Description du potentiel génétique; génomique comparative
Imagerie: microscopie électronique à balayage et microscopie électronique à transmission	Images	TIFF	Visualisation des nouveaux isolats
Protéomique	Spectre MS/MS (RAW ou msXML)	CSV; ; Graphiques en TIFF, PNG ou SVG	Protéomes obtenus dans différentes conditions cataboliques
Chromatographie ionique	Données brutes de chromatographie	XLSX ou CSV; ; Graphiques en TIFF, PNG ou SVG	Suivi des substrats, intermédiaires et produits du catabolisme au cours de la croissance
Sondes chimiques	RAW	XLSX ou CSV; ; Graphiques en TIFF, PNG ou SVG	Suivi des intermédiaires et produits du catabolisme au cours de la croissance
Transcriptomique	FASTQ	FASTA; XLSX ou CSV	Transcriptomes obtenus dans différentes conditions cataboliques
Phylogénomique/phylogénie	Séquences ADN (FASTQ ou FASTA) produites dans ce travail ou issues de banques de données publiques)	NEWICK	Phylogénie organismales Phylogénies basées sur des protéines du métabolisme du soufre

**Nature des données**

Données expérimentales

## 2. Documentation et qualité des données

### 2.1 Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

## Description

Les métadonnées ont pour but d'assurer la traçabilité des échantillons et des données générées. Elles incluent pour chaque échantillon ou sous-échantillon prélevé au niveau des sources hydrothermales océaniques profondes :

Le nom de la campagne océanographique, la date du prélèvement, les coordonnées géographiques et la profondeur du prélèvement, les identifiants uniques du prélèvement (e.g. numéro de pongée ROV ou Nautile), les identifiants uniques de l'échantillon (e.g. PBT), les identifiants uniques du sous-échantillon (e.g. roche), le mode de préservation de l'échantillon ou du sous-échantillon, le détenteur de l'échantillon ou du sous-échantillon.

Les données générées au cours des travaux de recherche sont consignées dans les cahiers de laboratoire, et restent la propriété du laboratoire après le départ des personnels non permanents ayant travaillé sur un projet de recherche. De même, l'ensemble des données numériques générées et des protocoles ayant été mis en œuvre sont archivés au laboratoire.

---

## 2.2 Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

### Description

L'ensemble des expérimentations sont réalisées en réplicats avec les témoins positifs et négatifs appropriés (e.g. triplicats pour les caractérisations physiologiques des souches; quadruplicata pour les analyses fonctionnelles/études protéomiques), et sont répétées au besoin.

Les appareils qui sont mis en œuvre dans le cadre de ce projet font l'objet d'une calibration régulière (e.g. chromatographie ionique...).

Les données produites sont interprétées par plusieurs membres du consortium.

---

## 3. Exigences légales et éthiques, code de conduite

### 3.1 Quelles seront les mesures appliquées pour assurer la protection des données à caractère personnel ?

#### Description

Le projet ne traite pas de données à caractère personnel.

---

### 3.2 Quelles sont les contraintes juridiques (sensibilité des données autres qu'à caractère personnel, confidentialité, ...) à prendre en compte pour le partage et le stockage des données ?

#### Description

Les données qui seront générées dans ce projet ne sont pas sensibles. Elles seront gérées de sorte à ce que ces données soient « Faciles à trouver, Accessibles, Interopérables et Réutilisables », c'est à dire selon les principes FAIR (*Findable, Accessible, Interoperable and Reusable*). Elles ne resteront fermées et accessibles uniquement aux membres du consortium, que jusqu'à la publication des travaux. Ensuite, elles seront aussi ouvertes que possible: dépôt des isolats dans des collections publiques (UBOCC et DSMZ), dépôt des génomes et transcriptomes dans des banques de données publiques (SRA, EMBL-EBI/GenBank/DBJ), publication dans des journaux en Open Access et/ou des archives ouvertes...

---

### 3.3 Quels sont les aspects éthiques à prendre en compte lors de la collecte des données ?

#### Description

Chacun des participants au projet est tenu de respecter les chartes d'éthique et de déontologie de son organisme d'affiliation dont les principes sont définis par la Charte nationale de déontologie des métiers de la recherche.

---

## 4. Traitement et analyse des données

#### 4.1 Comment et avec quels moyens seront traitées les données ?

<b>Description</b>	Les résultats du projet seront publiés dans des revues à comité de lecture. Les données sur lesquelles s'appuient les publications seront déposées dans des archives ouvertes, libre d'accès (GenBank, Seanoe,...). 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, Archimer).
--------------------	---

---

## 5. Stockage et sauvegarde des données pendant le processus de recherche

#### 5.1 Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

<b>Besoins de stockage</b>	Au cours du processus de recherche, les données sont stockées sur les postes de travail chiffrés des membres du projet. Des sauvegardes sur un serveur propre à chaque équipe (Lina pour BEEP) y sont effectuées plusieurs fois par jour.
<b>Volume estimé des données</b>	5
<b>Unité</b>	To
<b>Mesures prises pour la sécurité des données</b>	Les données sont stockées sur les postes de travail chiffrés des membres du projet et sur des serveurs propres à chaque équipe.

---

## 6. Partage des données et conservation à long terme

#### 6.1 Comment les données seront-elles partagées ?

<b>Modalités de partage</b>	<p>Les résultats du projet seront publiés dans des revues à comité de lecture. Les données sur lesquelles s'appuient les publications seront déposées dans des archives ouvertes, libre d'accès (EMBL-EBI/GenBank, SRA, Seanoe,...). 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, Archimer).</p> <p>Les souches isolées dans le cadre de ce projet seront déposées dans des collections de microorganismes publiques (UBOCC UBO culture collection; sous-collection: marine; <a href="https://nouveau.univ-brest.fr/ubocc/fr/page/presentation-des-collections">https://nouveau.univ-brest.fr/ubocc/fr/page/presentation-des-collections</a> // collection allemande de microorganismes DSMZ ; <a href="https://www.dsmz.de/">https://www.dsmz.de/</a>).</p>
-----------------------------	--

---

#### 6.2 Comment les données seront-elles conservées à long terme ?

**Justification**

Les métadonnées liées aux campagnes sont intégralement archivées par le SISMER (Système d'Informations Scientifiques pour la Mer) sans limitation de durée, et dans la base CAMPAGNES Océanographiques accessible *via* le Catalogue des campagnes océanographiques françaises (<https://campagnes.flotteoceanographique.fr/>) de manière pérenne.

Les données utilisées dans les publications sont archivées dans Seanoé, EMBL-EBI/GenBank/DDBJ, SRA sans limitation de durée.

**Volume estimé des données** 5

**Unité** To

**Archive** :