

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة أبو بكر بلقايد- تلمسان

Université ABOUBEKR BELKAID – TLEMCEM

كلية علوم الطبيعة والحياة، وعلوم الأرض والكون

Faculté des Sciences de la Nature et de la Vie, et des Sciences de la Terre et de l'Univers

Département Biologie



**Mémoire**

En vue de l'obtention du

**Diplôme de Master**

**En science biologiques**

Option :

**Biochimie**

Thème

***Etude des caractéristiques des peroxydases  
d'origine végétale par bio-informatique***

**Réalisé par :**

**AINA chaimaa**

Examiné le 10 /07/2021 par le jury composé de :

Examineur 1 : **Pr Rahmoun Mohammed Nadjib**

Université de Tlemcen

Examineur 2 : **Dr Bedrane Mohammed El Amine**

Université de Tlemcen

Encadrante : **Dr Benmansour Meriem**

Université de Tlemcen

**Année universitaire 2020/2021**

## *Remerciements*

Tout d'abord je remercie Dieu, tout puissant, Seigneur des mondes qui nous a accordé la santé, le courage et la force pour la réalisation et la finition de ce travail.

Je remercie tout particulièrement Mme Benmansour M., maitre de conférences classe B au département de biologie, faculté des sciences de la nature et de la vie et des sciences de la terre et de l'univers, Université Abou BekrBelkaid Tlemcen, pour avoir accepté de diriger ce travail, pour ses conseils, son aide, ses suggestions sur la rédaction de ce mémoire ainsi que la confiance qu'elle m'a témoigné tout au long de ce travail.

Mes remerciements vont également à Mr Rahmoun Mohammed Nadjib, « Professeur au Département de Biologie, Faculté des SNV-STU, Université Abou Bakr Belkaïd de Tlemcen » qui me fait l'honneur et le plaisir d'avoir accepté d'être examinateur de mon manuscrit.

Je remercie également Mr Bedrane Mohammed El Amine, maître assistant A au département de Biologie, Faculté des SNV-STU, Université Abou Bakr Belkaïd de Tlemcen, d'avoir accepté d'être examinateur de mon manuscrit.

Je remercie également Mr.Bendahmane A. pour ses compétences en bioinformatique et le partage de ses expériences.

## *Dédicace*

Je dédie ce modeste travail aux personnes qui me sont très chers

A mes très chers parents, qui m'ont toujours soutenue et encouragé à bien travailler dans tous ce que j'entreprends et plus particulièrement dans mes études.

Que ce travail soit le témoignage de mon amour.

A mes frère « Achref » « Islem » « Louay » « Younesse » pour leur soutien sa confiance et leur encouragement si précieux au quel je souhaite beaucoup de réussite et tout le bonheur du monde.

A tout ma famille pour leur soutien tout au long de mon parcours universitaire

A toutes mes copines « Wissem » « Kaouter » « Riheb » « Zokha » et tous mes collègues

A tous ceux qui me sont chers. Que ce travail soit l'accomplissement de vos, et le fruit de votre soutien infailible, Merci d'être toujours là pour moi.

**Résumé :**

Les peroxydases sont des enzymes qui catalysent la réaction d'oxydo-réduction par un mécanisme de radical libre permettant la transformation de plusieurs composés en produits oxydés ou polymérisés, qui peuvent peut-être isolés à partir de plusieurs sources végétales tel que le radis, le navet et l'ail.

L'objectif de ce mémoire consiste en l'étude des caractéristiques des peroxydases d'origine végétale par bio-informatique en déterminant les séquences protéiques afin de comparer les homologies de séquence des différentes enzymes observées.

L'alignement multiple des séquences des peroxydases, a été réalisé au moyen de programmes bio-informatiques, tel que le MEGA X pour aligner les séquences et ESPript 3.0 pour analyser et visualiser les résultats.

Les résultats de cet alignement ont montré que la peroxydase a plusieurs homologues mais notre étude, nous nous sommes intéressés qu'à celles ayant un fort pourcentage de similitude. Nous avons retenu pour cela quatre résultats dont le pourcentage supra-cité varie entre 91% et 100%, et où les acides aminés ; His170, His42, et Arg38 sont conservés autour de site actif.

**Mots clés :**

Peroxydases végétales , Alignement multiple, MEGA X , ESPript 3.0.

## Liste des figures

<b>Figure 1</b> : Schéma de la structure générale des peroxydases des plantes (Delannoy <i>et al.</i> , 2004).....	05
<b>Figure 2</b> : Structure cristalline de la peroxydase C de HRP (Veitch, 2004).....	06
<b>Figure 3</b> : Cycle catalytique des peroxydases (Pelmont, 1988).....	07
<b>Figure 4</b> : Modèle de la réaction de HRP-C avec H <sub>2</sub> O <sub>2</sub> comme unique substrat.....	08
<b>Figure 5</b> : Profil de la base de données UniProtKB.....	13
<b>Figure 6</b> : Profil de l'outil BLASTE.....	14
<b>Figure 7</b> : Application du serveur BLAST.....	15
<b>Figure 8</b> : Application du MEGA X.....	16
<b>Figure 9</b> : Etapes de téléchargement des séquences sous format FASTA.....	16
<b>Figure10</b> : Etapes de l'alignement.....	17
<b>Figure 11</b> : Etapes de l'alignements.....	17
<b>Figure12</b> : Etapes de l'alignements.....	18
<b>Figure 13</b> : Profil du serveur web Esript 3.0.....	19
<b>Figure14</b> : Séquence de la peroxydase au format FASTA.....	20
<b>Figure 15</b> : Les homologues structuraux de la peroxydase P00433.....	21
<b>Figure 16</b> : Résultats d'alignement de 100 homologues de la peroxydase P00433 par MEGA X. ....	22
<b>Figure 17</b> : Résultats d'alignement des 4 homologues de la peroxydase P00433 par MEGA X..	22
<b>Figure 18</b> : Alignement de séquences multiple de la peroxydase P00433 avec ses homologues structuraux.....	23

## Liste de tableau

**Tableau 1** : Quelques algorithmes utilisés pour l'alignement de séquences multiples.....12

**Tableau 2** : Les enzymes les plus homologues à la peroxydase P00433.....20

# Table des matières

## Introduction générale

## Synthèse bibliographique

### Partie 1 : Peroxydase

1.. Historique .....	03
2.. Définition .....	03
3.. Source et classification .....	03
4. Localisation .....	04
5.. Structure .....	04
6.. Site actif .....	06
7.. Mécanisme d'action .....	06
8.. Rôle physiologique .....	09
9.. Application biotechnologique des peroxydases .....	09

### Partie 2 : Analyse des protéines par bioinformatique

1. Notion de la bioinformatique.....	10
2. Banques de données .....	10
3. Alignement multiple des séquences.....	11

### Matériel et méthodes

1. Téléchargement et organisation des données .....	13
2. Recherche des homologues de séquence .....	13
3. Alignement de séquences multiple .....	15
4. Analyse des séquences.....	18

### Résultats et discussion

1. Le format FASTA de la séquence .....	20
2. Les homologues de la peroxydase horseradish .....	20
3. Résultats de l'alignement .....	22
4. Analyse de l'alignement multiple de séquences .....	23

<b>Conclusion</b> .....	24
-------------------------	----

### Références bibliographiques





# **Introduction générale**

### *Introduction*

Les enzymes sont des catalyseurs biologiques qui peuvent accélérer les réactions métaboliques. Une molécule enzymatique peut convertir jusqu'à des millions de molécules de substrat (Kühne et Wilhelm, 1877).

Parmi les milliers d'enzymes qui ont été identifiées et souvent étudiées dans les moindres détails, l'attention s'est portée sur la peroxydase. Cette enzyme se trouve chez les animaux, les plantes, les champignons et les bactéries aérobies. Elle possède une grande importance biologique grâce à son pouvoir d'oxyder les composés phénoliques étant toxiques pour les cellules et l'environnement [(Delannoy *et al.*, 2004) ; (Rampel, 2007) ; (Kalin, 2014)].

La source traditionnelle de la peroxydase à haute performance est le raifort cultivé, « *Armonioica rusticana* » communément appelée radis du cheval et en anglais horseradish (Krell, 1991), qui a été largement étudiée et commercialisée. Son utilisation dans diverses applications analytiques donne un rendement favorable. Vu l'intérêt que peut encore offrir cette enzyme dans plusieurs domaines biotechnologiques, l'axe de recherche reste toujours orienté vers la prospection d'autres sources riches en peroxydase (Diao *et al.*, 2019).

Il n'y a pas longtemps, une nouvelle discipline a été créée nommée « bioinformatique », située au carrefour de la biologie moléculaire, des mathématiques, des statistiques de l'informatique qui permet de modéliser, analyser et simuler l'information biologique. L'alignement multiple des séquencées est une tâche de base pour beaucoup d'application en bioinformatique utilisée par des biologistes pour analyser des séquences d'ADN (nucléiques) ou de protéine (protéique) pour déterminer leur degré d'homologie ou de divergence [(Sharma *et al.*, 2018) ; (Kumar *et al.*, 2020)].

C'est la raison pour laquelle, nous nous sommes intéressés dans cette étude à d'autres sources de peroxydases comme celles trouvées dans le radis rose ou le navet, comme alternatif à la HRP, en déterminant *in silico* les caractéristiques de ces protéines.

Notre objectif est de déterminer et comparer les séquences homologues de la peroxydase végétale HRP par la bioinformatique. L'alignement et l'analyse des séquences protéiques est réalisée à l'aide de deux programme MEGA X (Sudhir *et al.*, 2008) et ESPrit 3.0 (Kumar *et al.*, 2020).

Nous avons organisé notre mémoire comme suit :

- Dans le premier chapitre, une revue bibliographique constituée de deux parties

## *Introduction générale*

---

- La première partie rassemble les informations sur la peroxydase.
  - La deuxième traite la notion de la bioinformatique et de son développement en biologie.
- Le deuxième Chapitre matériel et méthodes, explique les méthodes et les techniques bioinformatiques appliquées dans ce travail.
- Le troisième Chapitre présente les résultats obtenus des différentes méthodes et leur interprétation.
- Enfin, cette étude se termine par une conclusion générale.

# CHAPITRE 1 : SYNTHÈSE

## BIBLIOGRAPHIQUE

# Partie 1 : La peroxydase

### 1. Historique

La peroxydase est parmi les premiers enzymes étudiés. En 1810, Planche a observé l'apparition d'une coloration de racine de Raifort en présence du peroxyde d'hydrogène ( $H_2O_2$ ) (Planche, 1810).

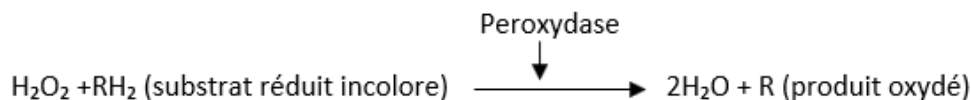
Ensuite, en 1855, Schönbein a noté que les extraits végétaux ou animaux étaient oxydés en présence d'une solution diluée de peroxyde d'hydrogène et de gaïacol (Schönbein, 1855). Plus tard, en 1898 Linosier en travaillant sur l'extrait de pus humain, qui a la même capacité à s'oxyder en détruisant le peroxyde, proposa le nom peroxydase (Linosier, 1855).

En 1976, la première image cristalline et la première séquence de protéine peroxydase sont apparues grâce aux travaux de Brathwaite et Welinder (Brathwaite *et al.*, 1976). En 1988, la première séquence nucléotidique a été déterminée par Fujiyama (Fujiyama, 1976).

### 2. Définition

Les peroxydases sont largement présentes chez tous les êtres vivants : animaux, végétaux et micro-organismes (Siúeciogluetal., 2010 ; Kumar et Jagannadham, 2011 ; Zheng, 2012).

La peroxydase est une enzyme intraet extracellulaire avec d'importants rôles dans les processus cellulaires (Everse et Grisham, 1990). Elle catalyse la réaction d'oxydation du  $H_2O_2$  et intervient dans le métabolisme des parois, le métabolisme hormonal, en réaction aux stress abiotiques et dans les interactions plantes/micro-organismes (Sreenivasulu *et al.*, 1999 ; Velikova., 2005). Elles catalysent la réaction d'oxydation d'une grande variété de substrats et participent également à la défense de la cellule contre les peroxydes toxiques (Yemenicio et Cestero, 1998 ; Rompel, 2007 ; KalÕn, 2014). La réaction générale s'écrit :



### 3. Source et classification

La peroxydase existe chez presque tous les êtres vivants : animaux, végétaux et microorganismes (champignons et bactéries) (Siúeciogluetal., 2010 ; Kumar et Jagannadham, 2011 ; Zheng, 2012). Elle ne manque que chez les anaérobies stricts car l'oxygène est un poison pour elle. En 1992, Welinder classifié les peroxydases selon leur séquence d'acides

aminés, en deux super-familles. La première pour les peroxydases du règne animal, tandis que la seconde inclut les plantes, les champignons et les bactéries. Ce dernier est divisé en trois classes : la classe I est composée de peroxydases intracellulaires, non glycosylées et sans calcium, dont le cytochrome C, les ascorbates et les peroxydases d'origine bactérienne à activité catalase. La classe II regroupe les peroxydases élaborées (extracellulaires) par les champignons telles que la lignine / manganèse peroxydase. La classe III renferme les peroxydases secrétées à partir des plantes (Welinder,1992).

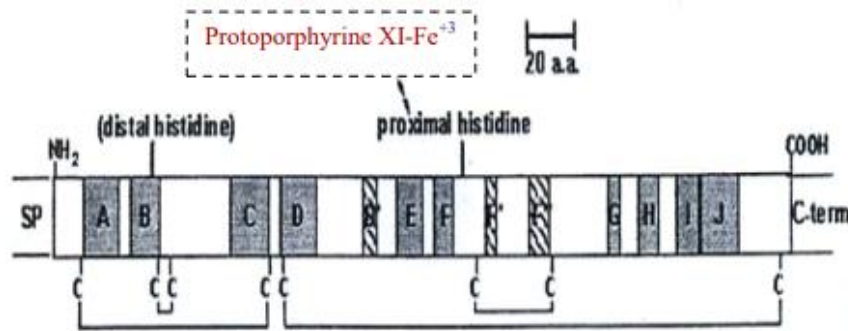
#### **4. Localisation**

Les peroxydases sont majoritairement présentes dans les tissus végétaux. En effet, les résultats des études histologiques de coloration de co-substrats, ont montré que les peroxydases étaient localisées dans différents compartiments : paroi pecto-cellulosique, réticulum endoplasmique, vacuole dictyosome et mitochondrie. Cette localisation peut varier selon les saisons et les espèces (Mc lellan et Robinson,1987).

#### **5. Structure**

Les peroxydases d'origine végétale ont un poids moléculaire entre 32 et 45 KDa (Gaspar *et al.*, 1982), alors que selon d'autres il est situé entre 29 et 59 KDa (Skharove *et al.*, 2000). La peroxydase présente trois domaines : deux dans la partie distale (N-Terminal) et le troisième dans la partie proximal (C-Terminal). Ces parties se lient à un ion  $Ca^{2+}$  qui joue le rôle de stabilisateur de la protéine repliée. Cette apoenzyme se lie d'une façon non covalente au groupe prosthétique " hème " défini par la présence d'un ion ferrique au centre d'une porphyrine (Szigeti *et al.*,2008 ; Dunford,1999 ; Heriksen *et al.*,1998).

Les ADN chromosomiques codant pour la peroxydase ont été distingué à partir de plusieurs espèces de plantes (Welinder *et al.*, 1996). Les gènes séquencés interagissent avec l'ion ferrique par la présence de deux molécules d'histidine : l'une est placée dans la cavité distale entourant l'hème (Histidine distale) et l'autre, située à l'opposé, à proximité de l'hème (His proximale) (Delannoy *et al.*, 2004).



**Figure 1** : Schéma de la structure générale des peroxydases des plantes (Delannoy *et al.*, 2004).

*Les hélices alpha sont indiquées (A à J) / SP : peptide signal / C-term*

A pH neutre, la structure cristalline, a montré la présence de quatre molécules d' $\text{H}_2\text{O}$  dans la poche du site actif. Une distance entre la molécule d'eau et l'atome ferrique qui lui est plus proche est de  $4 \text{ \AA}$  (Henriksen *et al.*, 2001). En 1993, Dunfond montre qu'il n'y a pas de liaison entre l'atome ferrique et cette molécule d'eau, alors cette dernière se trouve dans l'environnement polaire de l'hème.

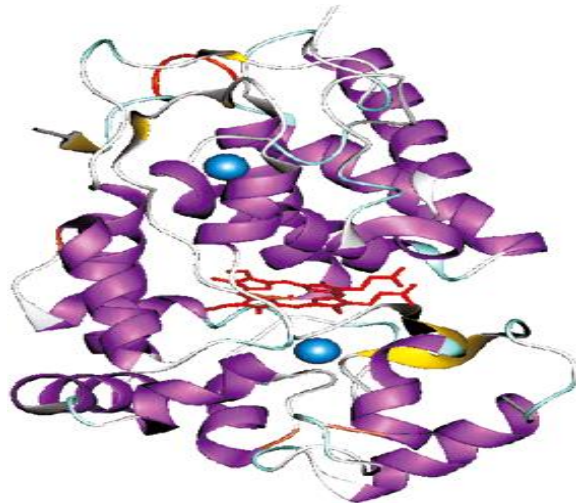
L'hème est localisée dans une poche réduite par des résidus d'acides aminés hydrophobes et d'His 42, His 170 et Arg 38 qui mènent leur groupement polaire vers le noyau hémique. Le fer tend à placer quatre liaisons de coordinance avec quatre atomes d'azote de la porphyrine, et deux autres avec le résidu His 170 (proximal). L'His 42 (distal), étant trop éloignée pour pouvoir établir la sixième coordinance (Dunfond, 1993).

- Structure primaire : une seule chaîne peptidique constituée d'environ 300 acides aminés. L'His 170, His 42 et Arg 38, sont les acides aminés conservés dans les différentes peroxydases. Pour la peroxydase C de horseradish (HRP-C), les résidus clés de la fixation de l'hème sont H42, His 170, Arg 38, Asn 70, Asp 247 et Phe 4, le monomère est constitué de 306 acides aminés (fig.2) (Veitch, 2004).

- Structure secondaire : la peroxydase présente trois hélices  $\alpha$  additionnelles. Le monomère typique consiste en dix segments d'hélice  $\alpha$  dont trois sont conservés dans les peroxydases des plantes.

- Structure tertiaire : la présence de deux domaines contenant chacun un atome de  $\text{Ca}^{2+}$  constitue le monomère typique des peroxydases. L'hème est inséré dans un creux formé par deux hélices antiparallèles.





**Figure 2 :** Structure cristalline de la peroxydase C de HRP (Veitch,2004).

*Un hème (colore en rouge) contient chacun un atome de calcium (représenté sous forme de sphère bleue). Les hélices  $\alpha$  et feuillet  $\beta$  sont représentées en mauve et jaune.*

- Structure quaternaire : le type de peroxydase est influencé par le nombre de sous unités.

Pour la HRP il y a six chaînes identiques ou monomères contenant chacune un groupement hème.

## 6. Site actif

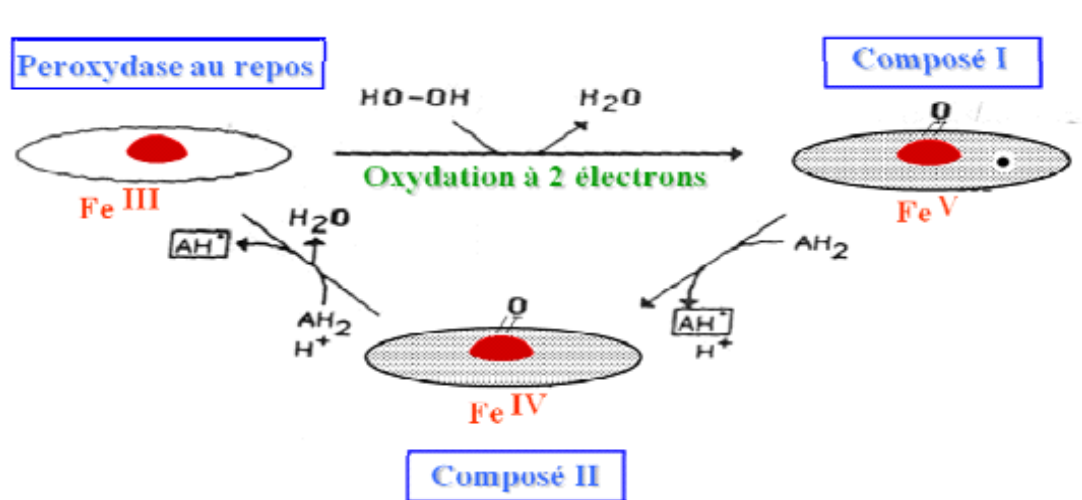
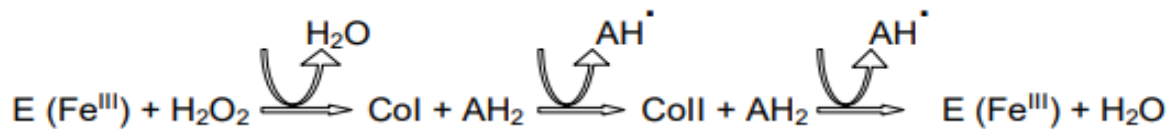
L'étude cristallographique de l'interaction de la peroxydase avec son substrat, réalisée par Henriksen et ses collaborateurs (1998), a permis de localiser le site actif chez les écrevisses. Le substrat fait face à l'hème, la partie non polaire fait face à la surface hydrophobe (Phe 179, Pro 141, Ala 140, Pro 139, Phe 68, Gly 69), et sa tranche polaire fait face à l'His 42, Arg 38 et est proche du fer. En effet, ces travaux ont montré que le remplacement de l'His 42 et l'Arg 38 a diminué l'affinité de l'enzyme pour le substrat. Par ailleurs, la substitution de l'Asn 70 par Val, entraîne une diminution de la vitesse maximale de plus de 90% (Henriksen *et al.*,1998).

## 7. Mécanisme d'action

Les peroxydases sont des enzymes michaeliennes composées d'une structure chimique « Hème » contenant du fer ionisé capable de se lier à l'oxygène catalysant l'oxydation d'une multitude de substrats (Dunford, 1999).

- En présence de substrat réducteur :

Les peroxydases catalysent l'oxydation de nombreux substrats de structures diverses bien qu'elles partagent le même cycle catalytique qui se déroule en trois étapes (figure 3) (Dunford, 1976).



**Figure 3** : Cycle catalytique des peroxydases (Pelmont,1988).

D'abord, le substrat oxydant (H<sub>2</sub>O<sub>2</sub>) est engagé dans la zone polaire de la crevasse, pour se lier aux résidus His 42, Arg 38 et à l'ion ferrique. Une fois dans le site actif, H<sub>2</sub>O<sub>2</sub> interagit avec des résidus d'acides aminés et de l'hème. Ceci mène à la formation d'une molécule de H<sub>2</sub>O et du composé I (Col), instable par l'oxydation de l'hème ferrique en oxyferryl (Fe<sup>4+</sup>=O<sup>+</sup>), H<sub>2</sub>O<sub>2</sub> est donc réduit et la sixième coordinance est bien occupée. Le Composé I n'est pas seulement un complexe enzyme-substrat classique, mais, précisément un intermédiaire enzymatique réactif d'un plus fort degré d'oxydation (+5) par rapport à celui de l'enzyme natif (+3) restitué (Dunford, 1999). Ainsi, le composé I est capable d'oxyder une gamme de substrats réducteurs par un mécanisme en deux étapes successives (Pelmont, 1998) :

## Synthèse bibliographique

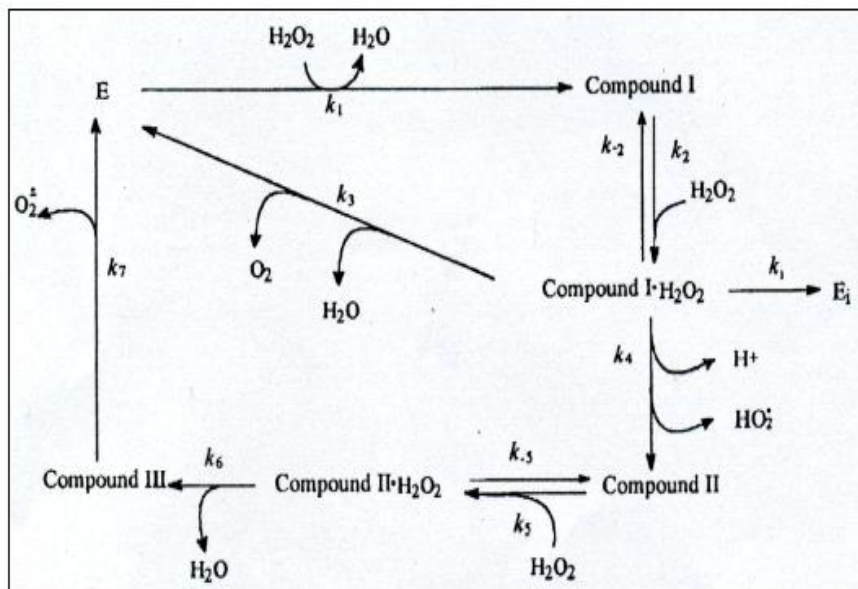
Le composé I (CoI) est réduit à la première étape, via une molécule de substrat donneur (AH) puis, en une deuxième forme enzymatique intermédiaire, le composé II (Coll) ayant un degré d'oxydation (+4).

Le composé (Coll) est réduit à la deuxième étape avec une autre molécule AH<sub>2</sub>, il perd sa structure radicalaire et retourne à sa forme native initiale E (Fe<sup>3+</sup>) avec libération d'une deuxième molécule de H<sub>2</sub>O.

- En absence de substrat réducteur :

Le H<sub>2</sub>O<sub>2</sub> peut jouer un triple rôle, en absence de substrats (AH<sub>2</sub>), un rôle de substrat oxydant dans la formation du Col, ou constituer un substrat réducteur de type donneur d'un électron et enfin agir en un agent désactivateur (Rodriguez *et al.*, 1997 ; Hernández-Ruiz *et al.*, 2001).

La figure 4 ci-dessous, décrit et englobe les deux voies catalytiques et aussi la voie d'inactivation identifiées. L'une des voies catalytiques est un processus direct de dismutation (à deux électrons) de H<sub>2</sub>O<sub>2</sub> en O<sub>2</sub>, semblable à celui de la catalase. L'autre étant un processus à deux étapes, qui met en jeu un électron par étape. Cette seconde voie catalytique génère la formation du composé III (Coll) et du radical superoxyde (O<sub>2</sub><sup>-</sup>) (Arnao *et al.*, 1990).



**Figure 4 :** Modèle de la réaction de HRP-C avec H<sub>2</sub>O<sub>2</sub> comme unique substrat (Arnao *et al.*, 1990 rapporté par Hernandez-Ruiz *et al.*, 2001)

*E* et *E<sub>i</sub>* sont respectivement l'enzyme native et l'enzyme inactive ; *Co I* *Co II* *Co III* sont des enzymes intermédiaires ; *Co IH<sub>2</sub>O<sub>2</sub>* et *Co IIIH<sub>2</sub>O<sub>2</sub>* sont les complexes entre les enzymes intermédiaires respectives et *H<sub>2</sub>O<sub>2</sub>*.

### 8. Rôle physiologique :

Après plusieurs études, les chercheurs ont démontré que la peroxydase participe à des fonctions importantes comme l'oxydation de l'acide indole-3-acétique (Hare, 1964) la formation de l'éthylène (Mattoo et Liebermann 1977) la polymérisation de la subérine et de la lignine (Espelie *et al.*, 1986 ; Ros Barcelo 1987) les réactions d'attachement des extensines et des pectines (Robert et Roland 1989) et l'oxydation des composés flavonoïques (Yamauchi *et al.*, 1997).

### 9. Applications biotechnologiques des peroxydases :

La peroxydase de raifort « horseradish peroxydase » (HRP), est l'enzyme la plus utilisée pour les applications analytiques. Toutefois, les peroxydases provenant d'autres sources apparentes peuvent être une bonne alternative en tant que substituts pour la HRP.

Les peroxydases spécialement celles des plantes sont utilisées pour produire des colorants à partir de phénols naturels. Ces colorants produits sont employés pour consolider la couleur des jus, des vins et des thés mais aussi des arômes (Egorov, 1995). Elles sont aussi utilisées comme indicateurs dans le blanchiment des légumes (Regalado *et al.*, 2004).

Dans l'industrie du papier, les peroxydases sont employées pour le blanchiment de la pulpe à papier et la biodégradation de la lignine (Wagner et Nicell, 2001 ; Regalado *et al.*, 2004).

En biologie moléculaire, les peroxydases peuvent présenter des marqueurs dans divers stress biotiques et abiotiques parce que la plupart des gènes sont exprimés soit de façon constitutive soit de façon inductive (Dicko, 2005).

Elles sont également utilisées comme biodétecteurs dans des kits d'identification et de diagnostic pour le dosage des sucres dans le sang et dans les urines ainsi que pour la quantification d'autres composés tel que l'alcool et la choline (Regalado *et al.*, 2004).

# Partie 2 : Analyse des protéines par bio-informatique

### 1. Notion de bio-informatique

Le terme de « bioinformatique » est apparu dans la littérature scientifique au début des années 1990. C'est un domaine multidisciplinaire qui regroupe la biologie (biochimie, biologie moléculaire, immunologie, virologie, etc.), l'informatique (algorithmes, programmes, réseaux, Internet, bases de données) et les mathématiques (modèles théoriques). Pour la plupart des membres de la communauté scientifique, la bioinformatique est l'étude de la façon dont la bioinformation est représentée et analysée dans les systèmes biologiques, en particulier les informations obtenues au niveau moléculaire (Sean et *al.*, 2014). Il y a une variété de définition, celle proposée par Andrade et Sander (1997) :

« *Bioinformatics is a science of recent creation that uses biological data, completed by computational methods, to derive new biological knowledge* ».

C'est la définition la plus moderne et large, qui présente que la bioinformatique ne se limite pas à l'analyse des séquences mais aussi des données concernant les marqueurs moléculaires (biomarqueurs) et les données phénotypiques, etc.

A partir de ces définitions, nous pouvons résumer les activités de cette discipline en trois points principaux (David, 2004) :

- Acquisition et organisation des données biologiques.
- Conception de logiciels pour l'analyse, la comparaison et la modélisation des données.
- Analyse des résultats obtenus par les logiciels.

### 2. Les banques de données :

La bioinformatique regroupe le stockage, l'organisation et la diffusion, qui sont les aspects importants de cette discipline. Toutes ces informations sont mises à disposition des chercheurs et utilisateurs du monde entier le plus rapidement possible (Beroud, 2010).

Ces banques de données sont divisées en deux groupes. D'une part nous avons, des banques dites « généralistes » de séquences nucléiques et protéiques, dans les quelles sont stockées les séquences provenant de tous les organismes. D'autre part, nous avons aussi des banques qui sont dites « spécialisées » sont créées pour des besoins spécifiques. Ils sont bien précis sans redondances et contiennent des données plus homogènes puisqu'elles traitent plutôt une thématique particulière ou un organisme donné comme par exemple la banque « PROSITE » (Falquet *et al.*, 2002) et « TFD » (Ghosh, 1993).

La principale banque de donnée protéique appelée Uniport (Universal Protein Resource) qui regroupe les données de plusieurs banques depuis quelques années. Elle est composée de deux sections (UniProt Consortium, 2007) :

- Swiss-Prot qui présente des données corrigées manuellement.
- TrEMBL qui présente des données non corrigées.

Les trois banques qui adoptent les séquences nucléotidiques sont :

- GenBank : banque diffusée par le NCBI « National Center for Biotechnology Information » (Benson *et al.*, 2016).
- EMBL : « European Molecular Biology Laboratory » (Stoesser *et al.*, 2003).
- DDBJ : « DNA Data Bank of Japan » développé par l'institut NIG (Tateno *et al.*, 2002)

### 3. Alignement multiple des séquences :

Parmi les techniques les plus utilisées dans l'analyse de séquences, l'alignement multiple des séquences d'ADN ou de protéines, qui jouent un rôle très important dans la biologie moléculaire. Il est utilisé pour un ensemble de tâches comme l'analyse phylogénétique, la fonction et la prédiction de la structure des macromolécules biologiques telles que l'ADN, l'ARN et les protéines. Un alignement est une écriture de deux séquences ou plus, l'une sous l'autre de façon à faire apparaître des identités, chaque alignement correspond à un score id% qui peut être calculé comme le pourcentage d'identité dès l'édition des séquences (Gilbert et Manalo, 2015).

Un alignement optimal pour une lecture de séquence nécessite un algorithme d'alignement qui tolère les correspondances imparfaites, où des variations génomiques peuvent se produire. De plus, l'algorithme doit être capable d'aligner des millions de lectures à une vitesse raisonnable (Bao *et al.*, 2014). Le tableau 1 représente certains algorithmes de l'alignement de séquences multiples.

**Tableaux 01** : Quelques algorithmes utilisés pour l’alignement de séquences multiples.

Algorithme	Alignement	Références
<b>Needleman-Wunsch</b>	Par paire, globale	(Needleman et Wunsch,1970)
<b>FASTA</b>	Recherches de base de données, par paires, locale	(Pearson et Lipman, 1988)
<b>Blast</b>	Recherches de base de données, par paire, local	Altschul et al., 1990)
<b>Clustal</b>	Multiple, global	(Higgins et Sharp, 1988)



**CHAPITRE 2 :**  
**MATÉRIEL ET MÉTHODES**

Dans ce travail, nous avons comparé les caractéristiques des peroxydases végétales par bio-informatique en réalisant un alignement des séquences multiple avec ses homologues structureaux.

### 1. Téléchargement et organisation des données :

Dans ce travail, nous avons choisi de travailler sur la peroxydase « horseradish », cette protéine présente dans la base de données UniProtKB/Swiss-Prot (<https://www.uniprot.org/>), qui contient des séquences protéiques annotées manuellement et obtenues à partir d'un large spectre d'organismes (figure 5) (Boutet *et al.*, 2016).



**Figure 5 :** Requête pour la base de données UniProtKB.

Les séquences protéiques sont écrites sous un format particulier appelé « FASTA », qui est le support de la plupart des moteurs de recherche et d'outils de traitement. Il permet de présenter les acides aminés par une succession de lettres codant tout ça pour donner une présentation d'une séquence protéique (Binz *et al.*, 2019).

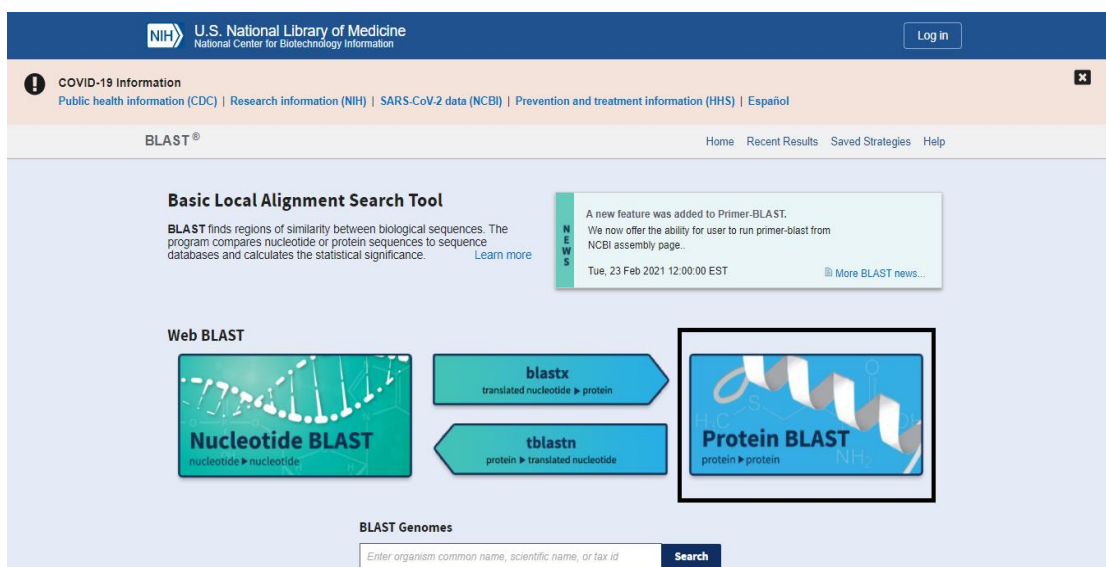
### 2. Recherche des homologues de la séquence :

Les homologues de la peroxydase Horseradish sont sélectionnés à partir du site web NCBI BLAST <https://blast.ncbi.nlm.nih.gov/Blast.cgi>, en utilisant l'algorithme blastx (figure 6) (Johnson *et al.*, 2008). On peut utiliser le même serveur pour la recherche des homologues à partir de PDB, une base pour les données structurales tridimensionnelles des protéines déterminées expérimentalement (Burley *et al.*, 2017).

Le serveur BLAST permet de rechercher dans des bases de données, les séquences répertoriées ayant des régions de similitude avec la séquence introduite par l'utilisateur. Cette

méthode est utilisée pour trouver des relations fonctionnelles entre les séquences et peut aider à identifier les membres d'une même famille (Gibson, 2004). Le principe de fonctionnement de BLAST, est décrit en trois étapes (Dardel et Képès, 2002) :

- La séquence analysée est d'abord découpée en segments quadruplets d'acides aminés chevauchants. BLAST constitue un dictionnaire contenant les segments répétitifs.
- Le balayage de la banque avec ce dictionnaire. BLAST tente d'étendre l'homologie en amont et en aval des quadruplets similaires trouvés ;
- Après extension de l'homologie, il évalue à partir du score obtenu la probabilité de similitude.



**Figure 6 :** Interface de l'outil BLAST

La figure 7 montre les étapes du fonctionnement du serveur BLAST. La séquence à analyser est introduite sous le format Fasta dans le casier (1) « Entrer la séquence de requête ». Une fois le choix des banques de données où se fait la recherche des homologues est réalisé (2), le programme est lancé en cliquant sur « BLAST » (4). L'algorithme blastP (3) est choisie automatiquement par BLAST une fois le type de séquence est indiqué.

The screenshot shows the BLAST web interface for a Standard Protein BLAST search. The form includes the following sections:

- Enter Query Sequence:** A text input field containing a protein sequence: `LNGNLSALVDFDLRPTIFDNKYVIVLEEQGLIQSDQELFSSPNATDTIPLVRSFANSTQITFFNAFVEAMDRMGNIPLTGTGGQIRLNCRVVNSNLLHDMVEVDFVSSM`. A 'Query subrange' section with 'From' and 'To' fields is also present.
- Choose Search Set:** A dropdown menu for the database is set to 'Non-redundant protein sequences (nr)'. Below it are fields for 'Organism' and 'Exclude' options.
- Program Selection:** A radio button selection for the algorithm, with 'blastp (protein-protein BLAST)' selected.
- BLAST Button:** A large blue button labeled 'BLAST' at the bottom left.

Figure 7 : Application du le serveur BLAST

### 3. Alignement et filtration des données :

Le logiciel MEGA (Molecular Evolutionary Genetics Analysis) est une application de bureau conçue pour l'analyse des séquences de gènes homologues provenant de familles multigéniques ou de différentes espèces l'accent sur l'inférence des relations évolutives et des modèles de l'évolution de l'ADN et des protéines (figure 8). En plus des outils d'analyse statistique des données, MEGA fournit de nombreuses facilités pratiques pour l'assemblage de données de séquence des fichiers ou des référentiels web, et il inclut des outils de présentation visuelle des résultats obtenus dans les arbres phylogénétiques interactifs et de matrices de distance évolutives. (Sudhir *et al.*, 2008).

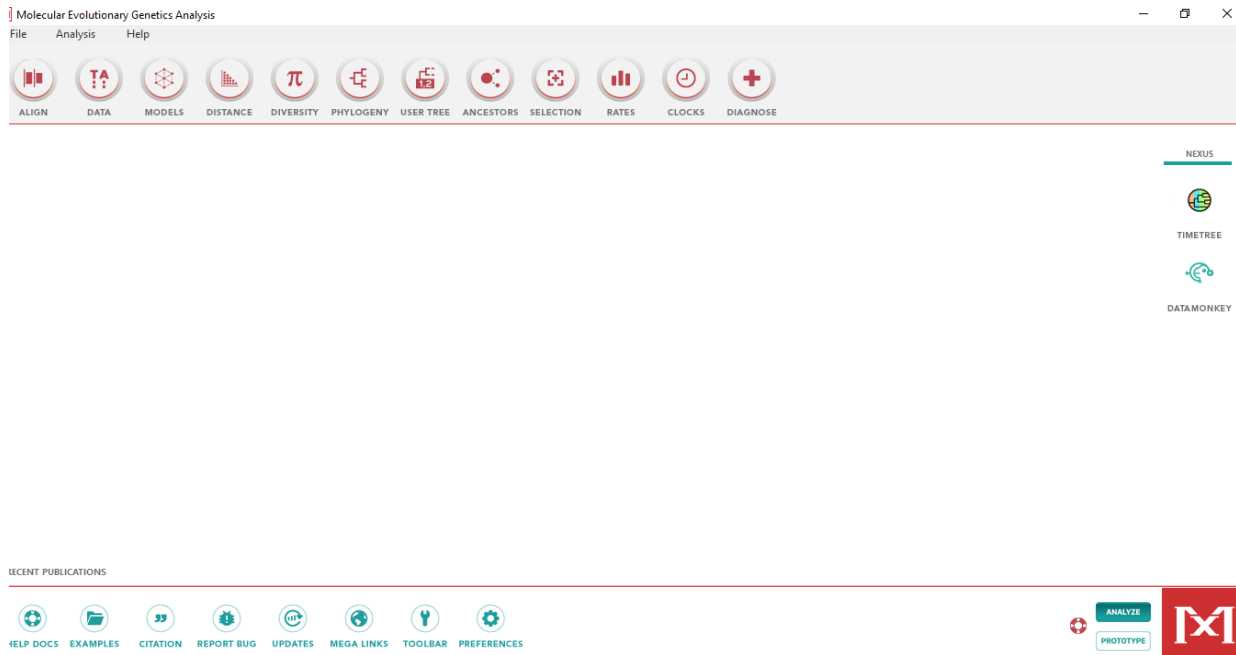


Figure 8 : Application du MEGA X

Les homologues obtenus à partir du serveur Blast, sont téléchargés sous leur format Fasta en cliquant sur Dowland (1) puis FASTA (aligned séquence) (2) (figure 9).

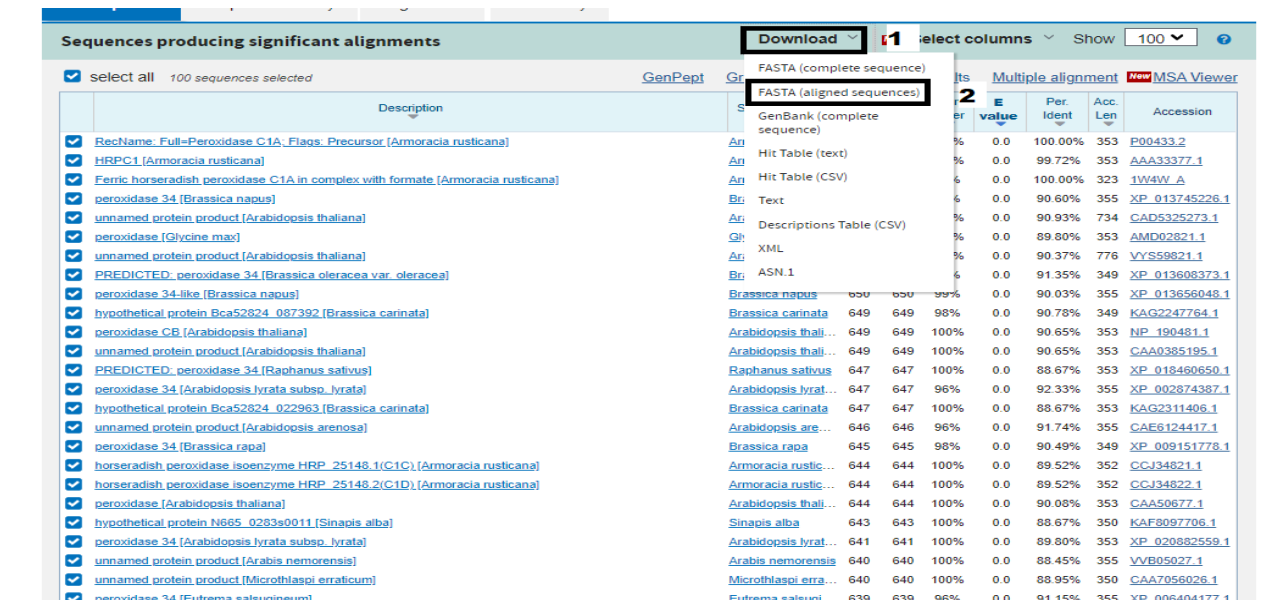


Figure 9 : Les étapes de téléchargement des séquences sous format FASTA

Les figures 10,11 et 12 montrent les étapes de fonctionnement sur MEGA X. Cliquer sur alignement (1), puis choisir créer un nouvel alignement (2) et enfin choisir protéine, coller les séquences téléchargées et cliquer sur alignement protéique (3).

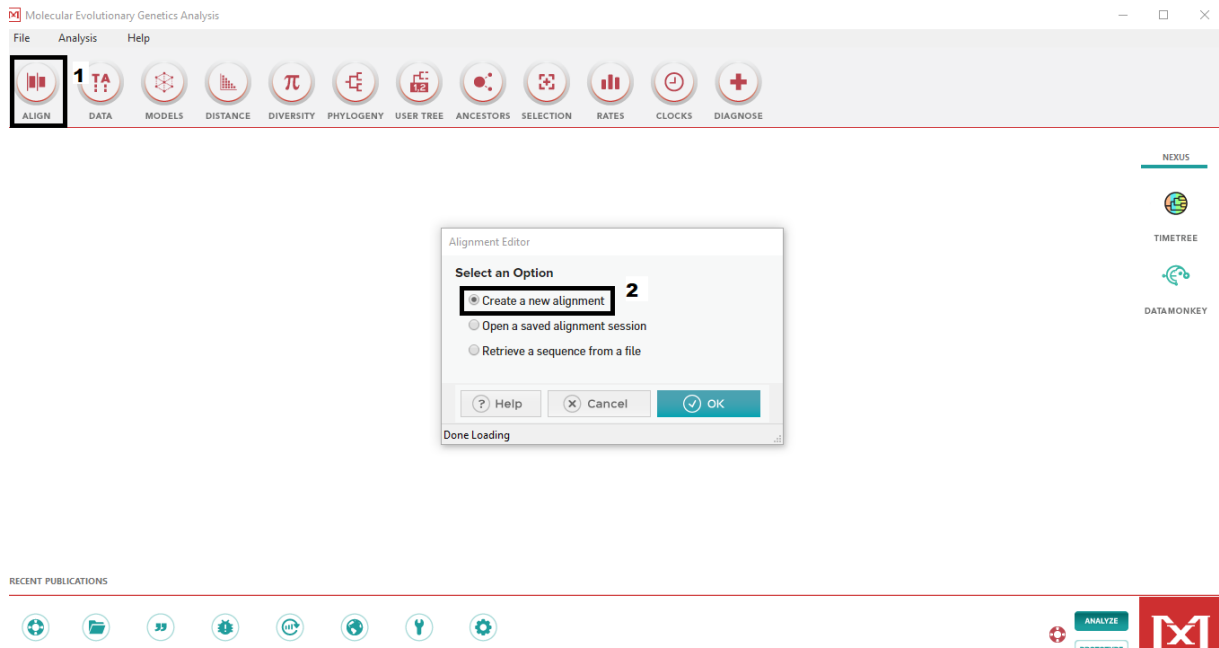


Figure10 : Les étapes de l'alignement

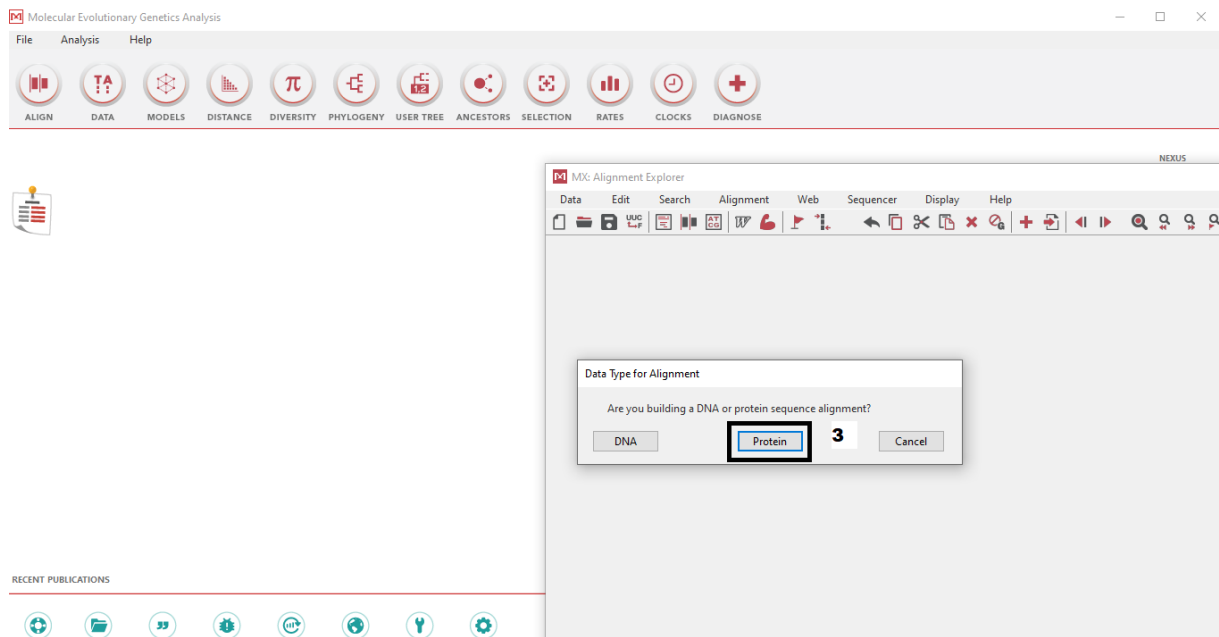


Figure 11 : Les étapes d'alignements.



**ESPrict 3.0**

SUBMIT
DISPLAY
RESULTS
DOC
BEG
ADV
EXP
LAYERS
SESSION
LOAD
SAVE
TIME
EXIT

ESPrict / ESPrict use popup windows to display results - please be sure to disable popup blockers before submitting a job  
 When publishing data resulting from usage of this server, please cite this reference article

**Aligned Sequences**

ALN file:  Aucun fichier n'a été sélectionné  
Exemple file • Tutorial

Number sequences:

**Secondary structure depiction**

Top secondary structures      Parameters

Input file:  Aucun fichier n'a été sélectionné  
Upload a file below OR click here: PDB

Chain ID:       Relative accessibility:

Sec. structure labels: a1:A1:2:3...  
 Sec. structure labels: a2:A2:3:4...  
 Sec. structure labels: a3:A3:4:5...

**Sequence similarities depiction parameters**

Similarity coloring scheme

NEquivalent     NMultiAln     NStrict     Rstar     RAN1250     BLOSUM62

Parameters

Global score:       Display consensus seq.:

**Defining groups**

Sequences to display

Figure 13 : Le profil du serveur web Esprict 3.0.



## CHAPITRE 3 :

# *RÉSULTATS ET DISCUSSION*

### 1. Le format FASTA de la séquence :

La figure 14 représente la séquence de la peroxydase de horseradish P00433 sous le format FASTA, à partir du serveur Blast.

```
>sp|P00433|PER1A_ARMRU Peroxidase C1A OS=Armoracia rusticana OX=3704 GN=PRXC1A PE=1 SV=2
MHFSSSTLFTCITLIPLVCLILHASLSDAQLTPTFYDNPCPNVSNIVRDTIVNELRSDP
RIAASILRLHFHDCFVNGCDASILLDNNTTSFRTEKDAFGNANSARGFPVIDRMKAAVES
CPRTVSCADLLTIAAQSVTLGGPSWRVPLGRRDSLQAFDLANANLPAPFFTLPLQLKD
SFRNVGLNRSSDLVALSGGHTFGKNQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCP
LNGNLSALVDFDLRPTIFDNKYYVNL EEQKGLIQSDQELFSSPNATDTIPLVRSFANST
QTFNFVFEAMDRMGNITPLTGTQGGIIRLNCRVVNSNSLLHDMVEVDFVSSM
```

**Figure14** : Séquence de la peroxydase au format FASTA.

### 2. Les homologues de la peroxydase horseradish :

Le serveur BLAST permet d'identifier toutes les protéines présentant une homologie structurale significative avec la peroxydase P00433. Nous avons obtenu comme résultats, 100 séquences homologues à celle de la peroxydase P00433 (figure 15).

Pour les enzymes dont les structures sont déterminées expérimentalement, les enzymes les plus homologues à la peroxydase P00433 sont la structure de la chaîne A, horseradish peroxidase C1A (Smith *et al.*, 1990), la horseradish C1 (Fujiyama *et al.*, 1988), peroxidase 34 *Brassic anapuset* et peroxidase 34 *Brassica oleracea var. oleracea* (tableau 2).

**Tableau 2** : Les enzymes les plus homologues à la peroxydase P00433.

Description	Identité	Accession
<b>ChainA,horseradish peroxidase c1a</b>	100%	1W4W_A
<b>Hrpc1 [armoracia rusticana]</b>	99.72%	AAA33377.1
<b>Peroxidase34 [brassica napus]</b>	92.33%	XP_013745226.1
<b>Peroxidase34[brassicaoleracea var. oleracea]</b>	91.35%	XP_013608373.1

sélectionner tout 100 séquences sélectionnées		GenPept	graphiama	Arborescence des distances des résultats	Alignement multiple	revoir	Visionneuse MSA		
	description	Nom scientifique	Score maximal	Total Score	Couverture de requête	Valeur	Ident par	Acc. Len	Accession
<input checked="" type="checkbox"/>	Reclhame_Full_-_Peroxidase_C1A_Drapeauc_Precurseur (Armoracia rusticana)	Armoracia rusti...	721	721	100%	0.0	100 000	353	NC_P00433.2
<input checked="" type="checkbox"/>	HRP_C1 (Armoracia rusticana)	Armoracia rusti...	717	717	100%	0.0	89 724	353	CA033377.1
<input checked="" type="checkbox"/>	Peroxydase ferrique de raifort C1A en complexe avec formate (Armoracia rusticana)	Armoracia rusti...	660	660	91%	0.0	100 000	323	1W4W_A
<input checked="" type="checkbox"/>	peroxydase 34 (Brassica napus)	Brassica napus	653	653	99%	0.0	90 606	355	XP_013745226.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis thaliana)	Arabidopsis tha...	653	653	100%	0.0	90 935	734	CA05325273.1
<input checked="" type="checkbox"/>	peroxydase (Glycine max)	Glycine max	652	652	100%	0.0	89 800	353	AM002821.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis thaliana)	Arabidopsis tha...	651	651	100%	0.0	90 374	776	VYSS682.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34 (Brassica oleracea var. oleracea)	Brassica olerac...	651	651	98%	0.0	91 355	349	XP_013608373.1
<input checked="" type="checkbox"/>	peroxydase 34-like (Brassica napus)	Brassica napus	650	650	99%	0.0	90 035	355	XP_013656048.1
<input checked="" type="checkbox"/>	Bca52824_087392 hypothétique de protéines (Brassica carinata)	Brassica carinata	649	649	98%	0.0	90 785	349	KAG2247764.1
<input checked="" type="checkbox"/>	peroxydase CH (Arabidopsis thaliana)	Arabidopsis tha...	649	649	100%	0.0	90 655	353	NP_190481.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis thaliana)	Arabidopsis tha...	649	649	100%	0.0	90 655	353	CA00035105.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34 (Raphanus sativus)	Raphanus sativus	647	647	100%	0.0	88 675	353	XP_018460550.1
<input checked="" type="checkbox"/>	peroxydase 34 (Arabidopsis lyrata subsp. lyrata)	Arabidopsis lyr...	647	647	96%	0.0	92 335	355	XP_002874387.1
<input checked="" type="checkbox"/>	Bca52824_022963 hypothétique de protéines (Brassica carinata)	Brassica carinata	647	647	100%	0.0	88 675	353	KAG2311406.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis arenosa)	Arabidopsis are...	646	646	96%	0.0	91 745	355	CAE6124417.1
<input checked="" type="checkbox"/>	peroxydase 34 (Brassica rapa)	Brassica rapa	645	645	98%	0.0	90 495	349	XP_009151728.1
<input checked="" type="checkbox"/>	peroxydase de raifort isoenzyme HRP_25148_1(C1C) (Armoracia rusticana)	Armoracia rusti...	24*	24*	100%	0.0	89 525	352	CGJ3482.1
<input checked="" type="checkbox"/>	peroxydase de raifort isoenzyme HRP_25148_2(C1D) (Armoracia rusticana)	Armoracia rusti...	24*	24*	100%	0.0	89 525	352	CGJ3482.1
<input checked="" type="checkbox"/>	peroxydase (Arabidopsis thaliana)	Arabidopsis tha...	24*	24*	100%	0.0	90 085	353	CAA50677.1
<input checked="" type="checkbox"/>	N665_0283e0011 hypothétique de protéines (Sinapis alba)	Sinapis alba	643	643	100%	0.0	88 675	350	KAF6097706.1
<input checked="" type="checkbox"/>	peroxydase 34 (Arabidopsis lyrata subsp. lyrata)	Arabidopsis lyr...	641	641	100%	0.0	89 800	353	XP_020892959.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis thaliana)	Arabidopsis tha...	640	640	100%	0.0	88 445	355	VX089802.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Microthlasia arvensis)	Microthlasia arr...	640	640	100%	0.0	88 955	350	CAA7056026.1
<input checked="" type="checkbox"/>	peroxydase 34 (Eutrema sativum)	Eutrema sativum	639	639	96%	0.0	91 155	355	XP_006404177.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis arenosa)	Arabidopsis are...	638	638	100%	0.0	90 085	353	CAE6075570.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Camelina sativa)	Camelina sativa	637	637	97%	0.0	90 415	356	XP_019085348.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Camelina sativa)	Camelina sativa	636	636	97%	0.0	90 415	356	XP_019103292.1
<input checked="" type="checkbox"/>	YADP2_HBP (construction synthétique)	construction sy...	636	636	87%	0.0	90 685	459	AV09214.1
<input checked="" type="checkbox"/>	Reclhame_Full_-_Peroxidase_C1B_Drapeauc_Precurseur (Armoracia rusticana)	Armoracia rusti...	635	635	100%	0.0	88 955	351	P15232.1
<input checked="" type="checkbox"/>	ATPCB/ATPRX34/PERX34 (Arabidopsis lyrata subsp. lyrata)	Arabidopsis lyr...	634	634	100%	0.0	89 275	354	EFH52207.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34 (Camelina sativa)	Camelina sativa	634	634	96%	0.0	91 155	355	XP_010503558.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Camelina sativa)	Camelina sativa	633	633	96%	0.0	88 645	345	XP_019114116.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34 (Camelina sativa)	Camelina sativa	634	634	96%	0.0	91 155	355	XP_010503558.1
<input checked="" type="checkbox"/>	erHRP (construction synthétique)	construction sy...	632	632	88%	0.0	90 045	345	ANS10151.1
<input checked="" type="checkbox"/>	sous-unité alpha-1 (E.C. 1.1.1.1.1.7) de la peroxydase de raifort synthétique [construction synthétique]	construction sy...	630	630	87%	0.0	100 005	309	AA72223.1
<input checked="" type="checkbox"/>	Structure de la peroxydase de raifort C1a Composé II (Armoracia rusticana)	Armoracia rusti...	630	630	87%	0.0	100 005	308	IHS5_A
<input checked="" type="checkbox"/>	PEROXYDASE DE RAIFORT RECOMBINANTE C1A ALA140GLY (Armoracia rusticana)	Armoracia rusti...	629	629	87%	0.0	90 685	309	IQWV_A
<input checked="" type="checkbox"/>	Complexe de peroxydase de raifort recombinant avec de l'acide benzhydroxamique (Armoracia rusti...)	Armoracia rusti...	629	629	86%	0.0	100 005	308	2ATJ_A
<input checked="" type="checkbox"/>	PEROXYDASE DE RAIFORT RECOMBINANTE C1A THR171SER EN COMPLEXE AVEC DE L'ACI...	Armoracia rusti...	629	629	87%	0.0	90 685	308	IQWZ_A
<input checked="" type="checkbox"/>	Peroxydase recombinante de raifort C1A ALA170GLN (Armoracia rusticana)	Armoracia rusti...	629	629	87%	0.0	90 685	309	IQWQ_A
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Raphanus sativus)	Raphanus sativus	629	629	100%	0.0	86 125	353	XP_018463254.1
<input checked="" type="checkbox"/>	PEROXYDASE DE RAIFORT RECOMBINANTE C1A PHE221MET (Armoracia rusticana)	Armoracia rusti...	628	628	87%	0.0	90 685	309	IQWT_A
<input checked="" type="checkbox"/>	MUTANT DISTAL DE POCHE D'HÈME (H42E) DE PEROXYDASE DE RAIFORT RECOMBINANTE...	Armoracia rusti...	628	628	87%	0.0	90 685	309	4ATJ_A
<input checked="" type="checkbox"/>	ATPCA/ATPRX33/PRX33 (Arabidopsis lyrata subsp. lyrata)	Arabidopsis lyr...	628	628	100%	0.0	86 695	349	EFH52203.1
<input checked="" type="checkbox"/>	Complexe de raifort recombinant peroxydase Phe209Ser avec de l'acide benzhydroxamique (Arm...	Armoracia rusti...	627	627	87%	0.0	90 685	309	IQXZ_A
<input checked="" type="checkbox"/>	peroxydase CA (Arabidopsis thaliana)	Arabidopsis tha...	627	627	96%	0.0	89 385	354	NP_190480.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis thaliana)	Arabidopsis tha...	627	627	96%	0.0	89 385	354	VYSS8819.1
<input checked="" type="checkbox"/>	PEROXYDASE DE RAIFORT RECOMBINANTE C1A (Armoracia rusticana)	Armoracia rusti...	627	627	86%	0.0	100 005	306	IATJ_A
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Camelina sativa)	Camelina sativa	626	626	96%	0.0	89 095	355	XP_010503561.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Camelina sativa)	Camelina sativa	626	626	96%	0.0	89 095	355	XP_010515256.1
<input checked="" type="checkbox"/>	peroxydase 33 (Capsella rubella)	Capsella rubéole	626	626	96%	0.0	88 795	354	XP_006291394.1
<input checked="" type="checkbox"/>	PRÉDICTED... peroxydase 33 (Camelina sativa)	Camelina sativa	625	625	100%	0.0	85 805	354	XP_010515258.1
<input checked="" type="checkbox"/>	Mutant distal de poche d'hème (R36S/H42E) de la peroxydase recombinante de raifort C. HRP C1 (...)	Armoracia rusti...	625	625	87%	0.0	90 355	308	IK2M_A
<input checked="" type="checkbox"/>	produit protéique sans nom (Microthlasia arvensis)	Microthlasia arr...	625	625	100%	0.0	86 725	354	CAA7056025.1
<input checked="" type="checkbox"/>	Peroxydase de radis de cheval... mutant S167Y (Armoracia rusticana)	Armoracia rusti...	625	625	86%	0.0	90 675	306	ZYJL_A
<input checked="" type="checkbox"/>	PRXCA (Arabidopsis thaliana)	Arabidopsis tha...	624	624	96%	0.0	89 095	354	PAF01267.1
<input checked="" type="checkbox"/>	Reclhame_Full_-_Peroxidase_C1C_Drapeauc_Precurseur (Armoracia rusticana)	Armoracia rusti...	623	623	94%	0.0	90 665	332	P15233.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Raphanus sativus)	Raphanus sativus	623	623	96%	0.0	88 205	354	XP_018441188.1
<input checked="" type="checkbox"/>	peroxydase 34 (Capsella rubella)	Capsella rubéole	622	622	96%	0.0	89 685	357	XP_006290443.1
<input checked="" type="checkbox"/>	Protéine de la superfamille de la peroxydase (Arabidopsis thaliana)	Arabidopsis tha...	622	622	100%	0.0	86 125	352	NP_890652.1
<input checked="" type="checkbox"/>	peroxydase 33 (Capsella rubella)	Capsella rubéole	622	622	96%	0.0	85 505	353	XP_006291396.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis thaliana)	Arabidopsis tha...	621	621	100%	0.0	86 125	352	CAA0363274.1
<input checked="" type="checkbox"/>	peroxydase 34 (Brassica rapa)	Brassica rapa	621	621	96%	0.0	87 395	354	XP_009149782.1
<input checked="" type="checkbox"/>	hypothetical protein Bca52824_000744 (Brassica carinata)	Brassica carinata	620	620	100%	0.0	84 705	351	KAG232956.1
<input checked="" type="checkbox"/>	peroxydase (Arabidopsis thaliana)	Arabidopsis tha...	619	619	100%	0.0	85 845	352	AAM64538.1
<input checked="" type="checkbox"/>	PRÉDIT... peroxydase 34-like (Camelina sativa)	Camelina sativa	619	619	96%	0.0	87 615	355	XP_010426415.1
<input checked="" type="checkbox"/>	peroxydase 32 (Capsella rubella)	Capsella rubéole	617	617	100%	0.0	85 275	352	XP_006279456.2
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 32 (Camelina sativa)	Camelina sativa	617	617	100%	0.0	86 125	372	XP_010495675.1
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 33 (Camelina sativa)	Camelina sativa	617	617	96%	0.0	87 615	354	XP_010426414.1
<input checked="" type="checkbox"/>	peroxydase 32 (Capsella rubella)	Capsella rubéole	617	617	100%	0.0	85 275	352	XP_006279456.2
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 32 (Camelina sativa)	Camelina sativa	617	617	100%	0.0	86 125	372	XP_010495675.1
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 33 (Camelina sativa)	Camelina sativa	617	617	96%	0.0	87 615	354	XP_010426414.1
<input checked="" type="checkbox"/>	peroxydase 34 (Capsella rubella)	Capsella rubéole	616	616	100%	0.0	84 705	353	XP_006290316.1
<input checked="" type="checkbox"/>	CARUB_v10016503mg hypothétique de protéines (Capsella rubella)	Capsella rubéole	616	616	100%	0.0	85 275	372	EOA1235.1
<input checked="" type="checkbox"/>	hypothetical protein N665_0166a0049 (Sinapis alba)	Sinapis alba	616	616	96%	0.0	87 325	354	KAF8104865.1
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 32 (Camelina sativa)	Camelina sativa	615	615	100%	0.0	85 555	352	XP_010425797.1
<input checked="" type="checkbox"/>	peroxydase 32 (Eutrema sativum)	Eutrema sativum	615	615	100%	0.0	85 555	357	XP_006409640.1
<input checked="" type="checkbox"/>	peroxydase 32 (Eutrema halophilum)	Eutrema haloph...	614	614	100%	0.0	85 555	353	AB093458.1
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 32 (Camelina sativa)	Camelina sativa	613	613	100%	0.0	85 845	372	XP_010514693.2
<input checked="" type="checkbox"/>	précurseur de type peroxydase 34 (Brassica napus)	Brassica napus	612	612	96%	0.0	87 025	354	NP_001302880.1
<input checked="" type="checkbox"/>	F2Q70_0000857 hypothétique de protéines (Brassica cretica)	Brassica cretica	610	610	100%	0.0	85 315	354	KAF2575876.1
<input checked="" type="checkbox"/>	AALP_AA30314000 hypothétique de protéines (Arabis alpeina)	Arabis alpeina	607	607	96%	0.0	86 185	349	KFK39978.1
<input checked="" type="checkbox"/>	hypothetical protein Bca52824_034502 (Brassica carinata)	Brassica carinata	607	607	100%	0.0	84 995	353	KAG2286090.1
<input checked="" type="checkbox"/>	F2Q69_00040348 hypothétique de protéines (Brassica cretica)	Brassica cretica	607	607	100%	0.0	84 755	354	KAF3503848.1
<input checked="" type="checkbox"/>	hypothetical protein N665_0071a0033 (Sinapis alba)	Sinapis alba	607	607	100%	0.0	83 625	354	KAF8111894.1
<input checked="" type="checkbox"/>	IQI04_033826 hypothétique de protéines (Brassica rapa subsp. Intolucaris)	Brassica rapa s...	606	606	100%	0.0	85 315	374	KAG5382356.1
<input checked="" type="checkbox"/>	peroxydase 32 (Brassica rapa)	Brassica rapa	605	605	100%	0.0	85 035	374	XP_009123209.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Arabidopsis arenosa)	Arabidopsis are...	604	604	93%	0.0	85 185	352	CAE6074342.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Brassica rapa)	Brassica rapa	604	604	100%	0.0	85 035	374	VDC58081.1
<input checked="" type="checkbox"/>	peroxydase 32 (Arabidopsis lyrata subsp. lyrata)	Arabidopsis lyr...	604	604	93%	0.0	88 185	352	XP_002877238.1
<input checked="" type="checkbox"/>	hypothetical protein N665_0082a0125 (Sinapis alba)	Sinapis alba	603	603	100%	0.0	84 425	353	KAF8112782.1
<input checked="" type="checkbox"/>	PREDICTED... peroxydase 33-like (Camelina sativa)	Camelina sativa	603	603	96%	0.0	86 145	354	XP_019100238.1
<input checked="" type="checkbox"/>	Raifort peroxydase isoenzyme HRP_1805 (Armoracia rusticana)	Armoracia rusti...	603	603	95%	0.0	86 395	354	CGJ34828.1
<input checked="" type="checkbox"/>	produit protéique sans nom (Brassica napus)	Brassica napus	602	602	100%	0.0	85 035	354	CAF2037385.1
<input checked="" type="checkbox"/>	DY000_02011874 hypothétique de protéines (Brassica cretica)	Brassica cretica	602	602	100				

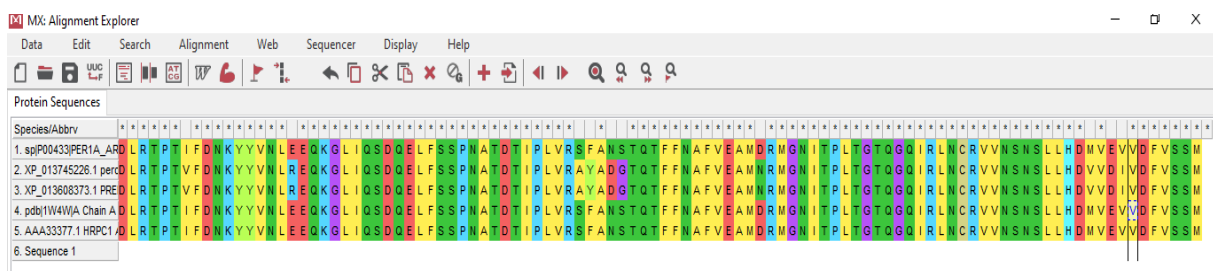
### 3. Résultats d'alignement :

La figure 16 montre l'alignement de séquence multiple de la peroxydase P00433 par rapport à ses homologues, déterminé à l'aide du serveur MEGA X. Les différentes couleurs représentent les 20 acides aminés.



**Figure 16 :** Résultats d'alignement de 100 homologues de la peroxydase P00433 par MEGA X

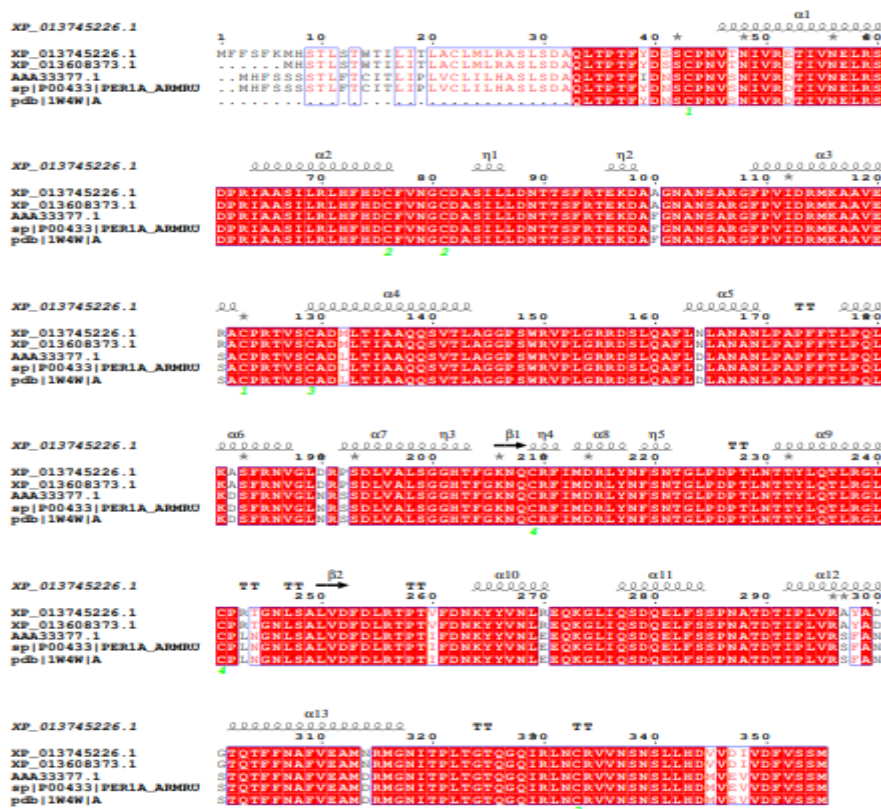
Pour comparer l'alignement des séquences, on n'a choisi seulement les 4 homologues de la peroxydase P00433 qui sont présentés dans le tableau 1 car elles ont un pourcentage d'identité très élevé (entre 91% et 100 %), (Hrpc1[armoracia rusticana], Peroxidase34 [brassica napus], Peroxidase34 [brassica oleracea var. oleracea]). Les couleurs pour les différents acides aminés et les étoile représentent les sites conservés (figure 17).



**Figure 17 :** Résultats d'alignement des 4 homologues de la peroxydase P00433 par MEGA X.

#### 4. Analyse de l'alignement multiple des séquences :

L'analyse des résultats par le logiciel ESPrift 3.0, permet de visualiser les séquences alignées rapidement (figure18). On a une séquence primaire identique de 91% jusqu'à 100% dans les quatre homologues par rapport à la peroxydase P00433. Les éléments de structure secondaire prédits de XP\_01374522.1 sont présentés en haute de l'alignement (les hélices alpha, les feuillets beta, et les tours stricts sont noté η, α). Les acides aminés His 42, His 170 et Arg 38 sont les plus conservés dans le site actif des différentes peroxydases et qui appartiennent à l'hélice alpha. Les résidus clés de la fixation de substrat His42, His170, Arg 38, Asn 70, Asp 247 et Phe 4. Les résidus Cystéine impliqués dans la liaison disulfure interchaîne sont indiqués par des nombres verts.



**Figure 18 :** Alignement de séquences multiples de la peroxydase P00433 avec ses homologues structuraux.

*Les résidus identiques sont marqués avec un fond rouge et les résidus hautement conservés sont représentés par une police rouge*



# CONCLUSION

La peroxydase est une enzyme qui catalyse l'oxydation de molécules donneuses de protons en présence du peroxyde d'hydrogène, qui sera réduit en molécules d'eau. Ces enzymes existent chez les animaux, les micro-organismes et les végétaux. Ces derniers sont regroupés dans la classe III et ont fait l'objectif de notre étude.

Le travail présenté dans ce mémoire porte sur l'étude de peroxydase végétale et de ces homologues à partir d'un alignement de séquences multiples.

A l'aide du programme Blaste, on a obtenu les 100 homologues de la peroxydase végétale HRP qui sont constitués d'environ 300 résidus d'acides aminés. Les résultats de l'analyse des séquences multiples réalisés à l'aide du programme MEGA X et ESPript 3.0, a montré que les acides aminés His 170, His 42 et Arg 38 sont les plus conservés autour de site actif dans les différentes peroxydases. Les résidus de cystéine impliqués dans la liaison disulfure inter-chaine sont hautement conservés.

Enfin, grâce à notre étude, nous avons réussi à entamer le domaine de la bio-informatique qui devient un procédé fiable pour la biologie afin d'acquisition de nouvelles techniques d'alignement.

En perspective, nous souhaitons approfondir notre étude en effectuant des alignements d'autres séquences protéiques de peroxydase à partir d'autres sources végétales. Aussi en testant d'autres programmes d'alignement.



RÉFÉRENCES

BIBLIOGRAPHIQUES

## Référence Bibliographique

---

### Référence bibliographique :

- Agostini E., Hernardez J., Arnao M.B., Milard S.B., Tigier H.A. and Acosta M. (2002). A peroxydase isoenzyme secreted by turnip (*Brassica napus*) hairy-root cultures : inactivation by hydrogen peroxide application in diagnostic kits. *Biotechnol. Applied Biochemistry* 35 :1-7.
- Andrade et Sander. (1997). From genome data to biological knowledge, *Current Opinion in Biotechnology*. *Journal of Bioinformatics*.
- Beroud, C. (2010). Bases de données et outils bioinformatiques utiles en génétique. Collège National des Enseignants et Praticiens de Génétique Médicale, Univ. Médicale Virtuelle Francophone, 3-6.
- Céline, B.-A. (2012). Introduction à la bioinformatique. Université Claude Bernard, Lyon1, Laboratoire de Biométrie et Biologie Evolutive (UMR 5558).
- David, W.-M. (2004). *Bioinformatics: Sequence and genome analysis*. Cold Spring Harbor Laboratory Press.
- Dunford H.B. (1999). *Hèmeperoxidases*. Wiley-VCH (H : Brian Dunford Ed.). New York :507.
- Delannoy, É., Marmey, P., Penel, C., & Nicole, M. (2004). Les peroxydases végétales de classe III. *Acta botanica gallica*, 151 :353-380.
- Egorov A. (1995). Peroxidase biotechnology and application. International workshop peroxidase biotechnology and application. Oral abstracts : parti. Moscow, Russia 1995.
- Espelie K.E., Kolattukudy P.E. (1985). Purification and characterization of an abscisic acid-inducible anionic peroxidase associated with suberization in potato (*Solanum tuberosum*). *Arch. Biochem. Biophys.* 240,539-545.
- Espelie K.E., Franceschi V.R., Kolattukudy P.E. (1986). Immunocytochemical localization and time course of appearance of an anionic peroxidase associated with suberization in wound-healing potato tuber tissue. *Plant Physiol.* 81,487-492.
- Everse, J., Everse, K. E., Grisham, M. B. (1990). *Peroxidases in chemistry and biology*, CRC Press, Boca Raton, Fla. vol. 1.

## *Référence Bibliographique*

---

- Gaspar T., Penel C., Thorpe T., Greppin H. (1982). Peroxidase 1970-1980. A survey of their Biochemical and Physiological Roles in Higher Plants. Uni v. Genève, Switserland, 324.
- Ghosh, D. (1993). Status of the transcription factors database (TFD). *Nucleic Acids Research*, 21(13), 3117-3118.
- Galston A.W., Dalberg L.Y. (1954). The adaptive formation and physiological significance of indoleacetic acid oxidase. *Amer. J. Bot*, 40, 373-380.
- Henriksen A., D.J.Schuller K.Meno K.G.Welinder A.T.Smith & M. Gajhede. (1998) . Structural interactions between horseradish peroxidase C and the substrate benzhydroxamic acid determined by X-ray crystallography. *Biochemistry*, 37, 8054-8060.
- Higuchi T. (1990). *Lignin biochemistry: Biosynthesis and biodegradation*. Wood Sei. Technol, 24, 23-63.
- Kalın, R., Atasever, A., Özdemir, H. (2014). The single-step purification of peroxidase by 4-aminobenzohydrazide from Turkish black radish (*Raphanus sativus* L) and turnip (*Brassica rapa* L) roots. *Food Chemistry*, 150, 335-340.
- Kohavi, R. et John, G.- H. (1997). Wrappers for feature subset selection. *Artif. Intell.*, 97, 273-324.
- Kumar, A., Mukhia, S., Kumar, N., Acharya, V., Kumar, S., & Kumar, R. (2020). A Broad Temperature Active Lipase Purified From a Psychrotrophic Bacterium of Sikkim Himalaya With Potential Application in Detergent Formulation. *Frontiers in bioengineering and biotechnology*, 8, 642.
- Mc Lellan KM, Robinson DS. (1987). The heat stability of purified spring cabbage peroxidase isoenzyme. *Food chem.* 26, 97-197.
- Paul, G.-H. et Teresa, K.-A. (2005). *Bioinformatics and molecular evolution*. Blackwell Science Ltd, Blackwell Publishing company.
- Regalado C., Garcia-Almendarez B.E and Duarte-Vazquez M.A. (2004). Biotechnologie applications of peroxidases. *Biochemestrey Rev.* 3 :243-256.

## Référence Bibliographique

---

- Rompel, A., Albers, M., NaserI, J, I., Gerdemann, C., Büldt-Karentzopoulos, K., Jasper, B., and Krebs, B. (2007). Purification, cloning and characterization of a novel peroxidase isozyme from sweet potatoes (*Ipomoea batatas*). *Biochimica et Biophysica Acta*, 1774, (11), 1422-1430.
- Ros Barcelo A. (1995). Peroxidase and not laccase is the enzyme responsible for cell wall lignification in the secondary thickening of xylem vessels in *Lupinus*. *Protoplasma* 186 ,41-44.
- Robert D., Roland J.C. (1989). *Biologie végétale*. Tome 1. Organisation cellulaire. Doin (Eds). page : 265.
- Sharma, A., Meena, K. R., & Kanwar, S. S. (2018). Molecular characterization and bioinformatics studies of a lipase from *Bacillus thermoamylovorans* BHK67. *International journal of biological macromolecules*, 107, 2131-2140.
- Sean, D.-M. , Jessica, D.-T. et Russ, B.- A. (2014). *Bioinformatics*. E.H. Shortliffe, J.J. Cimino (eds.), *Biomedical Informatics*, 695 DOI 10.1007/978-1-4471-4474-8\_24, Springer- Verlag London.
- Sicioglu, M., Gülçin, Çankaya, M., Atasever, A., ehitolu, M, H., Budak Kaya, H., et al. (2010). Purification and characterization of peroxidase from Turkish black radish (*Raphanus sativus* L.). *Journal of Medicinal Plants Research* , 4, 1187-1196.
- Sreenivasulu, N., Ramanjulu, S., Ramachandra-Kini, K., Prakash, H, S., Shekar-Shetty, H., Savithri, H, S., et al. (1999). Total peroxidase activity and peroxidase isoforms as modified by salt stress in two cultivars of fox-tail millet with differential salt tolerance. *Plant Science*, 141, 1 – 9.
- Skhanov I, Castillo J, Areza J.C, Galev I. (2000). Purification and stability of peroxidase of ricin oil plant (*Elaeis guineensis*). *Bioseparation*. 9(3), 125-132.
- Velikova, V., Pinelli, P., Loreto, F .(2005). Consequences of inhibition of isoprenes synthesis in *Phragmites australis* leaves exposed to elevated temperatures. *Agriculture, Ecosystems & Environment*, 106, 209-217.
- Veitch N.C. (2004). Horseradish peroxidase : a modern view of a classic enzyme. *Phytochemistry*, 65, 249-259.

## *Référence Bibliographique*

---

- Welinder. (1992). Comparison of structure and activities of peroxidases from *Coprinus cinereus*, *Coprinus macrorhizus* and *Arthromyces ramosus*. *Biochim. Biophys. Acta-Protein structure and Molecular Enzymology*, 1120, 248-256.
- Yamasaki H., Sakihama Y., Ikehara N. (1997). Flavonoid-peroxidase reaction as a detoxification mechanism of plant cells against H<sub>2</sub>O<sub>2</sub>. *Plant Physiol.* 115: 1405-1412.
- Yamauchi N., Xia X.M., Hashinaga F. (1997). Involvement of flavonoid oxidation with chlorophyll degradation by peroxidase in wasesatsuna mandarin fruits. *J. Jpn. Soc. Hort. Sei.*, 66, 283-288.
- Yemenicio, lu, A., Özkan, M., Cemer, lu, B. (1998). Partial purification and thermal characterization of peroxidase from okra (*Hibiscus esculentum*). *Journal of Agriculture Food Chemistry*, 46, 4158-4163.
- Zamocky, M., Janecek, S., Koller, F. (2000). Common phylogeny of catalase-peroxidases and ascorbate peroxidases. *Gene*, 256, 169-182.
- Zheng B., Li, J., Liu, C., Li, B., Yuan H and Yang, J. (2012). Identification and molecular characterization of a novel DyP-Type peroxidase from *Pseudomonas aeruginosa* PKE117. *Applied Biochemistry and Biotechnology*, 166, (3), 774-785.

## Résumé :

Les peroxydases sont des enzymes qui catalysent la réaction d'oxydo-réduction par un mécanisme de radical libre permettant la transformation de plusieurs composés en produits oxydés ou polymérisés, qui peuvent être isolés à partir de plusieurs sources végétales tel que le radis, le navet et l'ail.

L'objectif de ce mémoire consiste en l'étude des caractéristiques des peroxydases d'origine végétale par bio-informatique en déterminant les séquences protéiques afin de comparer les homologies de séquence des différentes enzymes observées.

L'alignement multiple des séquences des peroxydases, a été réalisé au moyen de programmes bio-informatiques, tel que le MEGA X pour aligner les séquences et ESPript 3.0 pour analyser et visualiser les résultats.

Les résultats de cet alignement ont montré que la peroxydase a plusieurs homologues mais notre étude, nous nous sommes intéressés qu'à celles ayant un fort pourcentage de similitude. Nous avons retenu pour cela quatre résultats dont le pourcentage supra-cité varie entre 91% et 100%, et où les acides aminés ; His170, His42, et Arg38 sont conservés autour de site actif.

## Mots clés :

Peroxydases végétales, Alignement multiple, MEGA X , ESPript 3.0.

## Summary:

Peroxidase are enzymes that catalyze the oxydo-reduction reaction by a free radical mechanism allowing the transformation of several compounds into products oxidized or polymerized, which may be isolated from several plant sources such as as radish, turnip and garlic.

The aim of this paper is to study the characteristics of peroxidase of plant origin by bioinformatics by determining the protein sequences in order to compare the sequence homologies of the different enzymes observed.

The multiple alignment of the peroxidase sequences was achieved by means of bioinformatic programs, such as the MEGA X to align sequences and ESPript 3.0 to analyze and visualize the results.

The results of this alignment showed that peroxidase has several counterparts but our study, we focused only on those with a high percentage of similarity. For this, we selected four results, the percentage of which varies between 91% and 100%, and where amino acids; His170, His42, and Arg38 are conserved around of active site.

## Key Words:

Vegetable peroxidases , Multiple alignment, MEGA X , ESPript 3.0.

## ملخص

بيروكسيداز هي إنزيمات تعمل على تحفيز تفاعل خفض الأوكسيداز من قبل أ آلية جذرية مجانية تسمح بتحويل عدة مركبات إلى منتجات المؤكسد أو المؤكسد الذي يمكن عزله عن عدة مصادر نباتية مثل الفجل، اللفت والثوم.

الهدف من هذه الدراسة هو دراسة خصائص البيروكسيداز من أصل نباتي بواسطة المعلوماتية الحيوية عن طريق تحديد تسلسل البروتين من أجل مقارنة تسلسل الإنزيمات المختلفة التي لوحظت وقد تحققت الموازنة المتعددة لتسلسل بيروكسيداز بواسطة برنامج المعلومات البيولوجية مثل MEGAX لتنسيق التسلسلات و ESPript3.0 لتحليل وتصور النتائج .

وأظهرت نتائج هذا التناقص أن لبيروكسيداز عدة نظراء ولكن دراستنا ، ركزنا فقط على أولئك الذين لديهم نسبة عالية من His170 حول من الموقع النشط وتحفظ. التشابه. ولهذا اخترنا أربع نتائج تتفاوت نسبتها المئوية وبين 91 في المائة و 100 في المائة ، وحيث الأحماض الأمينية ؛

، His42 ، Arg38

:الكلمات الرئيسية:

MEGX ، ESPript 3.0. بيروكسيدات الخضروات ، المحاذاة المتعددة