

THESIS / THÈSE

MASTER EN BIOCHIMIE ET BIOLOGIE MOLÉCULAIRE ET CELLULAIRE À FINALITÉ APPROFONDIE

Caractérisation d'une protéine potentiellement impliquée dans les mécanismes d'action de polluants aquatiques : la protéine disulfure isomérase chez *Danio rerio*

BAYARD , Alexandre

Award date:
2014

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

UNIVERSITE DE NAMUR

Faculté des Sciences

Caractérisation d'une protéine potentiellement impliquée dans les mécanismes d'action de polluants aquatiques: la protéine disulfure isomérase chez *Danio rerio*

**Mémoire présenté pour l'obtention
du grade académique de master en biochimie et biologie moléculaire et cellulaire**

Alexandre Bayard

Janvier 2014

Caractérisation d'une protéine potentiellement impliquée dans les mécanismes d'action de polluants aquatiques: la protéine disulfure isomérase chez *Danio rerio*

BAYARD Alexandre

Résumé

Le triclosan (TCS) est un composé chimique structurellement proche des polychlorobiphényles (PCBs) et du bisphénol A (BPA), connu pour être un polluant majeur de l'environnement aquatique et un perturbateur endocrinien. Notre hypothèse est que cette molécule interagit comme ses deux homologues structuraux, avec une cible enzymatique intervenant dans la régulation de nombreuses hormones telles que la triiodothyronine (T_3) et l'œstradiol (E_2), à savoir la protéine disulfure isomérase (PDI). La PDI est une protéine soluble de 4 domaines présente dans la lumière du réticulum endoplasmique. L'organisme modèle que nous avons choisi, afin de caractériser les effets du TCS sur la PDI, est le poisson zèbre, *Danio rerio*. L'objectif est de caractériser structurellement et fonctionnellement la PDI et de déterminer l'impact du TCS sur son activité et sur la méthylation de son gène.

Différentes PCR ont été réalisées afin d'isoler et d'amplifier la séquence codante. Les amplicons générés par PCR ont été insérés dans un plasmide de stockage et transformés chez des *E. coli*. Deux méthodes de bio-informatique, à savoir la modélisation par homologie (EsyPred3D) et l'amarrage moléculaire (GOLD), ont été employées afin de réaliser un modèle 3D de la PDI de *Danio rerio* à partir d'une structure cristallographique de la PDI humaine (4EL1) et caractériser son interaction avec différentes molécules (E_2 , 4'-OH-PentaCB, 4'-OH-BDE17, BPA et le TCS). À partir de la modélisation par homologie, nous avons caractérisé un site d'interaction sur l'un des domaines de la protéine (domaine b'). L'amarrage moléculaire a permis d'identifier deux résidus (H256 et F304) déterminants pour l'interaction de la protéine avec les différents composés testés.

Pour la partie épigénétique, la technique d'amplification des îlots CpGs par PCR a été optimisée et des mesures préliminaires ont été réalisées à partir du PyroMark Q24 Advanced pour évaluer le pourcentage de méthylation de CpGs présents dans 3 régions du gène de la PDI (2 dans l'intron 2 et 1 dans l'intron 4).

En conclusion, l'étude multidisciplinaire de l'impact du TCS sur la PDI du poisson zèbre apporte une analyse complète de la problématique ainsi que des pistes de réflexion dans l'analyse comparative de différents composés.

Mémoire de master en biochimie et biologie moléculaire et cellulaire

Janvier 2014

Promoteur: F. Silvestre

Co-promotrice: C. Michaux

Table des matières

| | |
|---|-----------|
| Table des matières | 3 |
| 1. Introduction | 5 |
| 1.1 Problématique de la pollution environnementale..... | 5 |
| 1.1.1 Les polluants: | 5 |
| 1.1.2 Classification des POPs | 5 |
| 1.1.3 Les PCBs | 6 |
| 1.1.4 Le BPA..... | 7 |
| 1.2 Le Triclosan..... | 7 |
| 1.2.1 Généralités | 7 |
| 1.2.2 Effets du Triclosan..... | 8 |
| 1.2.3 Dégradation du TCS | 9 |
| 1.3 La PDI..... | 10 |
| 1.3.1 Généralités | 10 |
| 1.3.2 Classification | 11 |
| 1.3.3 Structure | 11 |
| 1.3.4 Fonctions..... | 12 |
| 1.3.5 Régulation | 13 |
| 1.3.6 Interactions | 14 |
| 1.4 Danio rerio | 16 |
| 1.4.1 Généralités | 16 |
| 1.4.2 Impacts physiologiques et métaboliques de polluants | 16 |
| 1.4.3 Impacts épigénétiques du TCS chez Danio rerio..... | 17 |
| 2. Objectifs | 18 |
| 3. Matériel et méthodes | 19 |
| 3.1 Isolation et amplification de la séquence codante..... | 19 |
| 3.1.1 Extraction des ARN totaux d'un Danio rerio | 19 |
| 3.1.2 Rétrotranscription | 20 |
| 3.1.3 Conception des amorces | 20 |
| 3.1.4 PCR initiale..... | 21 |
| 3.1.5 RACE-PCR..... | 22 |
| 3.1.6 Full Length-PCR..... | 22 |
| 3.1.7 Électrophorèse de vérification | 23 |
| 3.1.8 Stockage et clonage en bactéries | 23 |
| 3.1.9 Extraction plasmidique | 24 |
| 3.1.10 Séquençage..... | 24 |
| 3.2 Analyse bio-informatique..... | 24 |
| 3.2.1 Analyse de séquences | 24 |
| 3.2.2 Modélisation par homologie | 24 |
| 3.2.3 Amarrage moléculaire (Docking) | 25 |
| 3.3 Analyse de la méthylation | 25 |
| 3.3.1 Extraction de l'ADN total | 25 |
| 3.3.2 Traitement | 25 |
| 3.3.3 Conception des amorces | 26 |
| 3.3.3 Amplification..... | 28 |
| 3.3.4 Pyroséquençage | 29 |
| 3.3.5 Analyse du séquençage | 29 |
| 4. Résultats | 29 |
| 4.1 Partie biochimique | 29 |

| | |
|--|----|
| 4.1.1 Isolation et amplification de la séquence codante | 29 |
| 4.1.2 Analyse de la séquence peptidique | 31 |
| 4.1.3 Modélisation de la PDI de <i>Danio rerio</i> par homologie | 33 |
| 4.1.4 Amarrage moléculaire ou docking | 35 |
| 4.2 Partie épigénétique | 36 |
| 5. Discussion et perspectives | 38 |
| 6. Conclusions | 41 |
| 7. Bibliographie | 42 |

Caractérisation d'une protéine potentiellement impliquée dans les mécanismes d'action de polluants aquatiques: la protéine disulfure isomérase chez *Danio rerio*

1. Introduction

1.1 Problématique de la pollution environnementale

La pollution des différents milieux terrestres et aquatiques a augmenté ces dernières années, atteignant un point critique dans certaines régions du globe avec l'accroissement démographique et l'expansion industrielle.¹ Cette augmentation est principalement associée aux rejets d'eaux usées d'origine domestique, industrielle ou encore issue de l'agriculture et touchant en premier lieu les environnements aquatiques. Un exemple concret est le cas de la Belgique où l'état des cours d'eau s'est dégradé ces dernières années. De nombreuses recherches évaluant les niveaux de pollution des eaux de surface ont été réalisées afin de caractériser l'impact éventuel des polluants sur la biodiversité ainsi que sur le métabolisme et la physiologie d'organismes modèles tels que le rat (*Rattus norvegicus*), le poisson zèbre (*Danio rerio*),...

1.1.1 Les polluants:

Ces composés présents dans les eaux usées sont classés en deux groupes majeurs: les polluants organiques persistants (POPs) et les métaux lourds. Chacun de ceux-ci possède leurs propres classifications.

Les POPs ont été identifiés pour la première fois après la Seconde Guerre mondiale par la communauté scientifique les définissant comme des xénobiotiques présents de manière ubiquitaire dans l'environnement.¹ On les retrouve au niveau de l'atmosphère, du sol et des sédiments présents dans l'eau, mais peuvent également être assimilés par les organismes se stockant dans les tissus adipeux (composés hydrophobes).² Une concentration élevée de ces composés entraîne des effets néfastes pour la faune et la flore à long terme ainsi que pour la santé humaine. Les POPs peuvent être d'origine naturelle ou anthropogénique.³ Parmi les principales sources naturelles de pollution, on peut citer notamment les feux de forêt ou l'activité volcanique qui ont en commun la combustion et sont à la base des relâchements de grandes quantités de dioxyde de carbone.³ Un bon exemple de polluants provenant de l'industrie est un insecticide très connu, le DDT (dichlorodiphényltrichloroéthane), produit en masse à partir de 1945 et dont la production a été interdite dans de nombreux pays après la reconnaissance de ses propriétés toxiques.³ Une autre caractéristique propre à ces composés est leur capacité à persister dans l'environnement, c'est à dire résistant durant de très longues périodes à la dégradation chimique, physique et à la photolyse.³

Dans ce travail, nous nous focaliserons sur la classification des POPs, et plus particulièrement sur les composés phénoliques et certains éléments anthropogéniques, dont nous décrivons ci-dessous la classification.

1.1.2 Classification des POPs

Ce groupe de polluants est très diversifié, reprenant les composés organométalliques, les hydrocarbures aromatiques polycycliques (HAP), les produits chimiques phénoliques (ex: le bisphénol A ou BPA), les composés organochlorés, ... Ils sont classés en deux groupes: les POPs «non-intentionnels» (produits de dégradation) ou «intentionnels» (synthèse en industrie).³

La première catégorie reprend les HAPs, les dioxines et les furanes qui sont des sous-produits formés par combustion de matières organiques (lors d'éruptions volcaniques, de feux de

forêt) ou par photodégradation et présentant une toxicité élevée pour les organismes.³ Par exemple les HAPs sont des composés formés de minimum deux phényles et n'ayant que des atomes de carbone ou d'hydrogène. Ils sont issus de la combustion naturelle ou encore de la pyrolyse de matières fossiles (fuel, pétroles). Les dibenzofuranes polychlorés (PCDF) et les dioxines polychlorés (PCDD), formés par combustion de céréales, de carcasses animales, de déchets chimiques ou domestiques, ont été montrés comme étant des cancérigènes. Dans le cas de la dioxine, il a également été montré qu'elle pouvait être produite par photodégradation de PCBs dans l'eau.³

La seconde catégorie reprend les composés issus de l'industrie chimique. Les pesticides organochlorés (OCPs) sont des substances qui peuvent être de nature chimique mais également biologique (produits par certains champignons, plantes, ...).³ L'une de leurs caractéristiques est d'être lipophiles, se stockant facilement dans les tissus adipeux. Ce sont également de puissants neurotoxiques. L'exemple le plus couramment repris pour décrire les OCPs est le dichlorodiphényltrichloroéthane (DDT). Le plus connu des OCPs anthropogénique est le groupe des polychlorobiphényles (PCBs), qui comme le DDT, a été utilisé en abondance avant que son utilisation soit interdite.⁴ Tous ces composés chimiques sont classés depuis maintenant quelques années comme des perturbateurs endocriniens. Ici, nous décrivons et analyserons trois polluants appartenant aux POPs et très proches structurellement: les PCBs, le bisphénol A (BPA) et le triclosan (TCS) qui sera décrit plus en détail (**Fig. 1**).

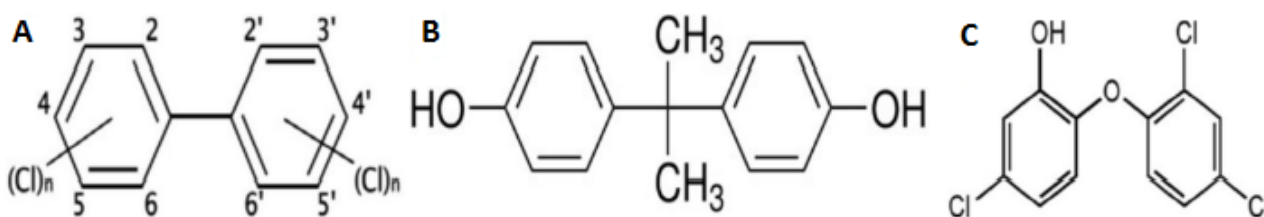


Figure 1: A) Polychlorobiphényles (PCBs), B) Bisphénol A (BPA), C) 5-chloro-2-(2,4-dichlorophénoxy)phénol ou triclosan (TCS)

1.1.3 Les PCBs

Utilisés à grande échelle durant les années 1930 à 1970 surtout pour leur propriété d'isolant électrique (transformateurs et condensateurs électriques), ils ont une répartition ubiquitaire dans l'environnement.⁵ Ces molécules sont une source de pollution majeure, certains d'entre eux ayant un temps de demi-vie pouvant être équivalent à plusieurs siècles. La plupart de ces produits se retrouvent au niveau des sols, mais on peut également en retrouver dans les milieux aquatiques, associés aux sédiments. Bien que les doses observées dans le milieu restent faibles, leurs capacités à se bioaccumuler dans les graisses animales et à se bioamplifier au travers de la chaîne alimentaire les rendent potentiellement très toxiques pour l'homme.⁶ Leur structure de base est le biphenyle qui peut être dérivé en de nombreuses isoformes de PCBs (Fig.1). De nombreuses recherches sont réalisées sur la toxicité de certaines isoformes comme celle de Wang *et al.* travaillant sur l'impact du PCB-77 sur le métabolisme du fer.⁷ On peut également citer l'article de Garner *et al.* portant sur l'impact du PCB-126 chez l'espèce modèle *Danio rerio*.⁸ Ce dernier est impliqué dans la diminution de l'expression d'un récepteur de la classe des AHRs (*aryl hydrocarbon receptors*). L'inhibition de ce récepteur chez l'embryon serait à l'origine de difformité cardiaque.⁸ Une recherche réalisée par Okada *et al.* (2009) identifie des PCBs comme inhibiteurs de la protéine disulfure isomérase (PDI), une protéine chaperone catalysant la formation et l'isomérisation des ponts disulfures.⁴ Ces résultats suggèrent notamment que les PCBs auraient un rôle indirect dans le repliement des protéines.

1.1.4 Le BPA

Comme les PCBs, le BPA a été utilisé de manière massive dans l'industrie plastique à partir des années 1960, servant de monomère pour la polymérisation du polycarbonate de bisphénol A. Bien que de plus en plus soumis à la législation, sa production n'a pas été interrompue, étant encore employé dans de nombreux produits. Certains pays ont d'ailleurs interdit son usage dans des produits de consommation tels que les biberons. De plus, l'année prochaine, son utilisation pour les emballages alimentaires sera interdite en Europe (directive 2011/8/UE). Une publication réalisée par Hiroi *et al.* (2006) a montré que le BPA pouvait affecter l'activité enzymatique de la PDI.⁹ Le BPA s'y lierait de manière compétitive, empêchant notamment son interaction avec deux de ses substrats: l'hormone 17 β -estradiol (E₂) et l'hormone thyroïdienne (T₃). Le BPA affecterait donc le système endocrinien et la maturation du système nerveux de l'organisme.⁹

1.2 Le Triclosan

1.2.1 Généralités

Parmi l'ensemble des polluants organochlorés, on retrouve un produit issu de l'industrie chimique ayant une structure très proche des PCBs et du BPA, le 5-chloro-2-(2,4-dichlorophénoxy)phénol également connu sous le nom de triclosan (TCS) (Fig.1).¹⁰ C'est ce composé que nous avons décidé de caractériser dans cette recherche. Comme ses deux homologues structuraux, les PCBs et le BPA, il est liposoluble, se stockant principalement chez l'individu dans les graisses et est également détectée dans les eaux usées, les eaux de surface, les sédiments et le sol.¹¹ On peut également noter que sa structure ressemble à la 3,3',5-triiodothyronine (T₃) qui est une hormone thyroïdienne impliquée à différents niveaux dans la physiologie et le métabolisme de nombreux organismes (Fig.2). Développé durant les années 1960, il est synthétisé industriellement depuis maintenant une trentaine d'années. On le retrouve notamment dans les produits pharmaceutiques et de soins personnels tel que le dentifrice, les déodorants, les jouets, les savons, ...¹⁰ Il est aussi employé comme antibactérien, antifongique, antiviral à large spectre et est utilisé dans certains produits comme agent de conservation. Le TCS est, au même titre que le BPA et les PCBs, un polluant majeur de l'environnement aquatique. En effet, une recherche d'Adolfsson-Erici *et al.* (2002) portant sur la mesure du TCS chez les poissons a montré qu'il se retrouvait chez les poissons exposés aux eaux usées mais également aux eaux traitées par les stations d'épuration.¹² Bien que ces dernières aient pour rôle de détruire ces composés, ceux-ci se retrouvent quand même en faible concentration à leur sortie.

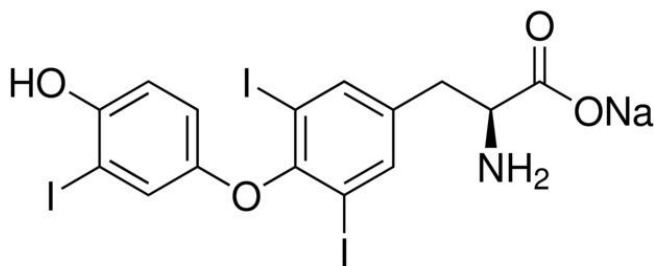


Figure 2: Représentation schématique de la 3,3',5-triiodothyronine (T₃).

1.2.2 Effets du Triclosan

1.2.2a Sur les organismes aquatiques

Plusieurs recherches ont montré que le TCS présente des effets délétères principalement sur la physiologie et le système endocrinien de nombreux organismes. L'une des recherches récentes (2009) réalisée par Oliveira *et al.* décrit les effets du TCS sur une espèce modèle de poisson, le poisson zèbre (*Danio rerio*) durant le stade larvaire et à l'état adulte.¹³ Cette étude se base sur une exposition de larves à des concentrations entre 0,1 et 0,5 mg/l de TCS durant une période de 6 jours, les adultes ayant été exposés aux mêmes concentrations mais durant une période plus courte de 96 heures.¹³ Les auteurs ont observé que pour une concentration de 0,5 mg/l, des déformations physiques apparaissent chez les larves et que l'ensemble des individus étaient morts après 120 heures d'exposition. Ils établissent qu'une concentration supérieure ou égale à 0,3 mg/l de TCS présente des risques importants pour les organismes aquatiques.¹³

Une autre recherche réalisée par Binelli *et al.* (2009) analyse l'effet génotoxique du TCS en exposant l'hémolymphe de moules zébrées (*Dreissena polymorpha*) à des concentrations entre 0,1 et 0,3 μM .¹⁴ La mesure de cet effet potentiel est réalisée par la méthode SCGE (*single cell gel electrophoresis*), ou méthode des comètes qui permet de déterminer la présence des cassures doubles brins de l'ADN.¹⁴ Les résultats montrent que le TCS a un effet significatif sur les hémocytes, et ceci à toutes les concentrations testées. En plus de déstabiliser l'ADN de ces cellules, la plus faible des concentrations en TCS (de 0,1 μM) est suffisante pour observer une apoptose.¹⁴

1.2.2b Sur les mammifères

Le TCS ne touche pas seulement les organismes aquatiques mais est également un polluant pour les organismes terrestres, comme le rat ou l'homme. Dans une étude réalisée sur les thymocytes de rat, par Kawanai *et al.* (2011), il est montré que le TCS peut induire une hyperpolarisation de la membrane cellulaire par l'augmentation de la concentration intracellulaire en Ca^{2+} .¹⁵ Plus précisément, leurs résultats montrent qu'à une concentration de 3 μM , la membrane devient hyperpolarisée suite au relargage du Ca^{2+} dans le cytoplasme. La réponse des cellules à ce stimulus est l'activation de canaux potassique (K^+), Ca^{2+} dépendants, qui augmente la perméabilité membranaire à l'ion K^+ . L'ensemble de ce processus perturbe fortement l'environnement cellulaire.¹⁵

Plusieurs recherches ont également démontré que ce polluant avait un impact sur l'homme. Dans une publication de Adolfsson-Erici *et al.* (2002), ils ont mesuré les concentrations en TCS dans 3 conditions: dans le lait des femmes enceintes, chez les poissons exposés aux eaux usées et à l'eau traitée par les usines d'épuration.¹² Ils observent que ce composé chimique est retrouvé dans les trois cas. Plus précisément, dans le cas du lait de femmes enceintes, 5 échantillons sont pris au hasard et analysés par chromatographie en phase gazeuse couplée à de la spectrométrie de masse et montrent des concentrations de 20 $\mu\text{g}/\text{kg}$ à 300 $\mu\text{g}/\text{kg}$.¹² Cette présence du TCS dans le lait maternel peut être associée à différents effets indésirables comme par exemple une augmentation de l'apparition d'allergies. Le fait que l'on observe ne serait-ce que de faibles concentrations dans le lait maternel est un signe alarmant de la contamination par le TCS de l'homme et pas uniquement les organismes aquatiques.

Une autre recherche menée par Clayton *et al.* (2010) a démontré un lien entre le TCS, le BPA et le système immunitaire de l'homme.¹⁶ Cette étude est basée sur des résultats obtenus durant les années 2003 à 2006 par le programme d'évaluation de l'état sanitaire et nutritionnel de la population américaine: le NHANES (*National Health And Nutrition Examination Survey*).¹⁶ Ils y comparent les concentrations urinaires du BPA et du TCS avec les symptômes d'immuno-diminution tels que les allergies, le rhume des foies et la présence de cytomégalovirus (CMV)

infectant les personnes immunodéprimées. Les résultats de cette recherche montrent que le BPA et le TCS affectent négativement l'immunité.¹⁶

1.2.3 Dégradation du TCS

Le TCS est instable et peut se dégrader de différentes manières notamment en deux polluants majeurs: les dioxines et les chlorophénols.

1.2.3a Dioxines

Les dioxines désignent des molécules hétérocycliques et aromatiques avec deux atomes d'oxygènes dans un cycle aromatique. Elles comprennent les polychlorodibenzo-para-dioxines (PCDDs, dans lequel se trouve la dioxine), les polychlorodibenzofuranes (PCDF) ainsi que des PCBs de type «dioxine». Dans la recherche menée par K. Aranami *et al.* (2007), il a été observé que le TCS est dégradé dans l'eau suite à une exposition prolongée à la lumière, formant un composé de la famille des PCDDs: le 2,8-dichlorodibenzo-para-dioxine (DCDD) (Fig.3).¹⁷ Il est aussi montré que l'efficacité de la dégradation par photolyse diffère entre l'eau pure (concentration en DCDD inférieur à 0,02 mg/l après 12 jours), l'eau douce (~0,06 mg/l) et l'eau de mer (~0,08 mg/l) où ce dérivé semble être plus stable.¹⁷ Le temps de demi-vie du TCS serait d'approximativement 4 à 8 jours.¹⁷ Notons qu'il s'agit d'une expérience *in vitro* basée sur des échantillons d'eau, ne reprenant pas les conditions exactes propres à l'environnement, où ce composé peut interagir avec les sédiments et s'y stabiliser.

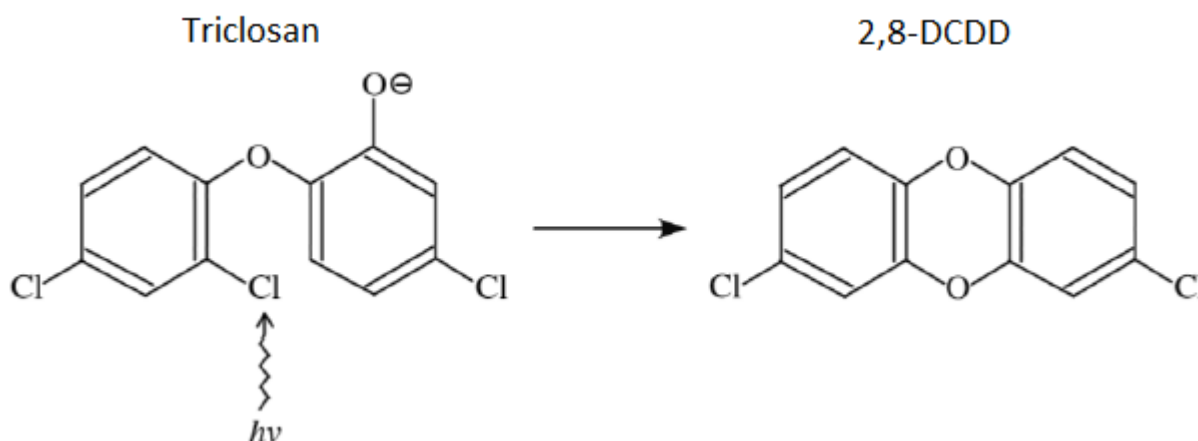


Figure 3: Dégradation du TCS par la lumière en 2,8-dichlorodibenzo-para-dioxine (2,8-DCDD).¹⁷

1.2.3b Chlorophénols

Le TCS ne se dégrade pas uniquement en DCDD, il peut également former d'autres composés toxiques tels que les chlorophénols. Il s'agit de composés aromatiques dérivés du phénol et présentant un ou plusieurs atomes de chlore. Dans une étude de Canosa *et al.* (2005) portant sur la dégradation du TCS en milieu aqueux, il est montré que celui-ci forme plusieurs sous-produits chlorés en présence d'échantillons d'eau chlorée.¹⁸ Les deux principaux sous-produits mesurés sont le 2,4-DCP et le 2,4,6-trichlorophénol (2,4,6-TCP) dérivé du 2,4-DCP sur lequel se lie un ion chlorure (Cl⁻). Le 2,4,6-TCP est un composé utilisé comme antifongique pour la protection du bois et est connu comme étant cancérigène. D'autres produits sont également formés dans l'eau chlorée comme le tetraclosan I et II ainsi que le Pentaclosan (Fig.4).¹⁸

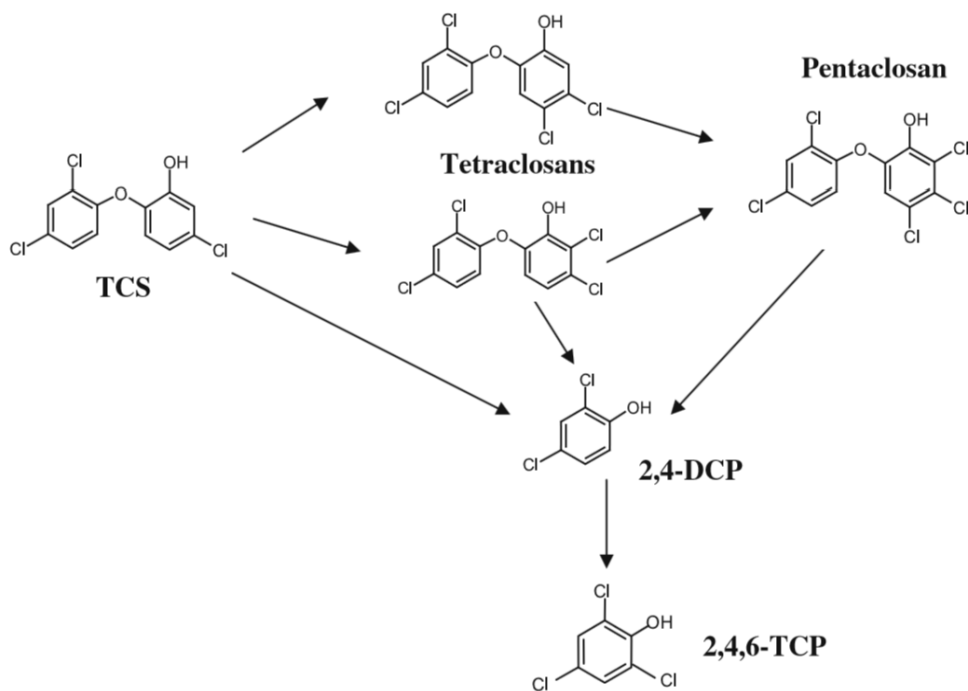


Figure 4: Dégradation du TCS en milieu aqueux chloré.¹⁸

Comme on peut le constater, le TCS semble être, au même titre que les PCBs ou le BPA, un polluant majeur de notre environnement, se retrouvant principalement dans les eaux usées. De plus, il présente différents effets toxiques pour les individus, étant un agent cancérigène, un neurotoxique et un perturbateur endocrinien. Notre hypothèse dans ce travail, est que cette molécule interagit avec les mêmes cibles enzymatiques que ses homologues structuraux (Voir 1.1.2 et 1.1.3) et notamment avec une protéine présente dans de nombreux organismes vivants: la protéine disulfure isomérase (PDI). Cette protéine semble avoir un rôle important au niveau du développement et est également une protéine chaperone pour de nombreuses hormones, intervenant donc dans le système endocrinien.

1.3 La PDI

1.3.1 Généralités

Le terme PDI est commun à de nombreuses enzymes ayant parfois des tailles et structures différentes. Nous nommerons par PDI les protéines encodées par le gène P4HB (*Prolyl-4-hydroxylase, subunit beta*), et connues également comme PDIA1 (*Protein disulfide isomerase family A, member 1*) qui sont exprimées de manière ubiquitaire dans les différents tissus.¹⁹ Elles ont l'avantage d'être très bien caractérisées dans les domaines de l'écotoxicologie, la protéomique et semblent également être associées à des études de carcinogenèse. La PDI d'une taille de 55 kDa (environ 500 acides aminés) possède 4 domaines (a, a', b, b') ainsi qu'un sous-domaine (c) et est présente en abondance dans le réticulum endoplasmique (RE) (Fig.5).²⁰ Catalysant le repliement des structures tertiaires de nombreuses protéines, elle est une protéine chaperone pour de nombreuses protéines et hormones (hormones thyroïdiennes majoritairement). De plus, cette enzyme catalyse l'oxydoréduction des ponts disulfure, permettant leur formation et leur isomérisation. De nombreuses recherches ont catégorisé de nombreux homologues de la PDIA1, par exemple les ERp44, PDIp, EndoPDI,...²¹

1.3.2 Classification

La PDI fait partie de la superfamille des thioredoxines. La thioredoxine est une petite protéine de 10 kDa ayant une activité oxydoréductase dépendante d'un motif précis CXXC (deux cystéines séparées par deux acides aminés).²¹ Celui-ci est nécessaire pour la formation et l'isomérisation des ponts disulfure. La famille des thioredoxines est très diversifiée, avec des protéines d'espèces très éloignées phylogénétiquement. On peut par exemple citer la protéine DsbA (*disulfide bond A*) exprimée par les *Escherichia coli* (bactérie), la Mpd1p exprimée chez *Saccharomyces cerevisiae* (levure), la PDIA exprimée chez les mammifères ou encore la P4HB exprimée chez le poisson zèbre (*Danio rerio*).²¹ Cette superfamille se divise en groupes, chaque enzyme étant classée selon son espèce. Parmi ses différentes familles se trouvent notamment celle des thioredoxines (PDIs) humaines reprenant 19 membres, ayant des tailles et localisations très variables.²² Par exemple, on peut citer les 4 protéines transmembranaires TMX (*thioredoxin-related transmembrane protein*=TMX, TMX2, TMX3, TMX4), alors que les 15 autres membres de la famille des PDIs sont solubles dans la lumière du RE.²²

1.3.3 Structure

Les 4 domaines de la PDI ont le repliement thioredoxine composé de 3 hélices α entourant 4 feuillets β , alors que le sous-domaine c n'a qu'une hélice α comme structure secondaire (Fig.6).²⁰

Les domaines a et a' (PDI-a et PDI-a') sont homologues, ayant chacun un site actif reprenant le motif CXXC (CGHC chez l'homme).²¹ Dans une publication récente de Wang *et al.* (2013), ils ont caractérisé la structure des 4 domaines de la PDI humaine (hPDI) par cristallographie, permettant de distinguer la forme réduite et oxydée du motif CGHC.²⁰

Les domaines b et b' (PDI-b et PDI-b') possèdent également le repliement thioredoxine mais, contrairement aux domaines a et a', pas le motif catalytique. La séquence polypeptidique de PDI-b' diffère fortement des autres domaines, étant principalement constitué d'acides aminés (a.a) hydrophobes. Différentes recherches de mutagenèse dirigée ont permis d'identifier des a.a nécessaires à l'interaction entre l'enzyme et ses substrats.

La structure du PDI-c est moins bien définie que les autres domaines, déterminée par cristallographie chez certaines espèces, comme dans l'article de Tian *et al.* (2005) qui représente ce domaine chez la levure comme une hélice α , dont l'extrémité est désordonnée.²³ Les composés à l'extrémité C terminale sont principalement des résidus acides comme l'acide aspartique (11/32 a.a) et glutamique (8/32 a.a), les derniers résidus étant le motif KDEL ou HDEL notamment chez la levure ou certains poissons (*Danio rerio*) qui est le motif d'adressage au RE. Dans une recherche de Lucero *et al.* (1998), ils mettent en évidence que le domaine c serait un site potentiel de liaison du calcium.²⁴

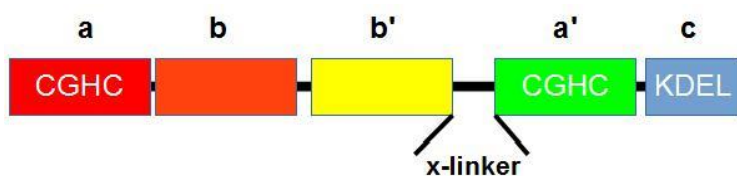


Figure 5: Représentation schématique des domaines de la PDI.

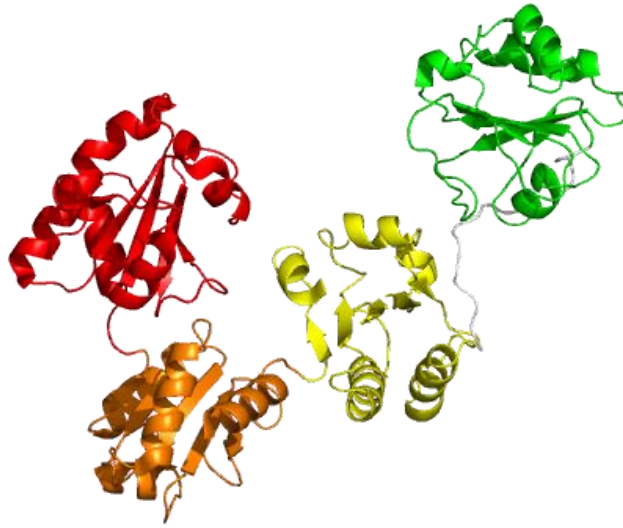


Figure 6: Structure cristallographique de la PDI humaine avec ses 4 domaines: a (rouge), b (orange), b' (jaune) et a' (vert) (PDB code: 4EL1). La forme globale de la protéine est de type globulaire, comme pour la majeure partie des protéines solubles, ayant une forme d'U (ou de bateau), PDI-b et PDI-b' à la base et PDI-a et PDI-a' présent de chaque côté.

1.3.4 Fonctions

On retrouve parmi les substrats de la PDI, des protéines telles que les immunoglobulines (système immunitaire), la RNase A (régulation traductionnelle) ou des hormones telles que l'hormone thyroïdienne (système endocrinien).²¹ Sous-unité β de la prolyl-4-hydroxylase, elle catalyse la formation du collagène et de la MTTP (*microsomal triglyceride transfer protein*) intervenant dans le repliement des lipoprotéines.¹⁹ La réaction catalysée par la PDI est une modification post-traductionnelle, intervenant directement durant la synthèse de la chaîne polypeptidique si le ribosome est ancré sur la membrane du RE ou après l'adressage de la protéine quand celle-ci se retrouve dans la lumière du RE. La formation et l'isomérisation des ponts disulfures ne peuvent pas être catalysées dans le cytosol (milieu réducteur). Les protéines ayant des ponts disulfures sont majoritairement sécrétées, ou adressées dans d'autres compartiments cellulaires (noyau, Golgi).

Deux études de Lappi *et al.* (2004 et 2010) ont identifié et caractérisé un résidu arginine (R120) présent dans le PDI-a et dont les mouvements de la chaîne latérale modulent le pKa du site catalytique.^{25,26} Ces résultats s'appuient sur l'analyse des structures RMN (résonance magnétique nucléaire) présentes dans la PDB (*Protein Data Bank*) et la mutagenèse dirigée. La mutation de cette arginine, que ce soit sur PDI-a ou PDI-a', est associée à une diminution de l'activité de la PDI.²⁶ Ces deux recherches suggèrent que le motif CXXC soit régulé par d'autres résidus proches structurellement et que la modulation du pKa du site actif par la chaîne latérale d'un résidu ou par les conditions cellulaires (pH), a un impact important sur l'activité de la PDI.^{25,26}

Dans le RE, le milieu est fortement oxydant, et bien que les conditions soient assez difficiles à reproduire *in vitro*, la PDI se trouve majoritairement sous forme oxydée.²² Comme on peut le voir sur la figure 7, le pont disulfure de la PDI interagit avec une fonction thiol du substrat formant un lien covalent entre les deux. Suivant la conformation du substrat et la présence d'une fonction thiol proche, il y aura formation d'un pont disulfure.²² Si ces conditions ne sont pas remplies, la PDI reformera un lien covalent entre les deux cystéines de son motif CXXC.

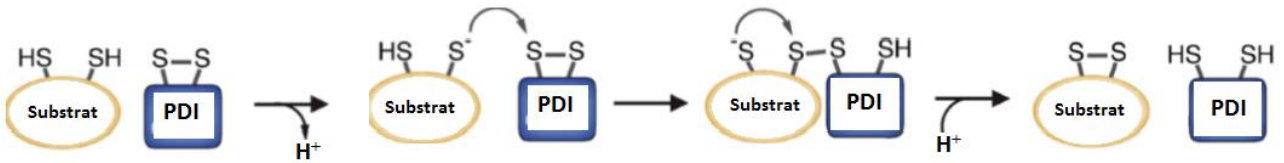
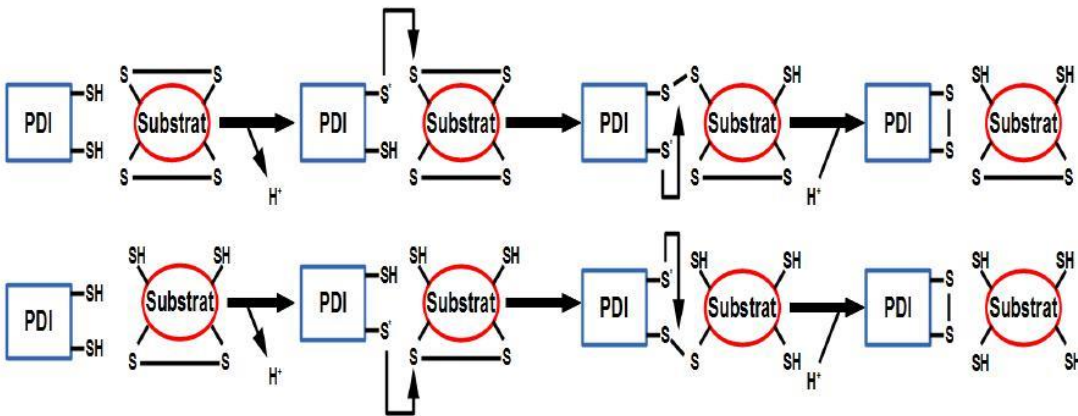


Figure 7: Représentation schématique de la réaction d'oxydoréduction catalysée par la PDI et induisant la formation d'un pont disulfure sur le substrat (protéine ou hormone).²²

La réaction d'isomérisation est légèrement plus complexe. Comme on peut le voir sur la figure 8, la PDI réduit dans un premier temps les ponts disulfures de son substrat avant d'oxyder les fonctions thiols pour reformer deux ponts disulfures mais dont l'agencement est différent.²¹ Cette étape de dépliement du substrat par réduction des ponts disulfures demande énormément d'énergie pour la PDI. Il semblerait que celle-ci possède une affinité 100 fois plus importante pour les protéines dépliées donc l'accès aux fonctions thiols est facilité.²¹

Isomérisation de deux ponts disulfures d'un substrat par la PDI

A. Réduction des ponts disulfures du substrat par la PDI



B. Oxydation des fonctions thiols du substrat par la PDI

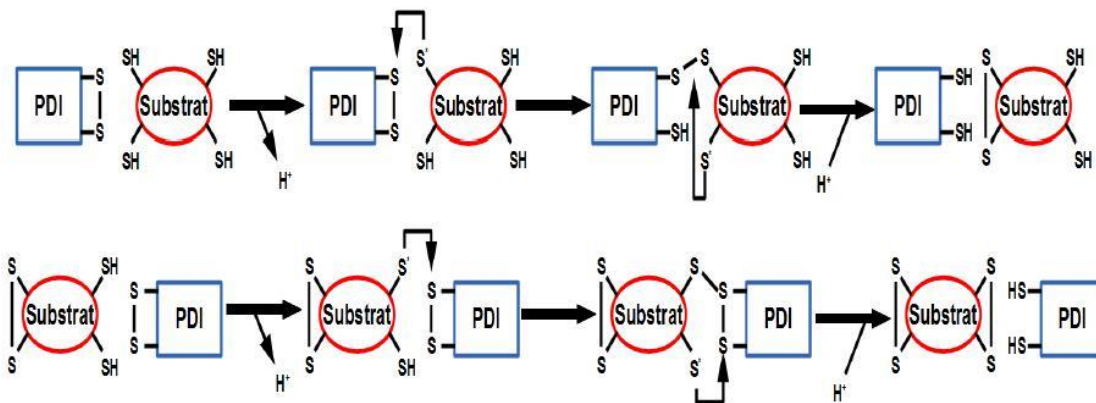


Figure 8: Isomérisation de ponts disulfures par la PDI

1.3.5 Régulation

L'état oxydé ou réduit de la PDI est régulé par deux facteurs majeurs: le milieu redox et la

flavoenzyme Ero1 (*Endoplasmic reticulum oxidoreductin 1*) α et β (Fig.9).²² Ces protéines interviennent dans le repliement des protéines, catalysant l'oxydation de la PDI à partir d'O₂ et formant comme sous-produit des espèces réactives de l'oxygène (ex: H₂O₂).^{22,27} Un autre régulateur majeur de l'état redox des protéines est le glutathion (GS). Il s'agit d'un tri-peptide (composé d'un acide glutamique, d'une cystéine et d'une glycine) formant un tampon redox (GSH/GSSG). Sa forme réduite s'oxyde facilement en cédant un électron à une protéine, comme la PDI, permettant sa réduction. Le GSH interagit également avec la GSH peroxydase, catalysant la réduction de fonctions peroxydes par la formation de GSSG. Cette forme oxydée est générée en abondance dans le RE contribuant ainsi au maintien de l'environnement oxydant.

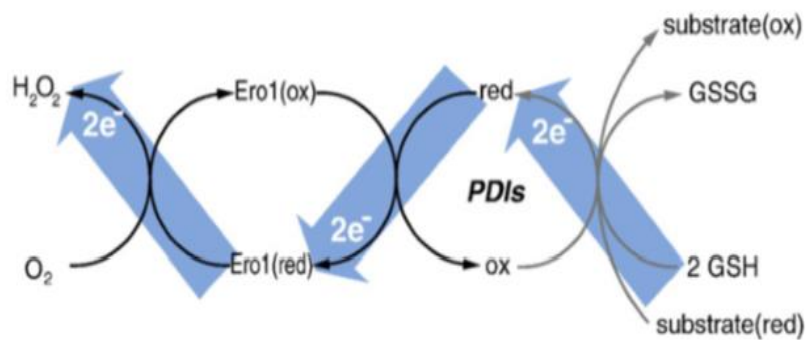


Figure 9: L'état oxydé et réduit de la PDI est régulé dans le RE par deux composants: Ero1 et le glutathion.²²

1.3.6 Interactions

1.3.6a Généralités

PDI-b' est resté durant de nombreuses années le seul domaine de la PDI humaine n'ayant pas de structure 3D déterminée expérimentalement. De nombreuses recherches récentes portant sur le domaine b' ont permis de caractériser sa structure (conformation thioredoxine) ainsi que sa fonction au sein de la PDI.²¹ En effet, ces différentes études suggèrent que PDI-b' serait associé à l'interaction de la PDI avec ses substrats. Dans une recherche réalisée par Nguyen *et al.* (2008), les différents résidus intervenant dans l'interaction de la protéine ont été identifiés notamment par RMN, fluorescence et mutagenèse dirigée. Les 16 résidus identifiés sont: F240, E242, A245, F249, H256, L258, F260, I289, I291, R300, I301, F304, F305, I318, L320, M324 (I=Isoleucine, E=Acide glutamique, A=Alanine, H=Histidine, L=Leucine, F=Phénylalanine, R=Arginine, M=Méthionine).²⁸ La numérotation employée correspond à la séquence peptidique de la PDIA1 humaine. Ces acides aminés forment dans le domaine b' une poche présentant un environnement hydrophobe.²⁸

1.3.6b Protéine/substrat

Dans la recherche de Nguyen *et al.* (2008), ils ont identifié également un polypeptide régulant l'accès à PDI-b': le *x-linker*.²⁸ Il s'agit d'une séquence peptidique de 16 acides aminés formant une boucle ou une hélice α entre PDI-b' et PDI-a'. Plus précisément, ils ont pu déterminer par NMR et fluorescence que le *x-linker* semble interagir avec la poche de PDI-b'. Ils ont également obtenu une structure cristallographique (code PDB=3BJ5) du domaine b' coiffé par le *x-linker* au niveau de son site d'interaction.²⁸ Une autre recherche de Wang *et al.* (2011) montre que la conformation globale de la PDI est dépendante de son état d'oxydation et de réduction.²⁹ Ils démontrent notamment que le site actif de PDI-a' régule, au même titre que le *x-linker*, l'accès à la

poche hydrophobe en entraînant un changement conformationnel de la protéine (Fig.10).²⁹ Lorsque le site catalytique est sous forme réduite, un résidu tryptophane (W396) interagit avec PDI-b', diminuant l'affinité de la PDI avec les polypeptides dépliés.²⁹ Cet acide aminé proche du motif CXXC se retrouve en dehors du site de liaison quand PDI-a' est sous sa forme oxydée.²⁹ Ils ont également obtenu une structure cristallographique (code PDB=3UEM) du domaine b, b', x-linker et a', avec une molécule de (4S,5S)-1,2-dithiane-4,5-diol cocrystallisée dans le site de liaison de PDI-b'.²⁹

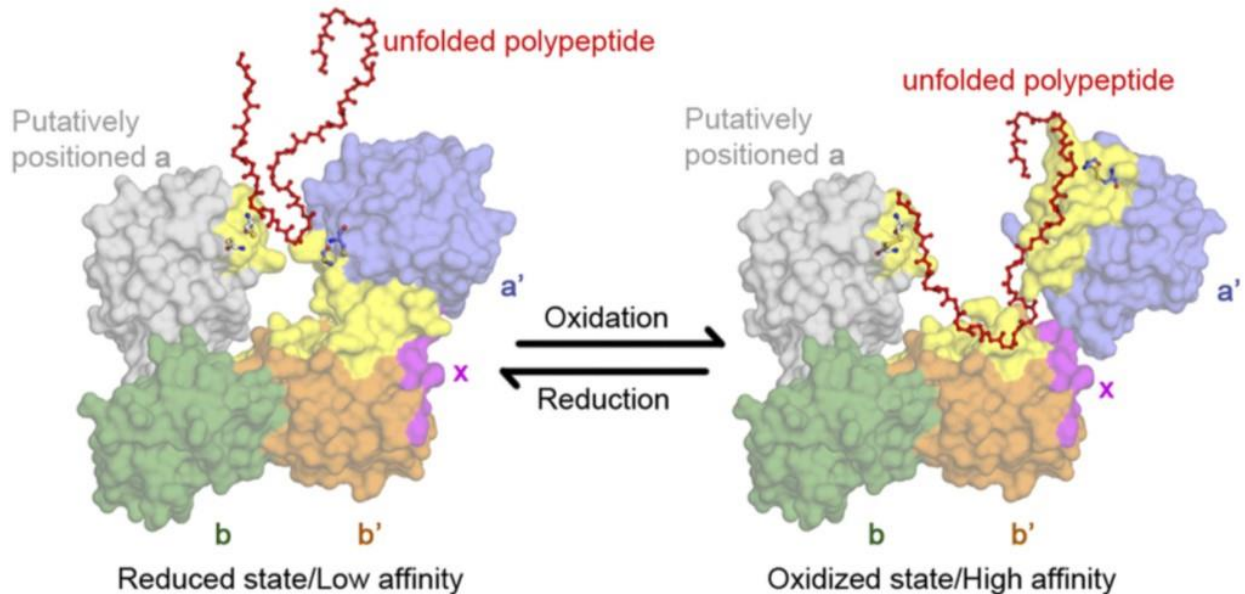


Figure 10: Représentation du changement de conformation de la PDI humaine entre sa forme oxydée et réduite.²⁹

1.3.6c Protéine/inhibiteurs

Une étude récente réalisée par Hashimoto *et al.* (2011) a mis en évidence l'interaction du BPA avec la PDI de rat.³⁰ À l'aide d'un test d'activité employant la RNase A, ils ont établi que la molécule inhibait l'activité de la protéine. Ils ont également réalisé de la mutagenèse dirigée sur certains résidus de PDI-b', identifiant H256 comme étant un résidu déterminant pour l'interaction du BPA avec la PDI, l'inhibition étant moins efficace pour cette PDI modifiée. H256 est l'un des acides aminés ayant été identifiés pour faire partie du site de liaison de PDI-b' à son substrat.³⁰ On peut supposer que cette interaction entraînerait un encombrement stérique de la poche de PDI-b', empêchant la protéine de pouvoir interagir avec son substrat.

Le BPA n'est pas le seul polluant à avoir été identifié comme un inhibiteur de l'activité de la PDI interagissant sur PDI-b'. En effet Hashimoto *et al.* a également montré dans une publication de 2012 des résultats montrant l'inhibition de la PDI de rat par des polybromodiphényléthers (PBDEs).³¹ Ces produits bromés sont utilisés principalement dans des produits ignifugés. Comme on peut le voir sur la figure 11, leur structure est très similaire au BPA, aux PCBs et surtout au TCS. Cette recherche reprend les méthodes de mutagenèse dirigée ayant permis de réaliser des formes tronquées de la PDI (b'a'c et a'c), sur lesquelles ont été réalisées des tests d'activités, identifiant le domaine b' de la PDI comme déterminant pour l'inhibition de l'enzyme par les PBDEs.³¹ De plus, il a également été observé que la forme bb' de la PDI précipitait en présence de la RNase A (substrat de la PDI), et que cette co-précipitation diminuait en présence de 4'-OH-BDE17 (une des isoformes de PBDEs).³¹ Ces deux recherches tendent à montrer que le BPA et les PBDEs, qui sont des composés très proches structurellement du TCS, interagissent avec le domaine b' de la PDI et

inhibent son activité catalytique.

Pour ce qui est des PCBs, une étude d'Okada *et al.* (2009) a montré que plusieurs formes de PCBs, pouvait perturber l'activité chaperone de la PDI. Ils démontrent notamment que comme le BPA, ces composés sont des inhibiteurs compétitifs, le 4'-OH-PentaCB semblant être la forme optimale à l'inhibition de la liaison de la PDI avec l'un de ses ligands, l'hormone thyroïdienne T₃. Bien qu'il n'identifie pas précisément le site de liaison de ses composés, il souligne que leur effet est similaire au BPA.⁴

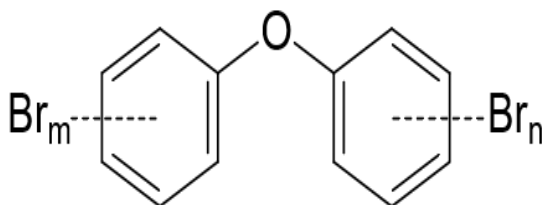


Figure 11: Structure de base des polybromodiphényléthers (PBDEs) dérivés en de nombreuses isoformes (209) suivant le nombre de brome (Br), d'hydroxyle, ... ajouté au diphényl.

1.4 *Danio rerio*

1.4.1 Généralités

Danio rerio, plus communément appelé poisson zèbre, est un organisme-modèle utilisé depuis plusieurs années dans les domaines de la recherche génétique, médicale ou encore en toxicologie. C'est un petit vertébré aquatique originaire d'Inde et appartenant à la famille des cyprinidés (*Cyprinidae*), groupe majeur de poissons d'eau douce reprenant plus de 2000 espèces différentes. Il présente de nombreux avantages tels qu'une petite taille (4-5 cm chez l'adulte), un développement rapide, une similarité génétique avec l'homme, une facilité à être élevé (coût très faible, peu d'entretien) et à se reproduire (100-200 œufs/ponte), ... Par ailleurs, de nombreuses recherches toxicologiques ont été publiées sur les effets de polluants (PCB, BPA et métaux lourds) sur la physiologie et le métabolisme du *Danio rerio*, permettant d'identifier les propriétés de ces molécules et leurs toxicités éventuelles avant d'extrapoler sur l'homme. On peut citer notamment un article de McGrath *et al.* (2008) qui a mis en évidence des similitudes entre les résultats portant sur le *Danio rerio* et ceux sur les mammifères.³² Ses larves peuvent assimiler les molécules diluées dans l'eau par simple diffusion à travers la peau, permettant d'étudier l'impact de ces composés chimiques sur un organisme entier et pas seulement sur un type tissulaire donné. De plus, les larves étant transparentes, elles permettent une étude facile du développement de l'organisme durant ses premiers stades.

1.4.2 Impacts physiologiques et métaboliques de polluants

De nombreuses publications très récentes ont permis de caractériser la méthode de bioaccumulation de composés chimiques issus de l'industrie, ainsi que leurs effets sur le développement, l'épigénétique et la protéomique.

On peut par exemple citer deux études récentes de Huang *et al.* (2011) portant sur le parathion-méthyle (PM) (pesticide organophosphoré) et de Ling *et al.* (2012) portant sur le cadmium, qui ont mis en évidence des différences dans le profil d'expression de protéines intervenant dans le stress oxydant chez le poisson zèbre.^{33,34} On constate donc que les polluants n'ont pas uniquement un impact direct sur une protéine mais peuvent également agir sur l'expression de protéines au niveau de leur transcription ou traduction.

Ces deux molécules ne sont pas les seules ayant été étudiées sur le poisson zèbre, on peut notamment citer l'étude de Lindholst *et al.* (2003), qui se sont intéressés à la cinétique de dégradation du sulfate de BPA et du BPA glucuronide.³⁵ Ils ont mis en évidence la relation entre ces deux composés et la réponse oestrogénique, plus particulièrement dans l'induction de vitellogénine.³⁵ Plus précisément, la vitellogénine est un marqueur oestrogénique synthétisé au niveau du foie des poissons femelles suite à l'activation d'un récepteur aux œstrogènes et est impliqué dans la féminisation des organismes mâles. Leurs résultats suggèrent qu'en présence de ces formes de BPA, qui sont des composés proches de l'œstradiol (E₂) au niveau de leurs structures, il y a activation de l'expression de la vitellogénine notamment chez les poissons zèbres mâles ou chez des jeunes organismes.³⁵ Cette expression de la vitellogénine par un polluant comme le BPA entraîne une féminisation des individus, responsable notamment de modifications de leur métabolisme, de leur physiologie et, à plus long terme, de leur succès reproducteur (et donc de leur fitness).

1.4.3 Impacts épigénétiques du TCS chez *Danio rerio*

L'épigénétique correspond à l'ensemble des mécanismes héréditaires affectant l'expression des gènes d'une cellule et qui ne se trouvent pas dans les séquences nucléotidiques. Ces changements permettent l'adaptation d'un organisme à différentes perturbations (stress, modulation de la température, pollution,...) survenant au cours de son développement et peuvent, dans certains cas, être transmises à sa descendance. Cette information épigénétique peut prendre différentes formes: la méthylation de l'ADN, les ARN non codants et la phosphorylation, méthylation ou acétylation des histones (modifications post-traductionnelles).³⁶

Dans le cadre de ce travail, nous nous concentrerons sur la méthylation de l'ADN qui est une modification épigénétique que l'on retrouve chez un très grand nombre d'organismes vivants (animaux, certaines plantes et champignons, ainsi que certaines souches de bactéries). Cette méthylation consiste en l'ajout d'un groupement méthyle sur l'une des quatre bases nucléotidiques de l'ADN, la cytosine étant majoritairement méthylée chez les eucaryotes alors que c'est l'adénine qui est méthylée chez les procaryotes. Plus particulièrement, la méthylation chez les animaux est symétrique sur les 2 brins de l'ADN, 70 à 80% des cytosines, formant un dinucléotide avec une guanine (CpG), étant méthylés.³⁶

Une étude récente de Strömquist *et al.* (2010) suggère que le niveau de méthylation de l'ADN pourrait être altéré par l'exposition d'organismes à des œstrogènes de synthèse telle que l'éthinylestradiol.³⁷ Plus précisément, ils ont analysé 3 îlots CpG identifiés dans la région 5' flanquant le gène de la vitellogénine, mesurant le niveau de méthylation dans le foie et le cerveau. Par la méthode de pyro-séquençage, ils ont ainsi comparé ces différentes mesures en fonction du sexe de l'organisme, du tissu et de l'exposition ou non à l'éthinylestradiol.³⁷ Ils ont montré que le niveau de méthylation diminuait suite à l'exposition à cette molécule. Ces résultats suggèrent donc que des composés proches des œstrogènes tels que l'éthinylestradiol (et potentiellement tels que les PCBs, le BPA ainsi que le TCS) pourraient modifier la méthylation des îlots CpG chez des gènes dépendant du système endocrinien.

Le *Danio rerio* est donc un organisme de choix pour des études d'écotoxicologie. En effet, au vu des résultats présentés (1.1.3, 1.2.2a, 1.4.2 et 1.4.3), nous constatons que de nombreux polluants aquatiques proches structurellement d'hormones (E₂), tels que le BPA, les PCBs mais également le TCS, sont responsables de dysfonctionnements au niveau de la physiologie et du métabolisme du poisson zèbre. Plus précisément, ces molécules semblent avoir un impact sur le système endocrinien de nombreux organismes, agissant potentiellement sur l'activité et l'expression de la PDI qui est un élément central intervenant dans le

métabolisme des hormones thyroïdiennes et des œstrogènes. L'étude de la PDI chez un organisme modèle comme *Danio rerio* permettrait une meilleure compréhension du mode de fonctionnement de polluants aquatiques tels que le TCS sur la physiologie et le métabolisme de ce poisson.

2. Objectifs

Ce mémoire s'inscrit dans un projet multidisciplinaire visant à étudier l'impact moléculaire et épigénétique de polluants, et notamment du TCS, sur la PDI de *Danio rerio*.

Dans la première partie, nous envisageons d'isoler et d'amplifier la séquence codante de la PDI par la méthode de PCR (*Polymerase chain reaction*), afin de permettre l'identification et la caractérisation de la séquence nucléotidique. Cette étape permettra, à plus long terme, d'insérer cette séquence dans un plasmide bactérien de surexpression afin de surproduire la PDI et de caractériser son activité et son inhibition *in vitro* par des polluants.

Dans la deuxième partie, un modèle théorique de la PDI de *Danio rerio* sera construit à partir de sa séquence en acides aminés. La méthode de modélisation par homologie est utilisée et la structure cristallographique de la PDI humaine est choisie comme modèle. Après avoir identifié le site de liaison du PDI-b', le mode de liaison de différents polluants est prédit au sein de ce dernier. Les composés 4'-OH-BDE17, BPA, E2, 4'-OH-PentaCB ont été sélectionnés car des données expérimentales ont montré qu'ils interagissaient avec le domaine b' de la PDI. La comparaison avec une molécule structurellement proche, le TCS, sera réalisée. Ces premières analyses permettront de mettre en évidence des relations structure-activité de polluants vis-à-vis de la PDI.

Ces deux premières étapes ont pour but de comprendre de manière approfondie le mécanisme d'action du TCS sur l'activité enzymatique de la PDI.

Dans la troisième partie, nous envisageons d'identifier et de caractériser des îlots CpG présents sur le gène de la PDI par la méthode de pyroséquençage. Le but à long terme est de déterminer si la méthylation des CpGs du gène de la PDI est affectée par l'exposition du *Danio rerio* à différentes concentrations en TCS. Nous pourrions alors avoir une vision globale et multidisciplinaire de l'impact d'un polluant notamment sur l'expression d'un gène et la biochimie de la protéine associée.

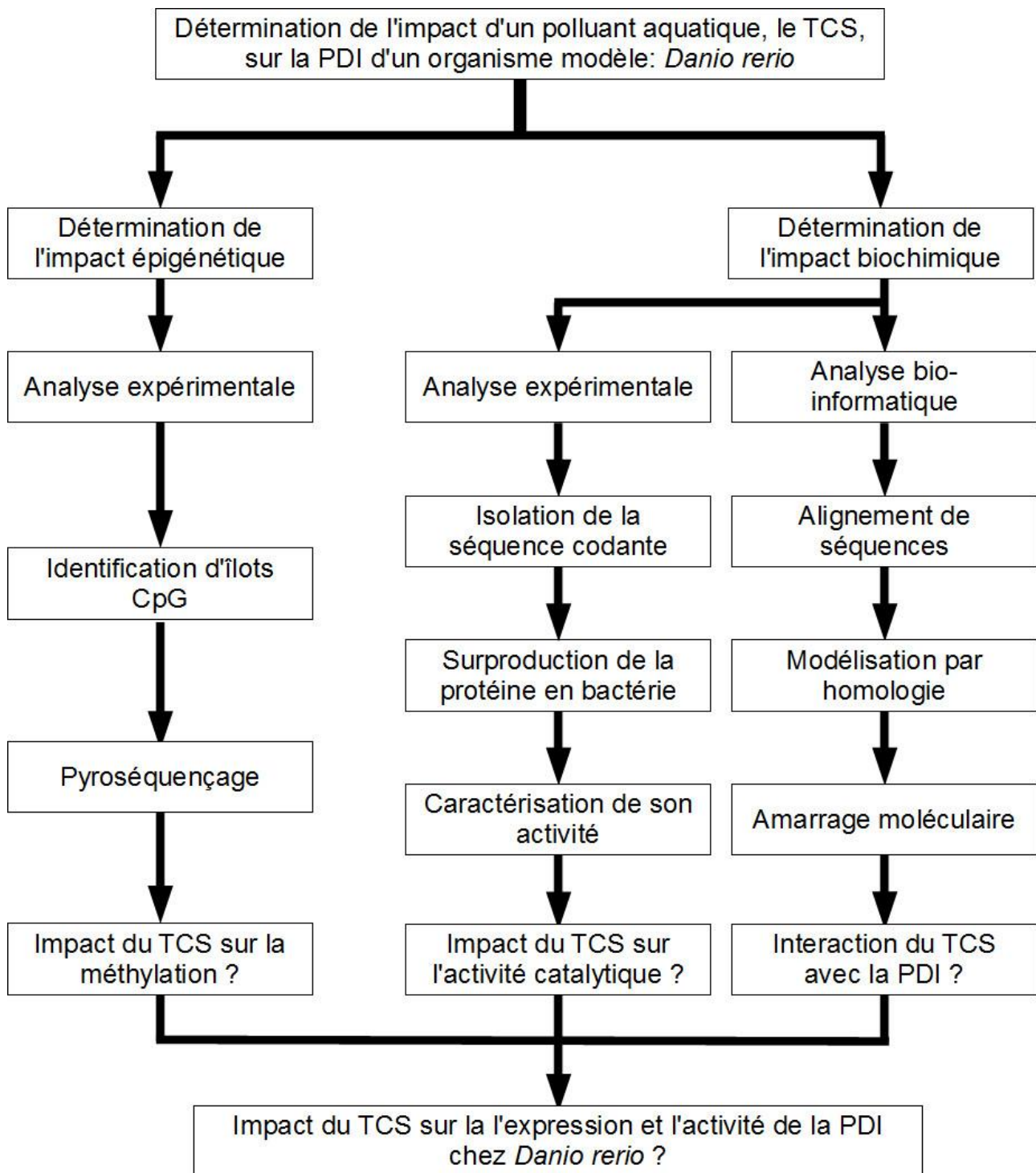


Figure 12: Objectifs du mémoire.

3. Matériel et méthodes

3.1 Isolation et amplification de la séquence codante

3.1.1 Extraction des ARN totaux d'un *Danio rerio*

-Kit: innuSPEED Tissued RNA© Kit (Analytik Jena, Allemagne)

Le matériel génétique a été extrait à partir d'un homogénat total réalisé sur un individu adulte de *Danio rerio*. Les ARN ont ensuite été isolés selon la technique de centrifugation par séparation de phases. Ils ont finalement été élués dans une solution d'eau RNase-free et stockés à -80°C.

3.1.2 Rétrotranscription

-Kit: RevertAid™ Minus First Strand cDNA Synthesis Kit (Fermentas, Allemagne).

-Appareillages: NanoDrop 2000 (ThermoScientific, USA) et TProfessional Standard PCR Thermocycler (Core Life Sciences Inc., USA).

Pour la suite des investigations, il a été nécessaire de rétrotranscrire les ARN totaux en ADN complémentaires (ADNc). La rétrotranscription est basée sur le principe suivant: une reverse transcriptase (ADN polymérase ARN-dépendante) va rétrotranscrire l'ARN en ADNc (ne comportant que des exons, caractéristique importante dans le cas d'une surexpression en bactérie ultérieure). Idéalement, la concentration du matériel génétique de base doit être comprise entre 0.1 ng/μl et 5 μg/μl. Celle-ci a été mesurée par spectrométrie UV/Visible.

1 μl d'ARN totaux a été utilisé auquel on a ajouté 1 μl d'amorces (universelles oligo(dT)₁₈ ou spécifiques), 4 μl de Reaction Buffer 5X, 1μl de RiboLock RNase Inhibitor, 2 μl de dNTP Mix 10 mM et 1 μl de RevertAid H Minus M-MuLV Reverse Transcriptase (200 u/μl). Le mélange a été complété par 12 μl d'eau RNase-free de manière à porter le volume à 20 μl.

Le mélange a été spinné avant d'être incubé pendant 60 minutes à 42°C. La réaction de rétrotranscription a été arrêtée en chauffant le mélange pendant 5 minutes à 70°C. Les ADNc ont été stockés à -20°C.

3.1.3 Conception des amorces

Avant d'amplifier une séquence nucléotidique, il est nécessaire de concevoir les amorces qui serviront à initier la réaction. Dans notre cas, trois «*Polymerase Chain reaction*» ont été nécessaires pour amplifier spécifiquement notre séquence.

Les amorces de la PCR initiale ont été construites à partir d'un alignement de séquences de PDI de plusieurs espèces de poissons disponibles dans la base de données NCBI (NIH, USA): *Takifugu rubripes* (XM_003972125), *Oerochromis niloticus* (XM_003448004) et *Danio rerio* (NM_213364). Celles-ci sont reprises dans le tableau 1.

| Amorces | Séquences |
|---------|-----------------------------|
| F5 | 5'-CTGAAATAGCCGAGGAAGAAG-3' |
| R51 | 5'-CTTACAACCTCATCGTGTCCC-3' |
| R52 | 5'-AGACAACGCCTTACAACCTC-3' |
| R53 | 5'-ACAGGAGACAACGCCTTAC-3' |

Tableau 1: Amorces de la PCR initiale.

Les amorces de la seconde PCR (RACE-PCR) ont été construites à partir de la séquence obtenue grâce à la PCR initiale (tableau 2).

| Amorces | Séquences |
|---------|-----------------------------|
| F1 | 5'-GAATGCCCGTCATCCGCCT-3' |
| F2 | 5'-TAACAGCAGAGAACATAATC-3' |
| F3 | 5'-TTCCTGAAGACTGGGATAAAA-3' |

| | |
|-----|-----------------------------|
| F51 | 5'-TGCCACGTTAKTAGYGGYAG-3' |
| R1 | 5'-CCTCAGCCTGTCGACCAGCA-3' |
| R2 | 5'-GATACCCACGAACACCAAAC-3' |
| R3 | 5'-TCAGCTTTTAGCATTCTGC-3' |
| R4 | 5'-CGACCCTCATCAAACCTTCTT-3' |
| R5 | 5'-CGGTTTAAATGAAGGCCTTA-3' |

Tableau 2: Amorces de la RACE-PCR.

Les amorces, nécessaire à l'amplification finale (FL-PCR) ont été conçues sur base du séquençage des produits de la RACE-PCR (tableau 3).

| Amorces | Séquences |
|---------|-----------------------------|
| FFL | 5'-TGCCACGTTATTAGTGGCAG-3' |
| RFL | 5'-CAGAGAGGTTACTTCTGGACT-3' |

Tableau 3: Amorces de la FL-PCR

L'ensemble des amorces utilisées ont été construites par Eurogentec S.A. (Belgique).

3.1.4 PCR initiale

-Kit: GoTaq® DNA Polymerase (Promega, Allemagne).

-Appareillage: TProfessional Standard PCR Thermocycler (Core Life Sciences Inc., USA)

La PCR, de manière générale, est basée sur la succession de réactions de réplication d'une matrice ADN double brin, grâce à des amorces (*forward* et *reverse*) et une ADN polymérase ADN-dépendante (*Taq* polymérase par exemple). Cette succession de réplifications se réalise via la répétition de cycles de transition de température (dénaturation de la matrice, hybridation avec les amorces, élongation des brins néoformés), permettant une amplification de l'ordre du million d'une séquence spécifique d'ADN.

Trois couples d'amorces (F5-R51, F5-R52, F5-R53) ont été utilisés pour amplifier partiellement l'ADNc codant pour notre séquence. 1µl d'ADNc totaux a été utilisé auquel on a ajouté 10 µl de GoTaq Reaction Buffer 5X, 1.5 µl de dNTP Mix 10 mM, 1 µl de l'amorce *forward* (F) 10 µM, 1µl d'amorce reverse (R) 10 µM et 0.5 µl de GoTaq DNA Polymerase. Le mélange a été porté à 50 µl de volume final en complétant avec 35 µl d'eau stérile.

Le mélange a ensuite été soumis à plusieurs cycles de dénaturations, hybridations et élongations, protocole classique en PCR (tableau 4).

| Étapes | | Temps | Températures |
|-----------------------|--------------|-------------|--------------|
| Dénaturation initiale | | 10 minutes | 95°C |
| Séries de 37cycles | Dénaturation | 1 minute | 95°C |
| | Hybridation | 45 secondes | 56°C |
| | Élongation | 2 minutes | 72°C |
| Élongation finale | | 10 minutes | 72°C |

Tableau 4: Protocole pour la PCR initiale.

Les amplicons ainsi obtenus ont été stockés à -20°C.

3.1.5 RACE-PCR

-Kit: 5'/3' RACE Kit, 2nd Generation (Roche, France).

-Appareillage: TProfessional Standard PCR Thermocycler (Core Life Sciences Inc., USA).

Cette technique particulière a été utilisée afin d'amplifier les extrémités non traduites (UTR= *untranslated regions*) de notre séquence d'intérêt (5'-UTR et 3'-UTR). Il est à noter que les deux extrémités sont amplifiées séparément (notamment en raison d'un traitement spécifique du matériel génétique dans le cas de la 5'-UTR). 9 couples d'amorces ont été utilisés, certains faisant intervenir une amorce universelle (UP): R1-UP, R2-UP, R3-UP, R4-UP, R5-UP, R4-F51, R5-F51, F2-UP et F3-UP.

Contrairement à une PCR «classique», la RACE-PCR nécessite une étape additionnelle avant l'amplification: l'adjonction d'une queue poly-A à l'extrémité 5'-UTR. Dans un premier temps, 19 µl d'ADNc, mélangés à 2.5 µl de Reaction Buffer 10X et 2.5 µl de dATP 2mM, ont été incubés pendant 3 minutes à 95°C. Dans un second temps, le mélange, auquel on a ajouté 1 µl de terminal transferase (80 u/µl), a été incubé pendant 20 minutes à 37°C avant d'être stoppé par chauffage pendant 10 minutes à 70°C.

5 µl d'ADNc (éventuellement traités dans le cas de la 5'-RACE) ont été utilisés pour réaliser la PCR. Le mélange comprenait en outre: 5 µl d'Expand High Fidelity Buffer 10X, 1µl de dNTP Mix 10 mM, 1 µl d'amorce universelle (UP) 10 µM, 1 µl d'amorce spécifique (F ou R) 10 µM et 0.75 µl d'Expand High Fidelity Enzyme Mix, le tout porté à 50 µl par de l'ajout d'eau stérile.

Le mélange a ensuite été soumis à plusieurs cycles de dénaturations, hybridations et élongations (tableau 5).

| Etapas | | Temps | Températures |
|-----------------------|--------------|-------------|--------------|
| Dénaturation initiale | | 10 minutes | 94°C |
| Séries de 37cycles | Dénaturation | 1 minute | 94°C |
| | Hybridation | 45 secondes | 58°C |
| | Élongation | 2 minutes | 72°C |
| Élongation finale | | 10 minutes | 72°C |

Tableau 5: Protocole pour la RACE-PCR.

Les amplicons ainsi obtenus ont été stockés à -20°C.

3.1.6 Full Length-PCR

-Kit: Platinum® Taq DNA Polymerase High Fidelity (Invitrogen, USA).

-Appareillage: TProfessional Standard PCR Thermocycler (Core Life Sciences Inc., USA).

La FL-PCR a été utilisée pour amplifier la totalité de notre séquence d'intérêt. 1.5 µl d'ADNc ont été mélangés avec 5 µl de High Fidelity PCR Buffer 10X, 1.5 µl de MgSO₄, 1 µl de dNTP Mix 10 mM, 1.5 µl de l'amorce FFL 10 µM, 1.5 µl de l'amorce RFL 10 µM, 0.25 µl de Platinum® Taq High Fidelity Polymerase et 37.75 µl d'eau stérile.

Le mélange a ensuite été soumis à plusieurs cycles de dénaturations, hybridations et élongations (tableau 6).

| Etapas | | Temps | Températures |
|-----------------------|--------------|------------|--------------|
| Dénaturation initiale | | 10 minutes | 94°C |
| Séries | Dénaturation | 1 minute | 94°C |

| | | | |
|----------|-------------------|-------------|---------|
| de | Hybridation | 45 secondes | 52-57°C |
| 37cycles | Élongation | 2 minutes | 72°C |
| | Élongation finale | 10 minutes | 72°C |

Tableau 6: Protocole pour la FL-PCR.

Les amplicons ainsi obtenus ont été stockés à -20°C.

3.1.7 Électrophorèse de vérification

Tous les résultats PCR ont été vérifiés par électrophorèse sur gel d'agarose.

L'électrophorèse sur gel d'agarose est basée sur la séparation des acides nucléiques sous l'effet d'un champ électrique (chargés négativement, ils migrent vers l'électrode positive). Cette séparation s'effectue à travers une matrice d'agarose: les molécules de plus petite taille se déplacent plus rapidement et migrent plus loin que les molécules de taille supérieure (un étalon permet d'évaluer les tailles des amplicons). Cette technique permet de vérifier si la PCR a bien amplifié la séquence cible.

12 µl de matériel PCR marqué (10 µl d'amplicons et 2 µl de Loading Dye), ainsi que 5 µl d'étalon (Ladder 100 bp), ont été déposés dans les puits d'un gel composé de 300 mg d'agarose, de 25 ml de tampon TAE 0.5X et de 5 µl de SYBR Safe.

Les amplicons marqués ont été laissés à migrer pendant 45 minutes sous une différence de potentiel de 100 V. Les résultats ont été visualisés sous UV.

3.1.8 Stockage et clonage en bactéries

-Kit plasmide: pGEM®-T Easy Vector System I (Promega, Allemagne).

-Kit cellules compétentes: Clonable™ Kit (Novagen, Allemagne).

Après vérification, les produits PCR ont été insérés dans un plasmide afin d'être stockés en bactéries. Ces organismes sont utilisés dans le clonage de gènes pour deux raisons: leur vitesse de multiplication (toutes les 20 minutes à 37°C chez *Escherichia coli*) et la sélection aisée de bactéries recombinantes. Le clonage est basé sur le principe de la transformation bactérienne. La séquence d'ADN cible est insérée dans un plasmide (ADN circulaire non chromosomique) suite à un choc thermique. Les bactéries possédant le plasmide sont sélectionnées grâce à la présence d'un gène de résistance à un antibiotique constitutivement présent sur ce plasmide. Dans certains cas, il est possible de sélectionner les clones recombinants grâce au test blanc/bleu.

Dans un premier temps, les produits PCR a été purifié. 20 µl d'amplicons ont été mélangés à 250 µl de Binding Buffer, transférés sur une colonne (dans un tube eppendorf) et centrifugés pendant 3 minutes à 12000 rpm. Ensuite, 10 µl d'Elution Buffer ont été ajoutés au culot, laissé à incuber pendant 1 minute à température ambiante avant une centrifugation finale d'une minute à 10000 rpm.

Dans un deuxième temps, la séquence d'intérêt a été insérée dans le vecteur pGEM. Le mélange de ligation, constitué de 3 µl de produit PCR purifié, de 5 µl de Rapid Ligation Buffer 2X, de 1 µl de vecteur pGEM-T Easy et de 1 µl de T4 DNA Ligase (3 u/µl), a été incubé pendant 1 heure à température ambiante avant la transformation.

Dans un troisième temps, la construction pGEM-PDI a été transfectée à la souche d'*E.coli* NovaBlue Singles™ Competent Cells. 1 µl du produit de ligation a été ajouté à 50 µl de cellules compétentes sur glace, avant un choc thermique de 30 secondes à 42°C. Redéposées sur la glace

durant 2 minutes, les bactéries transformées ont été mises à croître dans 250 µl de milieu SOC liquide pendant 1 heure à 37°C. Finalement, 90 µl ont été prélevés pour être déposés sur un milieu LB-Agar-Ampiciline (50 µg/ml)-X-Gal et laissés à incuber une nuit à 37°C.

Grâce à la présence de X-Gal dans le milieu et de l'opéron *lacZ* sur le plasmide, les bactéries ayant intégré un plasmide avec notre séquence d'intérêt apparaissaient blanches et ont été sélectionnées pour les manipulations suivantes. L'insertion effective de notre séquence dans pGEM a été vérifiée par PCR en bactéries et séquençage.

3.1.9 Extraction plasmidique

-Kit: GenElute™ Five-Minute Miniprep Kit (Sigma, USA).

Pour envoyer au séquençage une construction plasmidique, il est nécessaire d'extraire les plasmides des clones bactériens.

Pour ce faire, une dizaine de colonies effectivement transformées (blanches) ont été prélevées et mises à croître dans une solution de 5 ml de LB-Ampicilin (50 µg/ml).

400 µl de culture liquide ont été lysés par ajout de 40 µl de Lysis Solution et incubation pendant 2 minutes à température ambiante. On a ensuite rajouté 400 µl de Binding Solution avant de prélever 780 µl qui ont été déposés sur une colonne (tube eppendorf). Celle-ci a été spinée durant 20 secondes. Le culot a été rincé avec 700 µl de Wash Solution avant d'être spiné à nouveau pendant 20 secondes. Un nouveau rinçage a été effectué (200 µl de Wash Solution, 30 secondes). Le matériel génétique a finalement été élué dans 40 µl d'Elution Solution et spiné pendant 30 secondes. L'éluât a été envoyé au séquençage.

3.1.10 Séquençage

Le séquençage du matériel génétique a été confié à la firme MacroGen Inc. (USA) ainsi qu'à la firme Genoscreen Inc. (France). La séquence ainsi obtenue a ensuite été identifiée à l'aide du protocole BLAST disponible dans la base de données NCBI (NHI, USA).³⁸

3.2 Analyse bio-informatique

3.2.1 Analyse de séquences

Les solutions de fragments d'ADN obtenues par PCR ou par extraction plasmidique ont été envoyées à séquencer à la société MacroGen Inc. (Pays-Bas). L'analyse de ses résultats a été effectuée *in silico*, tous les fragments ayant été comparés entre eux (ClustalW2, EMBL-EBI, Royaume-Uni) et/ou par rapport à la base de données de NCBI en réalisant un BLAST (*Basic Local Alignment Search Tool*) (NIH, USA), permettant de déterminer si la séquence a bien été amplifiée. Le programme d'alignement ClustalW2 a permis de produire un alignement multiple des séquences de la PDI humaine, de rat et du poisson zèbre.^{38,39}

3.2.2 Modélisation par homologie

Plusieurs méthodes de modélisation, se basant sur la séquence en acides aminés de la protéine, sont disponibles: la prédiction *de novo*, la reconnaissance de *fold* ou la modélisation par homologie. Cette dernière est employée pour prédire la structure 3D d'une protéine ayant plus de 30% d'identité de séquence avec une structure obtenue par cristallographie ou RMN. Cette méthode se base sur l'alignement des deux séquences pour construire la structure générale de la chaîne

polypeptidique, modélisant par la suite les chaînes latérales de chaque acide aminé et optimisant ensuite le modèle créé. Le modèle a été validé par la construction d'un diagramme de Ramachandran avec le programme Molprobity (Harvard University, USA).⁴⁰

Ici, EsyPred3D a été utilisé pour modéliser la structure 3D de la PDI de *Danio rerio* par homologie avec des structures cristallographiques de la PDI humaine se trouvant dans la base de données protéique (PDB ; RCSB, USA).⁴¹ La particularité de ce programme est qu'il se base sur une méthode d'alignement consensus utilisant plusieurs programmes d'alignement, l'alignement étant une étape clé pour la modélisation par homologie. MacPyMol a été utilisé pour visualiser la structure obtenue et pour la comparer par superposition avec les structures de la PDB.⁴²

3.2.3 Amarrage moléculaire (Docking)

Il s'agit d'une méthode de modélisation permettant de prédire l'interaction entre deux molécules et de calculer à l'aide d'algorithmes leurs orientations spatiales l'une par rapport à l'autre. Cette technique de bio-informatique est souvent utilisée en pharmacologie afin de déterminer l'interaction entre une drogue et une protéine cible. De plus, cette méthode présente l'avantage de pouvoir donner un grand nombre de conformations possibles et de les classer, permettant ainsi de déterminer facilement les structures étant les plus intéressantes. Le programme choisi pour réaliser l'amarrage est GOLD (Cambridge, Royaume-Uni) qui utilise la méthode d'algorithme génétique pour générer ses solutions. Il présente comme avantage de pouvoir réaliser un *docking* en tenant compte de la flexibilité du ligand et partiellement de celle de la protéine.⁴³

3.3 Analyse de la méthylation

3.3.1 Extraction de l'ADN total

-Kit: QIAmp MiniElute (Quiagen, Allemagne).

-Appareillage: NanoDrop 2000 (ThermoScientific, USA).

Le matériel génétique a été extrait par la digestion d'une larve de *Danio rerio*.

Dans un premier temps, la larve a été lysée par ajout d'un tampon de lyse (180 µl d'ATL Buffer et 20 µl de protéinase K) et laissée à incuber pendant une nuit à 56°C.

Dans un second temps, l'ADN a été récupéré. Le lysat a été mélangé avec 200 µl d'AL Buffer et 1 µl de RNA Carrier avant d'être vortexé 15 secondes. 200 µl d'éthanol absolu froid ont ensuite été ajoutés et le mélange a été laissé à incuber 5 minutes à température ambiante. Après centrifugation rapide, le lysat a été déposé sur colonne (tube eppendorf) et centrifugé à 8000 rpm pendant 90 secondes. Le culot a ensuite été lavé deux fois par ajout d'AW Buffer et centrifugation (200 µl d'AW1 et 500 µl d'AW2). Une centrifugation à 14000 rpm pendant 7 minutes a permis l'élimination des résidus du tampon de lavage. Le matériel génétique a finalement été élué dans 30 µl d'AE Buffer avant incubation pendant 1 minute à température ambiante et une double centrifugation à 14000 rpm pendant 90 secondes. Une solution concentrée en ADN totale d'une larve a été obtenue.

L'ADN extrait a été dosé par spectrométrie (NanoDrop 2000). Le matériel obtenu a ensuite été stocké à -20°C.

3.3.2 Traitement

-Kit: EZ DNA Methylation-GOLD™ (Zymo Research, USA).

-Appareillage: TProfessional Standard PCR Thermocycler (Core Life Sciences Inc., USA).

Afin d'analyser le degré de méthylation de la séquence d'intérêt, il est nécessaire de lui faire subir un traitement bisulphite.

Pour ce faire, 400 ng d'ADN sont prélevés, auxquels on ajoute 130 µl de CT *conversion reagent*. Le mélange a ensuite été chauffé pendant 10 minutes à 98 °C et pendant 3 heures à 64°C. Cette étape de dénaturation de l'ADN permet le remplacement des cytosines non-méthylées de l'ADN en uraciles par réaction avec le bisulfite de sodium présent dans la solution. 600 µl de M-Binding Buffer ont été ajoutés avant de déposer le mélange sur la colonne. La colonne a ensuite été centrifugée à 14000 rpm pendant 30 secondes. Le culot a été lavé avec 100 µl de M-Wash Buffer (centrifugation à 14000 rpm pendant 30 secondes). 200 µl de M-Desuphonation Buffer ont été ajoutés au mélange et ont été laissés à incuber 15 à 20 minutes à température ambiante. Cette étape permet l'alkylation des cytosines. Finalement, le matériel traité a été doublement rincé avec 200 µl de M-Wash Buffer et élué dans 10 µl de M-Elution Buffer avant centrifugation à 14000 rpm pendant 30 secondes.

L'ADN ainsi traité a été stocké à -20°C.

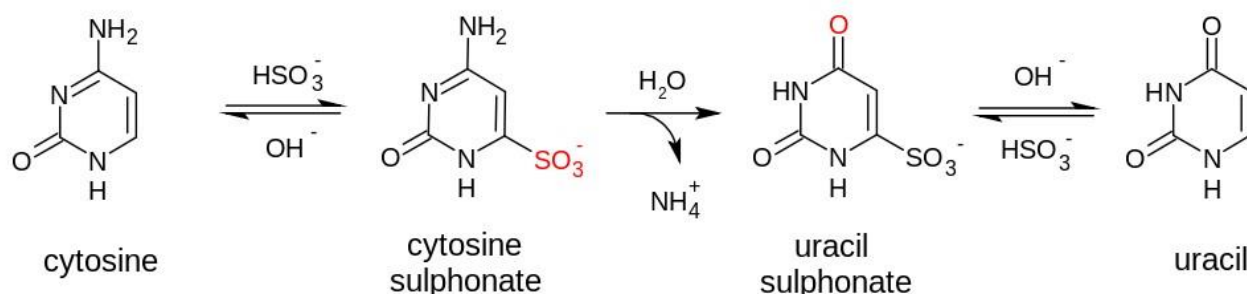


Figure 13: représentation de l'étape de traitement de l'ADN au bisulphite. Les cytosines ne présentant pas de fonction méthyle, réagissent en présence de bisulphite de sodium et forme de l'uracile.

3.3.3 Conception des amorces

Dans le cas d'un pyroséquençage, il est nécessaire de concevoir deux types d'amorces: un couple d'amorces permettant l'amplification de la séquence d'intérêt par PCR et une amorce employée pour l'étape de séquençage. Parmi les 2 amorces employées pour amplifier la séquence d'intérêt, l'une de celle-ci est biotinyllée à son extrémité 5'. Cette étape est nécessaire pour permettre l'étape du pyroséquençage, afin d'assurer la fixation de l'ADN sur des billes de sépharose. Le choix des séquences ainsi que des amorces a été réalisé à l'aide du programme Pyromark Assay Design 2.0 (Quiagen, Allemagne).

Ce programme permet de localiser des dinucléotides CpG et de convertir les bases cytosines, non suivie par une guanine, en thymidine pour une séquence de 10 kb maximum. Dans notre cas, la séquence complète du gène de la PDI de *Danio rerio* (obtenue sur NCBI) a été analysée et 3 régions de 100 à 200 nucléotides ont été sélectionnées. Les amorces construites avec le logiciel et ayant obtenu le plus haut score (dépendant des possibilités de chevauchement et de la température d'hybridation) sont reprises dans le tableau 7.

| Amorces | | Séquences | Score |
|-----------|-----|------------------------------------|-------|
| PDI4hb_12 | F12 | 5'-TGGAGTAAAGAGGAGTAGGTTAATAGTG-3' | 63 |
| | R12 | 5'-CTATAATATACTTTATCTACCTCATTC-3' | |

| | | | |
|-----------|-----|--------------------------------------|----|
| | S12 | 5'-TGGGAGTTTAGGTTAGGT-3' | |
| PDI4hb_22 | F22 | 5'-AGGGTGATTGTATGTAAAAAGATTT-3' | 58 |
| | R22 | 5'-CCAAAACCAATATCCTTTCCACAATCACTT-3' | |
| | S22 | 5'-TGGTTTATGTGAGATTAGAAAATT-3' | |
| PDI4hb_34 | F34 | 5'-GGTGTGATATTGGAAAGGTAGT-3' | 61 |
| | R34 | 5'-AACCATCCTAAATAAATTACACTACC-3' | |
| | S34 | 5'-GGTAGTTAAGGGGTGTGT-3' | |

Tableau 7: Amorces amplifiant les 3 régions CpG

Les séquences sélectionnées dans les 3 régions ainsi que les séquences amplifiées par chaque couple d'amorces après le traitement bisulphite et la séquence théoriquement obtenue par le pyroséquençage sont reprises dans le tableau 8.

| | Séquences | | | | |
|------------------------|--|--|--|--|--|
| PDI4hb_12 (F12-R12) | TGGAGCAAAG ATCAATAGCC CGGGCCGGGC CGTGTTTCCA CCCACAGAAC AATGAGGCAG | AGGAGTAGGT GCATTTCCAC CAATAGCCTG CTGTTGGGCT ATCCCCGAAT ATAAAGCATA | CAACAGTGGG TATCGGGCCA GGAGGTTGAG AGTAGCTCAC CAAACGTCAC CCATAG | CGTTCAGGCT GTGCGAGCCA AAATGAGGCC AGCGCGTCAC AACCCGCCCA CTCCAGCGGG | AGGCCCTGGG GGGCTTTAAT GAAATCATGT GCAAAACCCG CTCCAGCGGG |
| Traitée au bisulphite | TGGAGTAAAG ATTAATAGTY YGGGYGGGT YGTGTTTTTA TTTATAGAAT AATGAGGTAG | AGGAGTAGGT GTATTTTAT TAATAGTTG TTGTTGGGTT ATTTYGAAT ATAAAGTATA | TAAAGTGGG TATYGGGTTA GGAGGTTGAG AGTAGTTAT TAAAYGTAT TTATAG | YGTTAGGTT GTGYGAGTTA AAATGAGGTY AGYGYGTAY AATTYGTTTA TATTAGYGGG | AGGTTTGGG GGCTTTAAT GAAATCATGT GTAATAATY TTTATAGYGGG |
| Séquencée (S12) | T TTGGGATTA TTTAATYGGG TTATGTYGTG ATATYGTFTA AGYGGGAATG | ATAGTYGTAT TYGGGTTAAT TTTTTATTGT TAGAATATTT AGGTAGATAA | TTTTATTATY AGTTTGGGAG TGGGTTAGTA TYGAATTAAA AGTAT | GGGTTAGTGY GTTGAGAAAAT GTTTATAGYG YGTATAAAT YGTATAAAT | GAGTTAGGGT GAGGTYGAAA YGTAYGTAA YGTATTTT YGTATTTT |
| PDI4hb-22 (R22-F22) | AGGGTGATTG TTCCCTTATT CTCCTTTCAC ATTCGGCGAC TGGCCCGATT | TATGCAAAAA TATTTGTTT TCCGCCCGA CTGAAAAAGG AAACCCCGAA | GACCCAGAC GCTCATGTGC AGCCCCAGCT CCGGCCGCTG AGTGACTGTG | TATAAGGCCA GATCAGAAAA GGCCCTCCTT GCCCCAAGAA GAAAGGACAC | GATGTTTTCT CTACCCGCCT GGCCCAAGGT AGCCCCGCTT TGGCCTTGG |
| Traitée au bisulphite | AGGGTGATTG TTTTTTTATT TTTTTTTTTAT ATTYGGYGAT TGGTTYGATT | TATGTAAAAA TATTTGTTT TYGTTYGA TTGAAAAAGG AAATTTYGAA | GATTTTAGAT GTTTATGTGY AGTTTTAGTT TYGGTYGTTG AGTGAATTGTG | TATAAGGTTA GATTAGAAAA GGTTTTTTTT GTTTTAAGAA GAAAGGATAT | GATGTTTTTT TTATTYGTTT GGTTTAAGGT AGTTYGTTT TGCTTTTGG |
| Séquencée (S22) | ATTYGTTTTT TTTTAAGGTAT TTYGTTTTG GTT | TTTTTTATTT TYGGYGATTT GTT | YGTTTYGAAG GAAAAGGTY | TTTTAGTTGG GGTYGTTGGT | TTTTTTTGG TTTAAGAAAG |
| PDI4hb_34 (F34-R34) | GGTGTGACAT CCCACCCCGA ATCGTGTAAC TGCTTTCTGG | TGGAAAGGTA CACAGGATCG GTTAAAGCCC CAGATGTTTA | GTCACGGGGT TAGATCTGAG TAGTGACCAC GGTCACTTTT | GTGTCGCGAT TGTCGCGATT TGACCAGTTC GAATGCTGGT | TCTCCCTTAT TCGATTTTAT TTAGTTCCCTA GGTACTTTTA |

| | |
|-----------------------|---|
| | CTCTAGCGGT AGCTTCAAGA TCGAGTCCGC AACCCACACA AGTGGCTCAG GTAGTGCAAT TCATCCAGGA TGGCC |
| Traitée au bisulphite | GGTGTGATAT TGGAAAGGTA GTTAYGGGGT GTGTYGYGAT TTTTFTTAT TTTATTTYGA TATAGGATYG TAGATTTGAG TGYGYGATT TYGATTTAT ATYGTGTAAY GTTAAAGTTT TAGTGATTAT TGATTAGTTT TTAGTTTFTA TGTTTTTTGG TAGATGTTTA GGTATTTTTT GAATGTTGGT GGTATTTTTA TTTTAGYGGT AGTTTAAAGA TYGAGTTYGT AATTTATATA AGTGGTTTAG GTAGGTAAT TTATTTAGGA TGGTT |
| Séquencée (S34) | YGYGATTTTT TTTTATTTTA TTTYGATATA GGATYGTAGA TTTGAGTGTY GYGATTTYGA TTTTATATYG TGTAAYGTTA AAGTTT TAGT GATTATTGAT TAGTTTTTAG TTTTATGTT TTTTG |

Tableau 8: Séquences correspondant aux 3 îlots CpG sélectionnés.

3.3.3 Amplification

-Kit: PyroMark PCR kit (Quiagen, Allemagne).

-Appareillage: TProfessional Standard PCR Thermocycler (Core Life Sciences Inc., USA).

Pour l'amplification des régions sélectionnées, la PCR a été réalisée à partir de 1 µl d'ADN totaux traité au bisulphite qui a été testé avec 3 conditions de mix différentes (tableau X9). Ici la présence de sel (MgCl₂) et d'un marqueur de l'ADN (Coral Loading Dye) ont été testées par rapport à 2 solutions ne contenant pas le marqueur (l'une des deux possédant du MgCl₂). Ces mix permettent de déterminer l'importance du MgCl₂, qui est un cofacteur pour la Taq polymérase et de déterminer si la présence d'un marqueur de l'ADN influence la quantité d'amplicons obtenus.

| Mix | Composition |
|-----|---|
| A | 12.5 µl de Master Mix, 2 µl d'amorces (12, 22 ou 34), 1 µl de MgCl ₂ , 2.5 µl de Coral Loading Dye, 6 µl d'eau stérile |
| B | 12.5 µl de Master Mix, 2 µl d'amorces (12, 22 ou 34), 1 µl de MgCl ₂ , 8.5 µl d'eau stérile |
| C | 12.5 µl de Master Mix, 2 µl d'amorces (12, 22 ou 34), 9.5 µl d'eau stérile |

Tableau 9: Mélanges pour chacun condition (A, B et C).

Comme pour toute PCR, le mélange a été soumis à plusieurs cycles de dénaturations, hybridations et élongations (tableau 10).

| Étapes | | Temps | Températures |
|--------------------------|--------------|-------------|--------------|
| Dénaturation initiale | | 15 minutes | 95°C |
| Séries de 45cycles | Dénaturation | 30 secondes | 94°C |
| | Hybridation | 30 secondes | 56°C |
| | Élongation | 30 secondes | 72°C |
| Élongation finale | | 10 minutes | 72°C |

Tableau 10: Protocole de la PCR.

La vérification de la PCR est réalisée par électrophorèse sur gel d'agarose (voir 3.1.7). Les amplicons ainsi obtenus ont été stockés à -20°C.

3.3.4 Pyroséquençage

-Kit: PyroMark Q24 Advanced CpG Reagent (Quiagen, Allemagne).

-Appareillage: PyroMark Q24 Advanced et Vacuum Workstation (Quiagen, Allemagne).

Les amplicons ont été séquencés à l'aide des amorces précédemment construites, permettant à partir de ce résultat le calcul du pourcentage de méthylation des cytosines présentes dans la séquence amplifiée.

Dans un premier temps, 15 µl d'amplicons ont été mélangés avec 1 µl de Streptavidin sepharose HP, 40 µl de PyroMark Binding Buffer et 24 µl d'eau stérile. Le mélange a été vortexé à température ambiante durant 10 minutes à 1400 r.p.m. Cette étape permet la fixation des fragments d'ADN sur des billes de sépharose grâce à la présence de biotines sur l'une des amorces d'amplification.

La seconde étape a été de remplir la cartouche du PyroMark Q24 avec les solutions d'enzymes, de substrats, et de nucléotides (dATP, dGTP, dCTP, dTTP), les quantités ayant été ajoutées en fonction du nombre d'échantillons réalisés. Ces informations ont été fournies par le programme PyroMark Q24 Advanced.

La dernière étape a été de traiter l'ADN par une série de bains. Elle a consisté en l'immobilisation sur un filtre de l'ADN (fixé à des billes de sépharose) par aspiration. Le filtre a ensuite été placé dans des solutions d'éthanol (70%) et de solution de dénaturation permettant le traitement de l'ADN et a été lavé grâce à une solution de lavage. Les billes ont été relâchées par arrêt de l'aspiration et par immersion du filtre dans 20 µl d'amorces (0.375 µM) ayant été préalablement déposé sur une plaque pyromark de 24 puits. Après incubation de la plaque durant 5 minutes à 80°C, elle a été placée dans le PyroMark Q24 Advanced.

3.3.5 Analyse du séquençage

Le logiciel PyroMark Q24 Advanced a permis de réaliser l'analyse du séquençage ainsi que de programmer le protocole suivi par le PyroMark Q24 Advanced. Cette étape préalable a consisté à charger la séquence pour permettre son authentification par le logiciel. La séquence chargée a pu être modifiée pour y ajouter un faux négatif (ajout d'une base cystéine avant une base thymidine), permettant ainsi la vérification de l'étape du traitement bisulphite. Le protocole enregistré sur une clé USB a été insérée sur le PyroMark Q24 Advanced pour permettre le lancement du séquençage ainsi que la sauvegarde des résultats générés.

4. Résultats

4.1 Partie biochimique

L'objectif de cette première partie est d'étudier du point de vue moléculaire la structure et l'activité de la PDI de *Danio rerio*. D'une part, nous avons isolé et cloné le gène de la PDI afin de la surexprimer et étudier son activité enzymatique. D'autre part, nous avons modélisé sa structure tridimensionnelle afin d'investiguer les interactions entre notre protéine et certains polluants.

4.1.1 Isolation et amplification de la séquence codante

Sur les 3 couples d'amorces testés lors de la PCR initiale, seul le couple F5-B51 a permis l'amplification significative d'une séquence d'approximativement 1.5 kb (Fig.14). Suite au séquençage, celle-ci s'avère correspondre au fragment central de la PDI. Les amorces forward (F1,

F2, F3 et F51) et reverse (R1, R2, R3, R4 et R5) de la RACE-PCR ont été construites à partir de ce résultat.

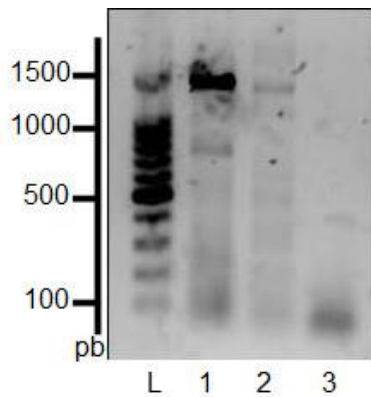


Figure 14: Gel d'électrophorèse de la PCR initiale (L = étalon, 1 = F5-B51, 2 = F5-B52 et 3 = F5-B53).

Les 2 couples d'amorces de la 3'-RACE-PCR ont généré des amplicons dont la taille est comprise entre 1 et 1.5 kb (Fig.15), pour une taille attendue estimée à 1.3 kb. Le séquençage fait correspondre ces fragments à la région 3'-UTR de la PDI et a permis la construction de l'amorce reverse de la FL-PCR (RFL).

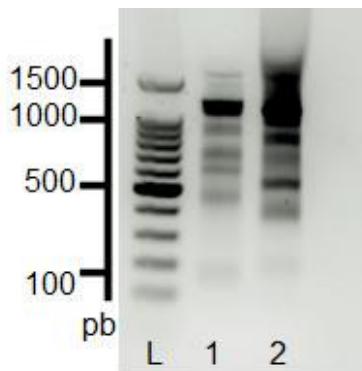


Figure 15: Gel d'électrophorèse de la 3'-RACE-PCR (L = étalon, 1 = F2-UP, 2 = F3-UP).

La 5'-RACE-PCR a nécessité l'emploi de 7 couples d'amorces différents (utilisés de manière successive). Seul le couple F51-R5 a permis de générer des amplicons correspondant, après séquençage, à la région 5'-UTR (Fig.16). La séquence ainsi obtenue a servi à la construction de l'amorce forward de la FL-PCR (FFL).

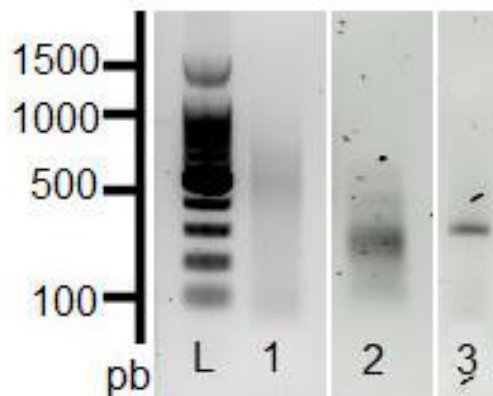


Figure 16: Gel d'électrophorèse de la 5'-RACE-PCR (L = étalon, 1 = R2-UP, 2 = R3-UP, 3 = F51-R5).

La FL-PCR a été réalisée avec un seul couple d'amorces. Les amplicons générés présentent une taille supérieure à 1.5 kb (Fig. 17A), correspondant potentiellement à la totalité de notre séquence d'intérêt (1.7 kb). Ces amplicons ont été ligués dans un plasmide, lui-même inséré en bactéries (*E. coli*). La figure 17B illustre la PCR en bactéries. Deux paires d'amorces ont été utilisées: FFL-RFL, qui sont les amorces spécifiques de notre séquence, et T7-SP6 qui sont les amorces universelles du plasmide. Le séquençage de ces amplicons est actuellement en cours.

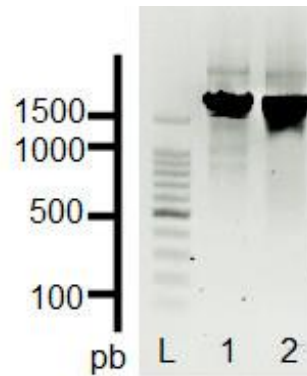


Figure 17A: Gel d'électrophorèse de la FL-PCR (L = étalon, 1 = GoTaq polymérase, 2 = Platinum Taq polymérase).

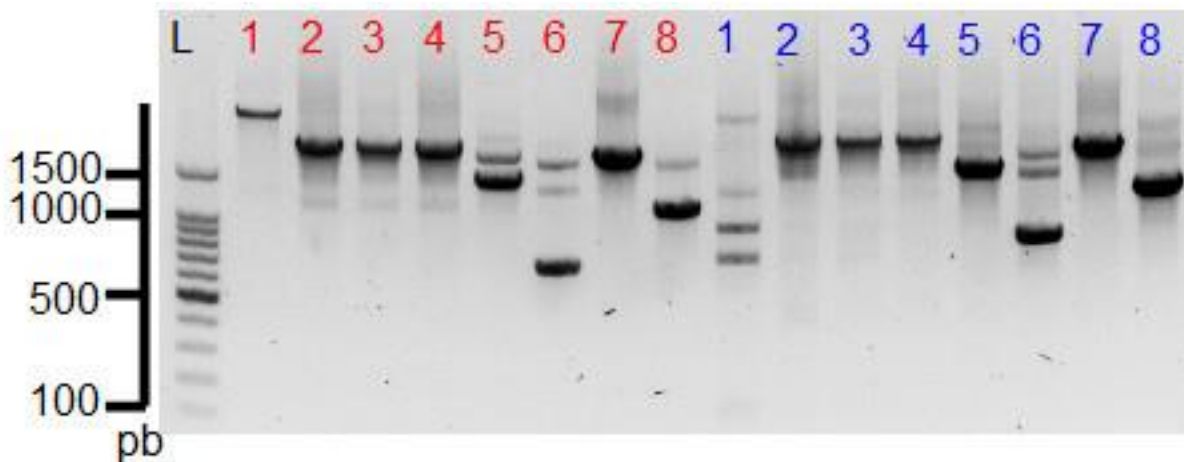


Figure 17B: Gel d'électrophorèse de la PCR en bactérie (L = étalon, 1-8 = colonies, rouge = FFL-RFL, bleu = T7-SP6).

4.1.2 Analyse de la séquence peptidique

Afin d'identifier les acides aminés conservés, un alignement de séquences a été réalisé entre la séquence de la PDI de *Danio rerio*, celle de l'humaine et celle du rat (Fig.18). Nous avons choisi ces deux organismes comme modèle car ils sont les deux seuls dont la structure et/ou l'activité de la PDI a été bien caractérisée (voir précédemment). Les pourcentages d'identité de séquences sont de 71,94 % et de 72.02% entre celle de *Danio rerio*, et celle de l'homme et du rat, respectivement. On constate que dans les PDI-a et -a', le motif CXXC (CGHC, représenté en bleu) est conservé entre

les 3 espèces ainsi que le résidu R120 (représenté en bleu) identifié par Lappi *et al.* (2004-2010) comme essentiel pour l'activité catalytique.^{25,26} Le motif d'adressage au RE (KDEL) est également présent chez *Danio rerio* à l'extrémité de la séquence peptidique. Au niveau du PDI-b', les résidus identifiés par Nguyen *et al.* (voir 1.3.6a) comme intervenant dans la liaison au substrat sont identiques chez les trois organismes.²⁸ Cela laisse suggérer que le mode de liaison des substrats ou des inhibiteurs dans le PDI-b' de ces trois espèces sera le même. Les régions présentant le moins d'identité de séquences entre les trois espèces sont celles du peptide signalé (20 premiers résidus) et la fin du domaine C terminal (20 derniers résidus).

```

Homme      MLRRALLCLAVA--ALVRDAPEEEDHVLVLRKSNFAEALAAHKYLLVEFYAPWCGHCKA 58
Rat       MLRRALLCLALAWAARVGADALEEEDNVLVLRKSNFAEP-AAHNYLLVEFYAPWCGHCKA 59
Danio     MLKFLIVCALAA---MSAAEIAEEEDVLVLRKSNFEEALKAHPNVLVEFYAPWCGHCKA 56
          **  :.*  *  :  *  ***:*****:**** *  **  :*****:*****
Homme      LAPEYAKAAGKLKAEGSEIRLAKVDATEESDLAQYGVRGYPTIKFFRNGDTASPKEYTA 118
Rat       LAPEYAKAAAKLKAEGSEIRLAKVDATEESDLAQYGVRGYPTIKFFKNGDTASPKEYTA 119
Danio     LAPEYSKAAGMLKAEGSDIRLAKVDATEESELAQEFGVRGYPTIKFFKGGEKGNPKEYSA 116
          *****:*** *****:*****:***** ***:*****:***** *****:*****
Homme      GREADDIVNWLKRTGPAATTLPDGAAESLVESSEVAVIGFFKDVESDSAKQFLQAAEA 178
Rat       GREADDIVNWLKRTGPAATTLSDTAAESLVDSSEVTVIGFFKDAGSSAKQFLLAEA 179
Danio     GRQAEDIVSWLKKRTGPAATTLNDVMQAESIIADNEVAVIGFFKDVESEDSKAFIKTAEA 176
          ***:*** ***** ***** * ***: ..* ***** ***:*** :***
Homme      IDDIPFGITSNSDVFSKYQLDKDGVVLFKKFDEGRNNFEGEVTKENLLDFIKHNQLPLVI 238
Rat       VDDIPFGITSNSDVFSKYQLDKDGVVLFKKFDEGRNNFEGEITKELLDFIKHNQLPLVI 239
Danio     VDDIPFGITSDSVFAKFEVAKDSVLFKKFDEGRNTFDGEVSKESLLNFIKANQLPLVI 236
          *****:..* *****: ***: *** *****:*** ***: *** ***** *****
Homme      EFTEQTAPKIFGGEIKHILLFLPKSVSDYDGKLSNFKTAESFGKGILFIFIDSDHTDN 298
Rat       EFTEQTAPKIFGGEIKHILLFLPKSVSDYDGKLSNFKKAEGFGKGILFIFIDSDHTDN 299
Danio     EFTEQTAPKIFGGDIKSHILMFVPKAAKDFQDKMDQFKKAEGFGKGILFIFIDSDVDDN 296
          *****:*** ***: ***: ***: ..* :* :* *** ***** ***** ***
Homme      QRILEFFGLKKEECPAVRLITLEEEMTKYKPESELTAERITEFCHRFLEGKIKPHLMSQ 358
Rat       QRILEFFGLKKEECPAVRLITLEEEMTKYKPESELTAEKITQFCHHFLEGKIKPHLMSQ 359
Danio     QRILEFFGLKKEECPVIRLITLEEEMTKYKPESSEITAENIISFCTSFVEGTLKPHLMSQ 356
          *****:*** *****:***** ***: *** * ** ***: *** *****
Homme      ELPEDWDKQPVKVLVGKNFEDVAFDEKNVFEFYAPWCGHCQLAPIWDKLGETYKDHE 418
Rat       ELPEDWDKQPVKVLVGKNFEEVAFDEKNVFEFYAPWCGHCQLAPIWDKLGETYKDHE 419
Danio     DIPEDWDKNPVKVLVGKNFEEVAFNPANNVFEFYAPWCGHCQLAPIWDQLGEKFKDNA 416
          :*** *** *****:***: *****:***** ***: ***:
Homme      NIVIAKMDSTANEVEAVKVHSFPTLKFFPASADRTVIDYNGERTLDGFKKFLESGGQDG- 477
Rat       NIVIAKMDSTANEVEAVKVHSFPTLKFFPASADRTVIDYNGERTLDGFKKFLESGGQDG- 478
Danio     NIVVAKMDSTANEIEAVKVHSFPTLKFFPAGDERKVIDYNGERTLDGFTKFLESGGKEGG 476
          ***:***** *****:***** ***: *****:***** *****:***** :*
Homme      --AGDDDDLEDEEAEEPDMEEDDDQAVKDEL 508
Rat       --AGDNDLD-LEEALEPDMEEDDDQAVKDEL 508
Danio     APVGDEEDEADDLDDAEEDSDVEEG HDEL 509
          .***:*** :* :* :* ***: :*

```

Figure 18: Alignement de séquences peptidiques de la protéine P4HB (ou PDIA1) de l'homme, du rat et du poisson zèbre. Les différents domaines sont représentés: le domaine a (rouge), b (cyan), b' (jaune), a' (vert) et c avec le x-linker en gris. Le symbole a) « * » indique les résidus identiques b) « : » indique les résidus ayant les mêmes propriétés (hydrophobe, aromatique, acide,...), c) « . » indique les résidus ayant le même encombrement stérique.

4.1.3 Modélisation de la PDI de *Danio rerio* par homologie

Afin de modéliser la structure 3D de la PDI de *Danio rerio*, nous avons utilisé les deux structures cristallographiques, oxydée (code PDB: 4EL1) et réduite (code PDB: 4EKZ), de la PDI humaine qui possède plus de 30% d'identité de séquence avec celle du poisson zèbre (68.7%). Ces deux structures ont l'avantage de représenter les 4 domaines majeurs de la PDI (a, b, b' et a').

Le modèle 3D de la PDI de *Danio rerio* a été construit à l'aide du programme EsyPred3D (Fig.19). Ce dernier et la structure de la PDI humaine se superpose très bien puisque la distance moyenne entre les atomes (ou RMSD, *root-mean-square-deviation*) est de 1.2 Å (Fig.19A). Comme attendu au vu des analyses réalisées chez l'homme (voir 1.3.6b), les formes oxydée et réduite de la PDI de *Danio rerio* sont différentes (RMSD=3.3Å) (Fig.19B). On peut voir de manière plus précise sur la figure 19C cette différence de conformation dans le domaine b' et le positionnement différent des résidus pointés précédemment comme importants pour la liaison au substrat. En se basant sur les résultats fournis par Wang *et al.*, nous avons décidé de choisir le modèle 3D de la forme oxydée pour étudier les interactions potentielles entre des polluants et le PDI-b'. En effet, dans sa forme réduite, ce dernier ne semble pas accessible.

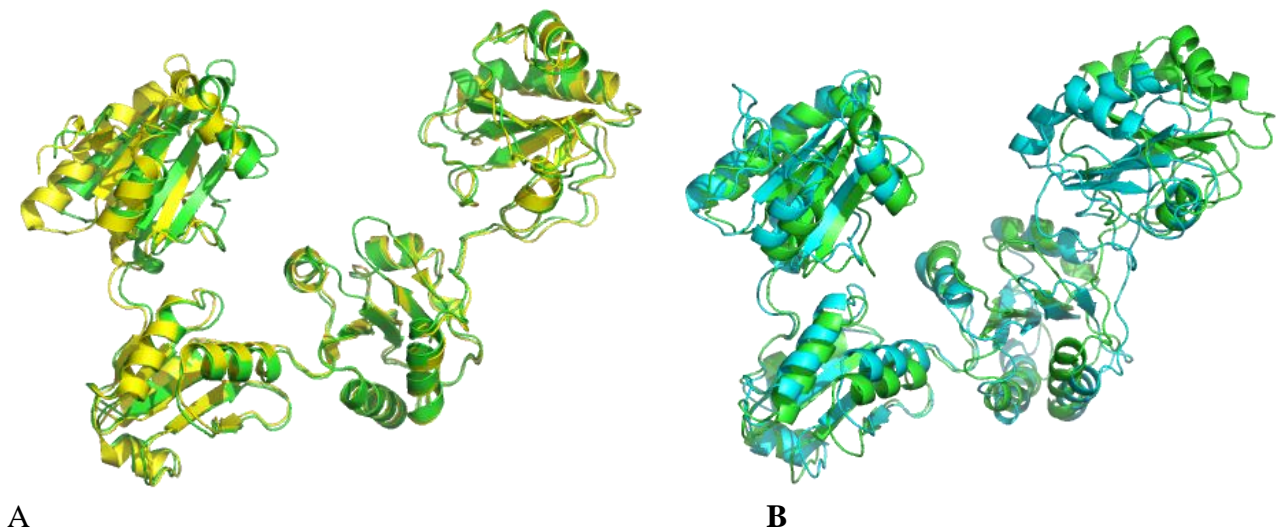


Figure 19: A) Superposition du modèle PDI de *Danio rerio* (PDID) (vert), avec la structure cristallographique de la PDI humaine (4EL1) (jaune). B) Superposition du modèle de la PDI oxydée (vert) et réduite (cyan) de *Danio rerio*.

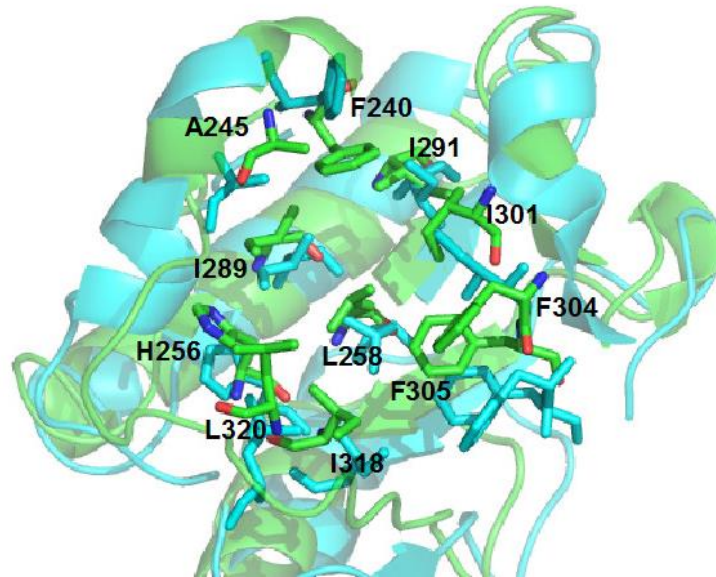


Figure 19: C) Représentation du domaine b' du modèle de la PDI oxydée (vert) et réduite (cyan) de *Danio rerio*

Afin de valider notre modèle, nous avons représenté les résidus sous la forme d'un diagramme de Ramachandran (Fig.20). 96.5 % des acides aminés se trouvent dans une conformation favorable (99.6 % étant dans une position permise).

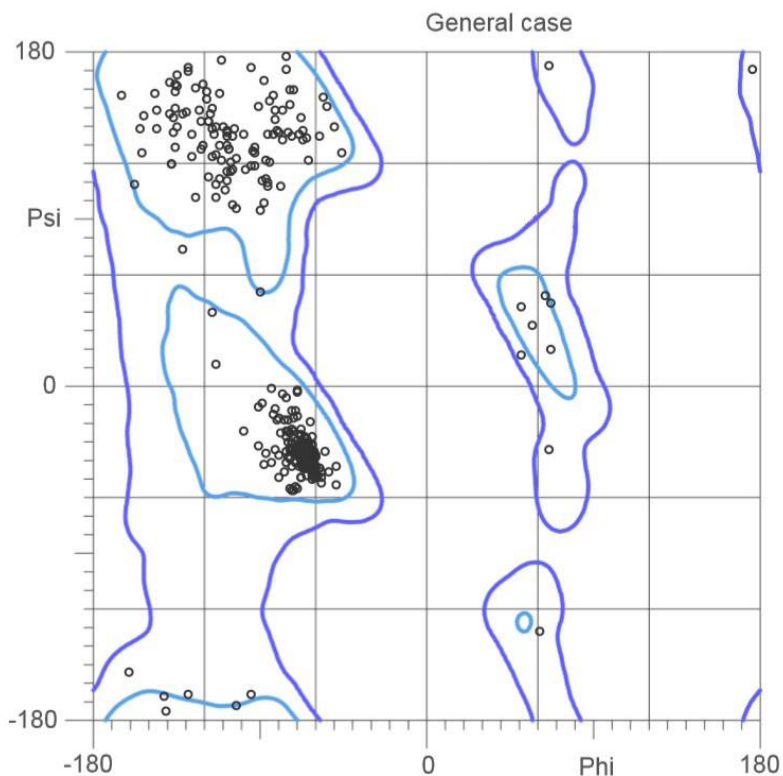


Figure 20: Diagramme de Ramachandran pour le modèle de la PDI de *Danio rerio* oxydée

4.1.4 Amarrage moléculaire ou docking

Le modèle obtenu nous a permis de simuler le mode de liaison de différents polluants dans le PDI-b' de *Danio rerio*. Afin de valider notre méthode de *docking*, nous avons d'abord évalué l'interaction entre le (4S,5S)-1,2-dithiane-4,5-diol (DD) et le domaine b' de la structure de la PDI humaine qui a été cocrystallisée (code PDB: 3UEM) avec ce composé. La figure 21 montre que la position et l'orientation du DD obtenues par *docking* (rose) est très proche de celle de la structure obtenue expérimentalement (cyan). DD se trouve bien localisée dans la poche hydrophobe du domaine b'.

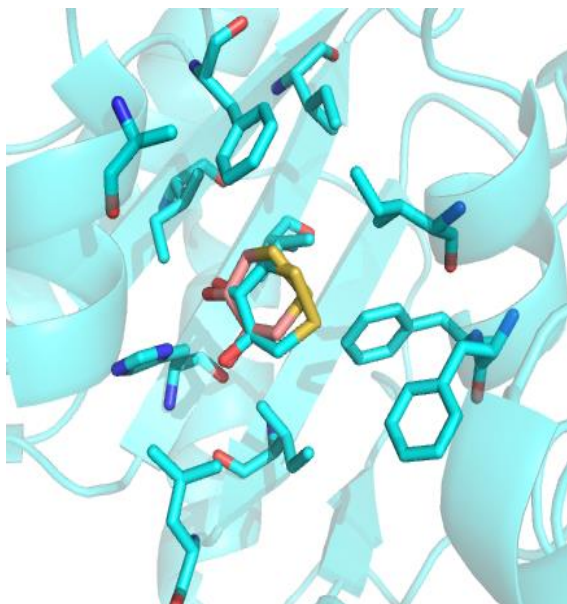


Figure 21: Superposition de la position du (4S,5S)-1,2-dithiane-4,5-diol interagissant avec la PDI-b' humaine, issue soit de l'expérience de cocrystallisation (cyan), soit de l'étude de docking (rose). Les hydrogènes ne sont pas représentés pour plus de clarté.

Après cette étape de vérification, nous avons simulé l'interaction de différents polluants identifiés expérimentalement comme inhibiteurs de la PDI, à savoir le BPA, le 4'-OH-PentaCB, le 4'-OH-BDE17 et l'oestradiol (E_2), avec le PDI-b' de *Danio rerio* (Fig.21). De manière générale, ces quatre molécules se positionnent au même endroit dans le domaine b'. Le mode de liaison du BPA et de l' E_2 (Fig.22A) est en accord avec les études de mutagenèse dirigée identifiant H256 comme étant un résidu déterminant pour l'interaction avec la PDI. En effet, on observe un pont hydrogène entre cet acide aminé et un des hydroxyles en position para du phényle des deux composés. La distance entre l'oxygène et l'azote est d'environ 2.2 Å. Les 4'-OH-PentaCB et 4'-OH-BDE17 ont la même orientation que l' E_2 et positionnent leurs deux groupements phényles de la même manière (Fig.22B). Ces derniers forment vraisemblablement des interactions pi-pi avec les résidus phénylalanine environnants (F240, F304, F305 ; numérotation de la PDI humaine). Par ailleurs, on observe à nouveau un pont hydrogène entre le 4'-OH et l'H256. Ces premiers résultats montrent que ces quatre molécules auraient le même mode de liaison au sein de la PDI.

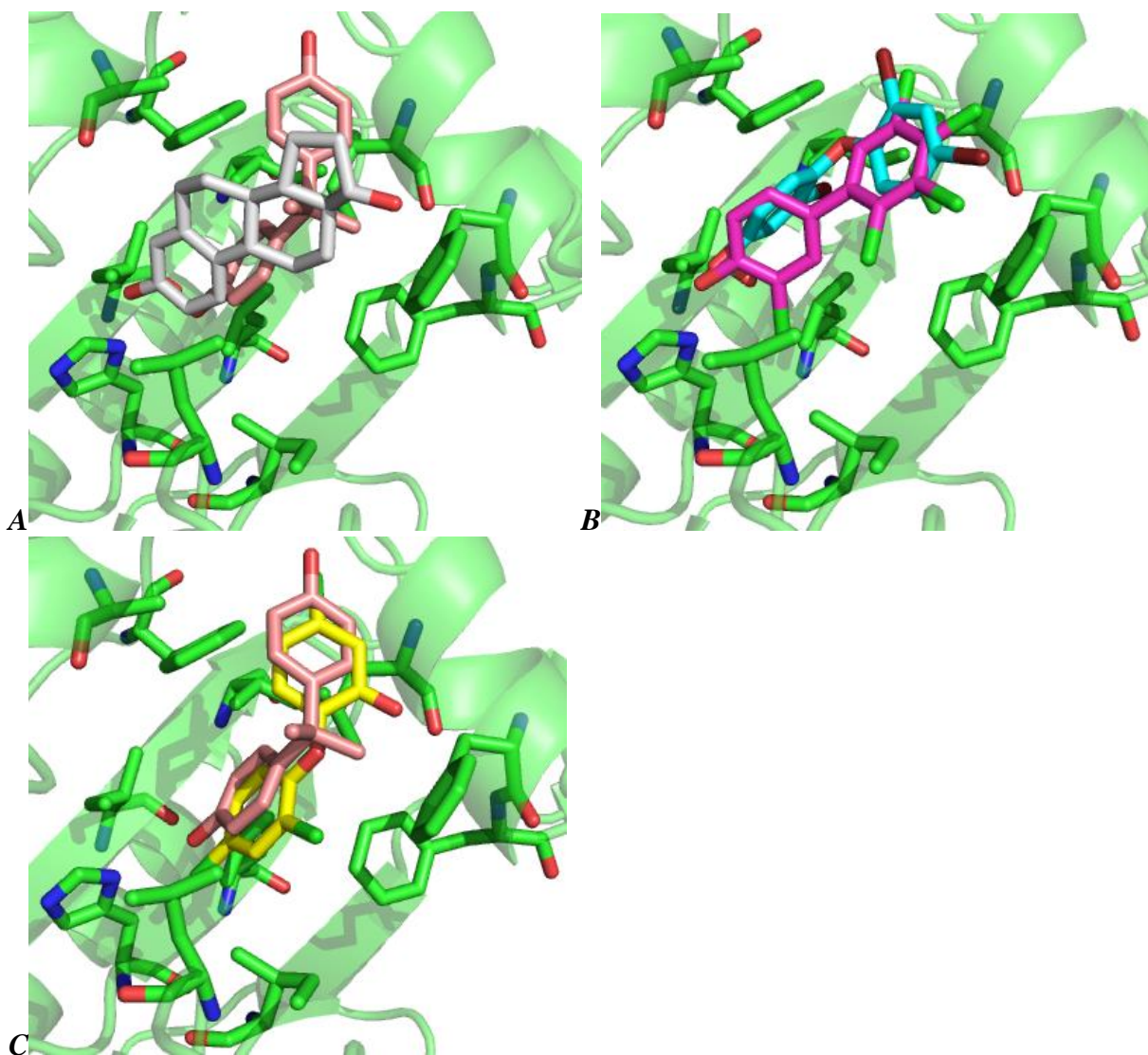


Figure 22: A: Superposition des résultats du docking pour la molécule le BPA (rose) et l'oestradiol (E_2) (blanc). B: Superposition des résultats de docking pour la molécule de 4'-OH-PentaCB (magenta) et de le 4'-OH-BDE17 (cyan). C: Superposition des résultats de docking pour la molécule de TCS (jaune) et de BPA (rose).

Après avoir caractérisé l'orientation de molécules démontrées expérimentalement comme interagissant avec la poche de PDI-b', nous avons simulé l'interaction du TCS dont le mode d'action est encore inconnu. Le TCS semble se positionner de la même manière que le BPA (Fig.22C) avec leurs deux groupes phényles dans la même orientation. La différence se trouve au niveau de la position de leurs substituants. L'hydroxyle du BPA interagissant avec le H256 est remplacé par un chlore dans le TCS alors que l'hydroxyle du TCS forme une interaction OH- π avec le F304, non présente dans le BPA. Ces résultats préliminaires suggèrent que le TCS serait un inhibiteur de la PDI et interagirait comme le BPA avec le PDI-b'.

4.2 Partie épigénétique

Afin de déterminer si le TCS a un impact sur le niveau de méthylation de cytosines situées dans le gène de la PDI de *Danio rerio* (ID: 406673), une analyse de la séquence nucléotidique a été

réalisée à l'aide du programme PyroMark Q24 Advanced. Il s'agit d'un gène d'une taille de 24896 pb constitué de 10 introns et 11 exons où on a pu identifier 3 îlots CpG d'une taille d'environ 200 pb. Pour ces 3 régions, se trouvant respectivement dans l'intron 2 (2 régions) et l'intron 4 (1 région), des amorces ont été construites afin de réaliser un pyroséquençage. La figure 23 représente le résultat de l'amplification de ces 3 régions (1, 2, 3) par PCR pour différentes conditions (A=rouge/vert, B=bleu et C=noir: voir 3.2.4). Comme on peut le voir, l'amplification de ces régions par PCR semble être un succès pour 2 des couples d'amorces, les amplicons devant correspondre à 2 îlots CpG identifiés dans l'intron 2 du gène de la PDI.

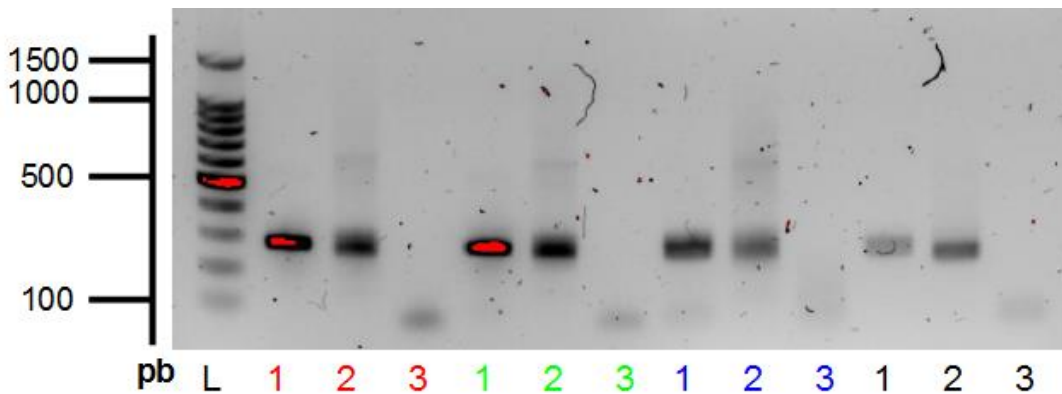
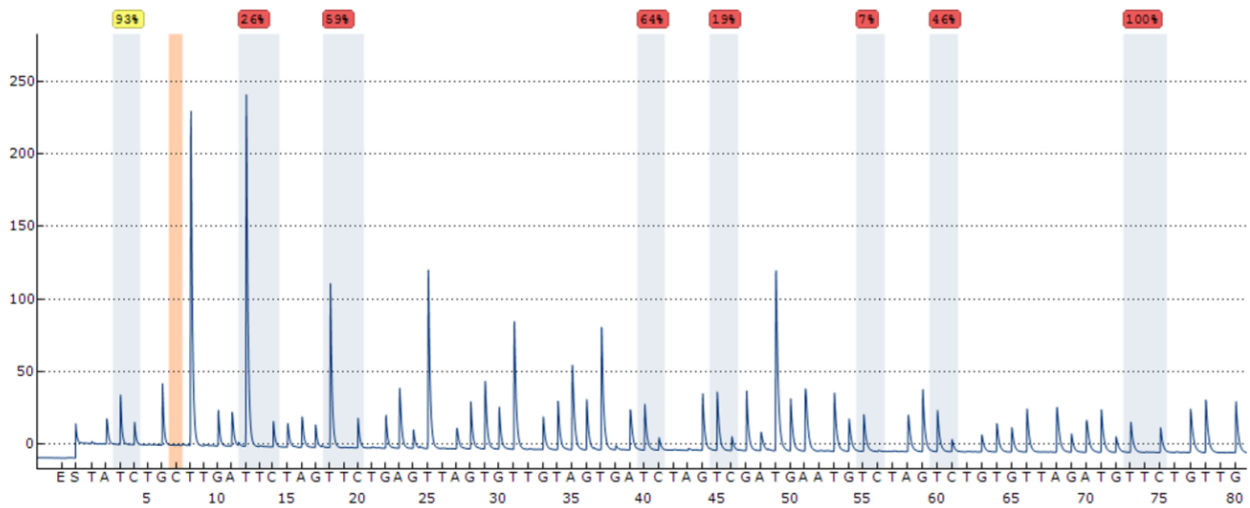


Figure 23: Gel d'électrophorèse réalisé à partir des produits PCR issus d'ADN traité au bisulphite: L=Ladder (étalon, indicateur pour la taille des fragments), 1 (amorces F12-R12), 2 (amorces F22-R22) et 3 (amorces F34-R34). 3 conditions (A, B et C, voir 3.2.3) différentes ont été testés (A=rouge et vert, B=bleu et C=noir).

Ces amplicons ont ensuite été séquencés à l'aide du PyroMark Q24 Advanced (Figure 24). On peut noter que le produit PCR obtenu pour le couple d'amorces F34-R34 n'a pas été séquencé. En effet, la taille du fragment observé sur le gel d'agarose est inférieure à 100 pb, indiquant que rien n'a été amplifié et plus précisément, que le signal doit correspondre aux amorces présentes dans la solution PCR initiale (Fig.23). Le séquençage par le PyroMark Q24 Advanced n'a été réalisé que pour les amplicons des régions 1 et 2. Cependant, il faut noter que la run dans le cas de la région 1 n'a pas fonctionné.

Si l'on regarde le pyrogramme pour la seconde région (2=PDI4hb_22, condition C, en noir sur la figure 23), on peut constater qu'il n'y a pas eu de détection d'un signal lumineux pour la première cystéine (bande orange). Cette cystéine correspond au contrôle négatif du traitement bisulphite et indique donc que cette étape a bien fonctionné. Cependant, bien que le séquençage ait fonctionné, l'intensité du signal lumineux ne semble pas encore être suffisante pour permettre l'identification du nombre exact de nucléotides incorporés. On constate que pour certains CpG, comme celui se trouvant en position 5 sur la séquence, la hauteur des pics n'est tout simplement pas suffisante pour permettre de déterminer le nombre de nucléotides incorporés. Ainsi pour la plupart des CpG, il n'est pas possible de déterminer précisément le pourcentage de méthylation parmi les amplicons présents en solution. Cependant, le séquençage a été réalisé pour quatre conditions différentes dans le cas de la région 2. Le tableau 11 reprend la moyenne des pourcentages de méthylation pour 7 de ces positions ainsi que l'écart-type. En effet dans le cas du CpG présent en position 1, seulement un pourcentage a pu être mesuré.

Pour conclure, il est donc nécessaire de réaliser à nouveau l'amplification des régions afin d'optimiser la qualité mais surtout quantité de produits PCR obtenus. Pour finir, on peut tout de même observer que le résultat du pyroséquençage donne des valeurs de méthylations pour les différentes positions de CpG allant de 7 à 100 % de méthylation.



| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|----|----|----|----|----|---|----|-----|
| Méth. (%) | 93 | 26 | 59 | 64 | 19 | 7 | 46 | 100 |

Figure 24: Pyrogramme obtenu pour les amplicons de la première et seconde région. 13 CpG sont présent sur la séquence pour la région PDI4hb_12 et 8 pour la région PDI4hb_22 qui sont présent avec leur pourcentage respectif de méthylation.

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------------------|---|------|------|------|------|-------|-------|-----|
| Moyenne de Méth. (%) | - | 12 | 43.5 | 37 | 8.25 | 22.25 | 28.5 | 100 |
| Écart-type | - | 8.25 | 9.01 | 15.6 | 6.23 | 8.87 | 10.33 | 0 |

Tableau 11: Moyenne des pourcentages de méthylation des 4 amplicons obtenus par PCR, à chaque position CpG ainsi que l'écart-type.

5. Discussion et perspectives

L'un des objectifs de ce mémoire était de caractériser la PDI de *Danio rerio* expérimentalement en la surexprimant en bactérie. Pour ce faire, il a été nécessaire d'isoler la séquence codante (1530 pb). Concernant une protéine de séquence connue, la stratégie standard fait intervenir une PCR classique, avec des amorces spécifiques permettant d'amplifier l'entièreté de la séquence d'intérêt. Dans notre cas, cette étape a été réalisée au tout début des expérimentations avec différents couples d'amorces mais n'a pas donné de résultats probants.

Devant cette problématique, une stratégie multi-PCR a donc été mise en place, celle-ci étant habituellement utilisée lorsque la séquence d'intérêt est absente des bases de données. Trois PCR ont été réalisées: une amplification initiale afin d'isoler la partie centrale de la séquence d'intérêt, une RACE-PCR pour obtenir les régions non traduites (5'-UTR et 3'-UTR), et une FL-PCR de manière à amplifier l'entièreté de la séquence. Le séquençage séparé des 3 régions (centrale et extrémités) a été réalisé mais cependant le séquençage global est toujours en cours.

L'un des objectifs de ce mémoire était de déterminer à l'aide de méthodes bio-informatiques

si le TCS pouvait interagir avec la PDI chez l'espèce *Danio rerio*. Nous avons d'abord réalisé un alignement de la séquence peptidique de cette dernière avec celle de deux espèces couramment étudiées : l'homme et le rat. Bien que l'alignement avec d'autres espèces de cyprinidés ait pu être intéressant, notamment pour déterminer la phylogénie du *Danio rerio* pour cette protéine, ces organismes ne présentent pas l'avantage d'avoir leurs séquences peptidiques complètement annotées. De plus, avec les informations expérimentales publiées par Wang *et al.* et Lappi *et al.* utilisant la mutagenèse dirigée, nous avons pu désigner différents résidus potentiellement importants pour l'activité de la PDI de *Danio rerio*.^{26,29} Nous avons observé notamment que ces résidus, formant une poche hydrophobe au sein du domaine b' de la PDI, étaient tous conservés entre les 3 espèces. À partir de ces premiers résultats, nous avons pu identifier un site potentiel d'interactions pour le TCS, permettant une analyse ciblée dans le cadre d'un amarrage moléculaire. Cependant, il est important de noter que cette similitude de séquences entre le *Danio rerio*, l'homme et le rat, ne signifie pas que les autres espèces de cyprinidés ou d'organismes aquatiques possèdent les mêmes résidus conservés. En effet, il est possible que suivant les organismes, certains résidus soient ou non mutés. Cette possibilité pourrait être analysée par un alignement d'un plus grand nombre de séquences peptidiques de la PDI.

Ensuite, nous avons construit un modèle 3D pour pouvoir simuler l'interaction de polluants, et notamment le TCS, avec la protéine. Ce modèle est fiable car il a été réalisé à partir de la structure cristallographique de la PDI humaine dont l'identité de séquence avec celle de *Danio rerio* est supérieure à 70%. Par ailleurs, le diagramme de Ramachandran montre que plus de 96% des résidus ont une position favorable au sein de la protéine. La structure modélisée par homologie peut donc être utilisée pour simuler l'interaction entre la protéine et plusieurs polluants. Notons que nous avons choisi la forme oxydée de l'enzyme car il a été montré la forme active de la PDI serait celle dont le site catalytique est oxydé. Plus précisément, le changement d'orientation des domaines b' et a' entraînerait lors de la réduction de la protéine une obstruction du site de liaison empêchant l'interaction protéine-ligand et ainsi l'activité de la PDI.

Avec les différents résultats générés par l'algorithme de *docking* GOLD, nous avons pu visuellement analyser l'interaction des résidus formant la poche hydrophobe avec 5 molécules (E2, 4'-OH-BDE17, 4'-OH-PentaCB, BPA et TCS). Ces résultats corroborent les analyses expérimentales de différentes recherches.^{4,31,44} On constate notamment que H256 ayant été identifiée par Okada *et al.* pour être déterminante dans l'inhibition de la PDI par le BPA semble également nécessaire pour stabiliser l'interaction avec le 4'-OH-BDE17, E2 et 4'-OH-PentaCB. Le TCS se positionnant de la même manière que le BPA, nous pouvons émettre l'hypothèse que ce dernier interagit avec PDI-b'. Les résultats de recherches réalisés sur le BPA pourraient donc servir de base dans le cadre d'une étude de l'effet toxicologique du TCS sur les organismes. On peut d'ailleurs citer l'étude précédente de Clayton *et al.* qui indiquait un lien entre l'exposition au BPA et TCS avec une immunodépression. Nos résultats de *docking* montrent également que les interactions du TCS ne sont pas exactement les mêmes (OH-pi avec F304) que pour les autres composés (OH avec H256), indiquant que H256 aurait peut-être un impact moindre dans sa liaison au domaine b'. Il serait intéressant d'étudier l'impact du TCS sur l'activité de la PDI par rapport au BPA dans une étude de mutagenèse dirigée afin de confirmer cette hypothèse.

On constate que les différentes méthodes de bio-informatique employés dans ce travail nous ont permis d'émettre des hypothèses quant aux interactions protéine-polluants. De plus, les résultats obtenus par ces méthodes permettent d'ouvrir de nouvelles pistes de réflexion au niveau de l'étude fonctionnelle de la PDI.

En effet, dans les recherches d'Okada *et al.* ainsi que dans celle de Wang *et al.*, il a été nécessaire d'établir des protocoles afin de surproduire la protéine, de la purifier et d'en caractériser l'activité de la PDI *in vitro*.^{31,45} Cette approche réalisée pour la structure de la protéine pourrait être

également employée afin de déterminer des protocoles pour étudier l'activité de la PDI. Par exemple, dans une recherche de Fu *et al.* réalisée en 2011, l'activité de la PDI humaine et de la PDIp (homologue pancréatique) ont été caractérisées à l'aide d'une RNase. Plus précisément, la PDI en présence de RNase réduite, catalyse la formation de ponts disulfures, activant la protéine.⁴⁶ cette activation se mesure par l'hydrolyse de cCMP (cyclic cytidine monophosphate) qui est dosée par mesure de l'absorbance à 296 nm. Il ne s'agit pas de la seule recherche ayant employé cette méthode, celle-ci étant couramment utilisée pour caractériser l'activité oxydo-reductase de la PDI.⁴⁷

L'un des autres objectifs de ce mémoire était de caractériser à l'aide du pyroséquençage, les pourcentages de méthylation pour des CpGs présents dans le gène de la PDI de *Danio rerio*. On constate que les pourcentages de méthylation obtenus pour les différentes positions de CpG dans la région 2 sont très variables d'une position à l'autre. De plus, cette variabilité est également importante entre les résultats correspondant à des conditions différentes. Si l'on regarde des résultats obtenus par Strömquist *et al.* (2012), on constate que dans le cas d'une mesure du pourcentage de méthylation entre 3 CpG d'une même région, à savoir la région 5' du gène de la vitellogénine I, il n'obtient pas de variabilité des pourcentages, ceux-ci étant presque identique dans la condition contrôle. Bien que les conditions n'aient pas été optimales, nos résultats nous apportent néanmoins une indication sur la variabilité de la méthylation des CpGs au sein de cette région. Il serait donc intéressant de voir si l'on observerait encore ces résultats après optimisation de la méthode de pyroséquençage.

Comme on le constate dans la figure 23, l'amplification en utilisant la condition A, à savoir un mélange composé de MgCl₂ et d'un marqueur de l'ADN (Coral Loading Dye), est celle ayant le mieux fonctionné pour les 2 premières régions. On constate cependant que la quantité d'amplicons obtenue n'est pas suffisante pour permettre un séquençage de qualité. Une hypothèse qui pourrait être envisagée consisterait à ajouter des quantités plus élevées en ADN traité au bisulfite. Une légère modification de ces concentrations de base permettrait la production d'un plus grand nombre d'amplicons. On peut également noter que l'augmentation de la concentration en MgCl₂ n'est pas envisageable pour optimiser le mix PCR. En effet bien que ce composé soit un cofacteur pour la *Taq* polymérase, il stabilise également l'ADN double brin. Une trop grande concentration en MgCl₂ pourrait donc avoir un effet inverse, diminuant l'efficacité de la PCR. Au niveau des étapes précédant l'amplification, celle-ci semblent avoir fonctionné parfaitement, comme peut notamment le montrer avec le contrôle négatif du traitement bisulfite. Pour le résultat correspondant à la région 3, il n'y a pas eu de détection d'amplicons marqués sur gel d'agarose. Comme les 2 autres régions, cette absence d'amplification pour le couple d'amorces F34-R34 pourrait être dû à une trop faible concentration initiale en ADN traité. Cependant, il faut noter que même si le score donné par le programme PyroMark Q24 Advanced est de 61, cette valeur étant relativement élevée (les amorces construites pour la région deux avait un score de 58), il est fort possible que ce problème d'amplification soit dû à un mauvais appariement des amorces à la séquence d'ADN. Dans le cas où l'optimisation ne fonctionnerait pas pour la région 3, il faudrait envisager de construire de nouvelles amorces.

On remarque que dans les études réalisées par Okada *et al.* (sur les PCBs ou le PBDEs), plusieurs dérivés d'un composé ont été caractérisés afin de déterminer si des modifications structurelles tels que l'ajout d'un groupement fonctionnel pouvait être associé à un gain de toxicité.^{4,31} Bien que nous ne disposions pas encore de protéine surexprimée, il serait intéressant de réaliser des tests d'activité comparant les effets du TCS avec des produits déjà connus pour inhiber la PDI tels que les PCBs ou le BPA ainsi qu'avec des produits issus de sa dégradation tels que le DCDD.

Cette réflexion pourrait également être appliquée à la mesure de la méthylation. En effet, il a déjà été montré dans une étude de Sales *et al.* (2013), portant sur les neuroblastomes de souris, que

la méthylation globale de l'ADN de ces cellules pouvait être affecté par la présence de perturbateurs endocriniens.⁴⁸ Plus précisément, l'exposition de ces cellules à des composés tels que le BPA, le PCB-153 (2,2',4,4',5,5'-hexachlorobiphenyl), le TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin), le BDE-47 (2,2',4,4'-tetrabrominateddiphenyl ether),... a été associé à une déméthylation globale de l'ADN.⁴⁸ Dans un premier temps, il serait intéressant de réaliser un *docking* de ces composés afin de déterminer si d'autres formes de PCBs et PBDEs sont potentiellement capables de venir interagir avec le domaine b' de la PDI et d'en inhiber l'activité. Outre caractériser les différentes formes de PCBs et PBDEs pouvant peut-être inhiber la PDI, on pourrait également s'attarder sur le TCDD. Ce composé est très proche du DCDD, produit formé par dégradation du TCS à la lumière, et pourrait être au même titre que celui-ci un inhibiteur potentiel de la PDI. À l'aide de la bio-informatique, nous pourrions ainsi prévoir différents composés intéressants à tester afin d'identifier et caractériser de nombreux inhibiteurs. Une autre hypothèse, qui pourrait être vérifiée par l'analyse de la méthylation, serait de déterminer si ces composés structurellement proches modulent similairement la méthylation des CpGs du gène de la PDI et peut être plus généralement la méthylation globale de l'ADN. Cette analyse multidisciplinaire permet d'établir des liens avec différents domaines de la biologie, apportant une vision globale de la problématique et débouchant sur de nouvelles perspectives de recherche.

6. Conclusions

Après un peu moins d'un an de travail, il est maintenant temps de conclure ce mémoire. Plusieurs objectifs avaient été posés en début de mémoire afin de déterminer si le TCS avait un impact potentiel sur une protéine disulfure isomérase de *Danio rerio*. Un début de réponse a été apporté à cette question.

En effet, grâce à la méthode de modélisation par homologie, un modèle tridimensionnel de la structure de la PDI de *Danio rerio* a pu être réalisé. Ce modèle a permis de caractériser un site d'interaction présent sur le domaine b' de la protéine qui avait été identifié notamment dans une étude de mutagenèse dirigée.²⁸ À l'aide de la méthode d'amarrage moléculaire, plusieurs simulations d'interactions protéine-ligands ont été réalisées, permettant d'une part, de confirmer des résultats d'interaction avec de petites molécules mis en évidence dans plusieurs recherches et d'autre part, de simuler l'interaction de la PDI avec le TCS. Cette dernière étape a permis d'identifier 2 résidus importants, H256 et F304, nécessaire pour l'interaction de la PDI humaine, ainsi que celle de *Danio rerio*, avec différents ligands (E2, 4'-OH-PentaCB, 4'-OHBDE17, BPA, TCS). Ces résultats ont permis également de donner plusieurs pistes pour la réalisation expérimentale de tests d'activité afin de caractériser le fonctionnement de la PDI de *Danio rerio*.

Un autre de nos objectifs consistait à isoler la séquence codante de la PDI du poisson zèbre et à l'insérer dans un plasmide de surexpression afin de produire et purifier la protéine en bactérie. Cette étape a été partiellement atteinte, la séquence semblant avoir bien été amplifiée par PCR, mais la vérification par séquençage étant toujours en cours, il est pour le moment impossible de le confirmer.

Le troisième objectif de ce mémoire était de caractériser l'impact du TCS sur la méthylation d'îlots CpG présents dans le gène de la PDI. Comme pour l'objectif précédent, ce point est partiellement atteint. En effet, bien que les résultats ayant été générés par pyroséquençage ne puissent pas être interprétés, ces premiers résultats bruts ont permis de déterminer des pistes d'optimisation afin d'obtenir des résultats interprétables.

Bien que les objectifs expérimentaux soient partiellement atteints, les résultats théoriques de la modélisation apportent de nouvelles pistes de réflexion sur l'analyse de l'effet du TCS sur la PDI de *Danio rerio* et la méthylation de son gène, qui pourrait être comparé avec d'autres types de

polluants tels que le BPA. De plus on constate que les polluants structurellement proches comme le BPA et TCS, ont un impact relativement identique sur des organismes tels que le rat et l'homme. Plus généralement, l'impact de ces polluants pourrait être comparé sur d'autres organismes aquatiques. Remarquons que ces pistes de réflexion sont surtout apportées par l'analyse des résultats générés dans plusieurs domaines différents de la biologie, ce qui n'aurait peut-être pas pu être mis en évidence sans l'aspect multidisciplinaire de cette recherche.

Remerciements

Tout d'abord, je tiens à remercier sincèrement mon promoteur Frédéric Silvestre et ma copromotrice Catherine Michaux, pour l'aide apporté tout au long de la durée de mon mémoire. Leurs conseils et leurs critiques m'ont été d'une très grande utilité afin d'évoluer au cours de l'année passée en leur compagnie.

Je souhaite également remercier mes deux encadrants, Emmanuel Tinti et Elodie Falisse pour leurs conseils, le temps qu'ils m'ont accordé ainsi que pour leurs amabilités.

Je remercie également mes promoteurs ainsi que mes encadrants pour leurs corrections apportés dans la réalisation de ce travail.

Pour terminer, je tiens à remercier les membres du département d'URBE ainsi que les membres de l'UCPTS et plus particulièrement Guillaume Roussel et le professeur Eric Perpete pour vos nombreux conseils et votre sympathie.

Merci encore à toutes et tous.

7. Bibliographie

1. Kallenborn, R. Persistent organic pollutants (POPs) as environmental risk factors in remote high-altitude ecosystems. *Ecotoxicol. Environ. Saf.* **63**, 100–107 (2006).
2. Jones, K. C. & de Voogt, P. Persistent organic pollutants (POPs): state of the science. *Environ. Pollut.* **100**, 209–21 (1999).
3. El-Shahawi, M. S., Hamza, a, Bashammakh, a S. & Al-Saggaf, W. T. An overview on the accumulation, distribution, transformations, toxicity and analytical methods for the monitoring of persistent organic pollutants. *Talanta* **80**, 1587–97 (2010).
4. Okada, K., Hashimoto, S., Funae, Y. & Imaoka, S. Hydroxylated polychlorinated biphenyls (PCBs) interact with protein disulfide isomerase and inhibit its activity. *Chem. Res. Toxicol.* **22**, 899–904 (2009).
5. Gomes, H. I., Dias-Ferreira, C. & Ribeiro, A. B. Overview of in situ and ex situ remediation technologies for PCB-contaminated soils and sediments and obstacles for full-scale application. *Sci. Total Environ.* **445-446**, 237–60 (2013).

6. Gladen, B. C. *et al.* Development after exposure to polychlorinated biphenyls and dichlorodiphenyl dichloroethene transplacentally and through human milk. *J. Pediatr.* **113**, 991–995 (1988).
7. Wang, L. *et al.* PCB-77 disturbs iron homeostasis through regulating hepcidin gene expression. *Gene* **532**, 146–51 (2013).
8. Garner, L. V. T., Brown, D. R. & Di Giulio, R. T. Knockdown of AHR1A but not AHR1B exacerbates PAH and PCB-126 toxicity in zebrafish (*Danio rerio*) embryos. *Aquat. Toxicol.* **142-143C**, 336–346 (2013).
9. Hiroi, T., Okada, K., Imaoka, S., Osada, M. & Funae, Y. Bisphenol A binds to protein disulfide isomerase and inhibits its enzymatic and hormone-binding activities. *Endocrinology* **147**, 2773–2780 (2006).
10. Bhargava, H. N. & Leonard, P. A. Triclosan : Applications and safety. *Am. J. Infect. Control* **24**, 209–218 (1996).
11. Dann, A. B. & Hontela, A. Triclosan: environmental exposure, toxicity and mechanisms of action. *J. Appl. Toxicol.* **31**, 285–311 (2011).
12. Adolfsson-Erici, M., Pettersson, M., Parkkonen, J. & Sturve, J. Triclosan, a commonly used bactericide found in human milk and in the aquatic environment in Sweden. *Chemosphere* **46**, 1485–9 (2002).
13. Oliveira, R., Domingues, I., Koppe Grisolia, C. & Soares, A. M. V. M. Effects of triclosan on zebrafish early-life stages and adults. *Environ. Sci. Pollut. Res. Int.* **16**, 679–688 (2009).
14. Binelli, a, Cogni, D., Parolini, M., Riva, C. & Provini, a. Cytotoxic and genotoxic effects of in vitro exposure to triclosan and trimethoprim on zebra mussel (*Dreissena polymorpha*) hemocytes. *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.* **150**, 50–6 (2009).
15. Kawanai, T. Triclosan, an environmental pollutant from health care products, evokes charybdotoxin-sensitive hyperpolarization in rat thymocytes. *Environ. Toxicol. Pharmacol.* **32**, 417–22 (2011).
16. Rees Clayton, E. M., Todd, M., Dowd, J. B. & Aiello, A. E. The Impact of Bisphenol A and Triclosan on Immune Parameters in the U.S. Population, NHANES 2003–2006. *Environ. Health Perspect.* **119**, 390–396 (2011).
17. Aranami, K. & Readman, J. W. Photolytic degradation of triclosan in freshwater and seawater. *Chemosphere* **66**, 1052–6 (2007).
18. Canosa, P. *et al.* Aquatic degradation of triclosan and formation of toxic chlorophenols in presence of low concentrations of free chlorine. *Anal. Bioanal. Chem.* **383**, 1119–26 (2005).
19. Sun, S. *et al.* Inhibition of prolyl 4-hydroxylase, beta polypeptide (P4HB) attenuates temozolomide resistance in malignant glioma via the endoplasmic reticulum stress response (ERSR) pathways. *Neuro. Oncol.* **15**, 562–77 (2013).
20. Ren, J. *et al.* Structural Insights into the Redox-Regulated Dynamic Conformations of Human Protein Disulfide Isomerase. *Antioxid. Redox Signal.* 120709084834006 (2012). doi:10.1089/ars.2012.4630

21. Wilkinson, B. & Gilbert, H. F. Protein disulfide isomerase. *Biochim. Biophys. Acta* **1699**, 35–44 (2004).
22. Appenzeller-Herzog, C. & Ellgaard, L. The human PDI family: versatility packed into a single fold. *Biochim. Biophys. Acta* **1783**, 535–548 (2008).
23. Tian, G., Xiang, S., Noiva, R., Lennarz, W. J. & Schindelin, H. The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell* **124**, 61–73 (2006).
24. Lucero, H. A. & Kaminer, B. The Role of Calcium on the Activity of ERcalcristorin / Protein-disulfide Isomerase and the Significance of the C-terminal and Its. **274**, 3243–3251 (1999).
25. Lappi, A. K. *et al.* A conserved arginine plays a role in the catalytic cycle of the protein disulphide isomerases. *J. Mol. Biol.* **335**, 283–295 (2004).
26. Karala, A.-R., Lappi, A.-K. & Ruddock, L. W. Modulation of an active-site cysteine pKa allows PDI to act as a catalyst of both disulfide bond formation and isomerization. *J. Mol. Biol.* **396**, 883–892 (2010).
27. Laurindo, F. R. M., Pescatore, L. a & Fernandes, D. D. C. Protein disulfide isomerase in redox cell signaling and homeostasis. *Free Radic. Biol. Med.* **52**, 1954–69 (2012).
28. Nguyen, V. D. *et al.* Alternative Conformations of the x Region of Human Protein Disulphide-Isomerase Modulate Exposure of the Substrate Binding b' Domain. *J. Mol. Biol.* **383**, 1144–1155 (2008).
29. Wang, C. *et al.* Human protein-disulfide isomerase is a redox-regulated chaperone activated by oxidation of domain a'. *J. Biol. Chem.* **287**, 1139–49 (2012).
30. Hashimoto, S., Shiimoto, K., Okada, K. & Imaoka, S. The binding site of bisphenol A to protein disulphide isomerase. *J. Biochem.* **151**, 35–45 (2012).
31. Hashimoto, S. *et al.* Effects of polybrominated diphenyl ethers (PBDEs) and their derivatives on protein disulfide isomerase activity and growth hormone release of GH3 cells. *Chem. Res. Toxicol.* **25**, 656–63 (2012).
32. McGrath, P. & Li, C.-Q. Zebrafish: a predictive model for assessing drug-induced toxicity. *Drug Discov. Today* **13**, 394–401 (2008).
33. Huang, Q.-Y., Huang, L. & Huang, H.-Q. Proteomic analysis of methyl parathion-responsive proteins in zebrafish (*Danio rerio*) brain. *Comp. Biochem. Physiol. Toxicol. Pharmacol. CBP* **153**, 67–74 (2011).
34. Ling, X., Lu, Y. & Huang, H. Differential protein profile in zebrafish (*Danio rerio*) brain under the joint exposure of methyl parathion and cadmium. (2012). doi:10.1007/s11356-012-1037-3
35. Lindholst, C., Wynne, P. M., Marriott, P., Pedersen, S. N. & Bjerregaard, P. Metabolism of bisphenol A in zebrafish (*Danio rerio*) and rainbow trout (*Oncorhynchus mykiss*) in relation to estrogenic response. *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.* **135**, 169–177 (2003).

36. Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226–232 (1975).
37. Strömqvist, M., Tooke, N. & Brunström, B. DNA methylation levels in the 5' flanking region of the vitellogenin I gene in liver and brain of adult zebrafish (*Danio rerio*)--sex and tissue differences and effects of 17alpha-ethinylestradiol exposure. *Aquat. Toxicol.* **98**, 275–81 (2010).
38. Johnson, M. *et al.* NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).
39. Here, E. T., Us, A., Feedback, H., Applications, R. & Sequence, M. ClustalW2 Results. *Mol. Biol.* **189**, 2012 (2011).
40. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
41. Lambert, C., Léonard, N., De Bolle, X. & Depiereux, E. ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**, 1250–6 (2002).
42. Rother, K. Introduction to PyMOL. *Methods Mol. Biol. Clift. Nj* **635**, 0–32 (2005).
43. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **52**, 609–623 (2003).
44. Hashimoto, S., Okada, K. & Imaoka, S. Interaction between bisphenol derivatives and protein disulphide isomerase (PDI) and inhibition of PDI functions: requirement of chemical structure for binding to PDI. *J. Biochem.* **144**, 335–342 (2008).
45. Gao, M. *et al.* Synthesis of radiolabeled protein disulfide isomerase (PDI) inhibitors as new potential PET agents for imaging of the enzyme PDI in neurological disorders and cancer. *Appl. Radiat. Isot.* **74**, 61–9 (2013).
46. Fu, X.-M. & Zhu, B. T. Human pancreas-specific protein disulfide isomerase homolog (PDIp) is an intracellular estrogen-binding protein that modulates estrogen levels and actions in target cells. *J. Steroid Biochem. Mol. Biol.* **115**, 20–29 (2009).
47. Xiao, R., Lundström-Ljung, J., Holmgren, A. & Gilbert, H. F. Catalysis of thiol/disulfide exchange. Glutaredoxin 1 and protein-disulfide isomerase use different mechanisms to enhance oxidase and reductase activities. *J. Biol. Chem.* **280**, 21099–21106 (2005).
48. Bastos Sales, L. *et al.* Effects of endocrine disrupting chemicals on in vitro global DNA methylation and adipocyte differentiation. *Toxicol. In Vitro* **27**, 1634–43 (2013).