

2025-2026

Thèse

pour le

Diplôme d'État de Docteur en Pharmacie

Développement de modèles pharmacophoriques et criblage virtuel de bibliothèques chimiques pour l'identification d'inhibiteurs de la prostaglandine E2 synthase 1 microsomale (mPGES-1)

--

Development of pharmacophore models and virtual screening of chemical libraries for identification of microsomal prostaglandin E synthase 1 (mPGES-1) inhibitors

Gossye Thomas |

Né le 05 avril 1989 à Séclin (59)

Sous la direction de M. Helesbeux Jean-Jacques |

Membres du jury

Denis Séraphin | Président

Jean Jacques Helesbeux | Directeur

Chau Phi Dinh | Membre

Soutenue publiquement le :
17 décembre 2025



**FACULTÉ
DE SANTÉ**

UNIVERSITÉ D'ANGERS

LISTE DES ENSEIGNANTS DE LA FACULTÉ DE SANTÉ D'ANGERS

(mise à jour 06/10/2025)

Doyen de la Faculté : Pr Cédric ANNWEILER

Vice-Doyen de la Faculté et directeur du département de pharmacie : Pr Sébastien FAURE

Directeur du département de médecine : Pr Vincent DUBEE

PROFESSEURS DES UNIVERSITÉS

ABRAHAM Pierre	PHYSIOLOGIE	Médecine
ANGOULVANT Cécile	MEDECINE GENERALE	Médecine
ANNWEILER Cédric	GERIATRIE ET BIOLOGIE DU VIEILLISSEMENT	Médecine
ASFAR Pierre	REANIMATION	Médecine
AUBE Christophe	RADIOLOGIE ET IMAGERIE MEDICALE	Médecine
AUGUSTO Jean-François	NEPHROLOGIE	Médecine
BAUFRETON Christophe	CHIRURGIE THORACIQUE ET CARDIOVASCULAIRE	Médecine
BELLANGER William	MEDECINE GENERALE	Médecine
BELONCLE François	REANIMATION	Médecine
BIERE Loïc	CARDIOLOGIE	Médecine
BIGOT Pierre	UROLOGIE	Médecine
BONNEAU Dominique	GENETIQUE	Médecine
BOUCHARA Jean-Philippe	PARASITOLOGIE ET MYCOLOGIE	Médecine
BOUET Pierre-Emmanuel	GYNECOLOGIE-OBSTETRIQUE	Médecine
BOURSIER Jérôme	GASTROENTEROLOGIE ; HEPATOLOGIE	Médecine
BOUVARD Béatrice	RHUMATOLOGIE	Médecine
BRIET Marie	PHARMACOLOGIE	Médecine
CAMPONE Mario	CANCEROLOGIE ; RADIOTHERAPIE	Médecine
CAROLI-BOSC François-Xavier	GASTROENTEROLOGIE ; HEPATOLOGIE	Médecine
CASSEREAU Julien	NEUROLOGIE	Médecine
CLERE Nicolas	PHARMACOLOGIE / PHYSIOLOGIE	Pharmacie
COLIN Estelle	GENETIQUE	Médecine
CONNAN Laurent	MEDECINE GENERALE	Médecine
COPIN Marie-Christine	ANATOMIE ET CYTOLOGIE PATHOLOGIQUES	Médecine
COUTANT Régis	PEDIATRIE	Médecine
CUSTAUD Marc-Antoine	PHYSIOLOGIE	Médecine
CRAUSTE-MANCIET Sylvie	PHARMACOTECHNIQUE HOSPITALIERE	Pharmacie
DE CASABIANCA Catherine	MEDECINE GENERALE	Médecine
DERBRE Séverine	PHARMACOGNOSIE	Pharmacie
DESCAMPS Philippe	GYNECOLOGIE-OBSTETRIQUE	Médecine
D'ESCATHA Alexis	MEDECINE ET SANTE AU TRAVAIL	Médecine
DINOMAS Mickaël	MEDECINE PHYSIQUE ET DE READAPTATION	Médecine
DUBEE Vincent	MALADIES INFECTIEUSES ET TROPICALES	Médecine
	BACTERIOLOGIE-VIROLOGIE ; HYGIENE	
DUCANCELLE Alexandra	HOSPITALIERE	Médecine
	PEDOPSYCHIATRIE	Médecine
DUVERGER Philippe	BACTERIOLOGIE-VIROLOGIE	Pharmacie
EVEILLARD Matthieu	PHARMACOLOGIE PHYSIOLOGIE	Pharmacie
FAURE Sébastien	ANATOMIE	Médecine
FOURNIER Henri-Dominique	CHIRURGIE THORACIQUE ET CARDIOVASCULAIRE	Médecine
FOUQUET Olivier	CARDIOLOGIE	Médecine
FURBER Alain	PNEUMOLOGIE	Médecine
GAGNADOUX Frédéric	PSYCHIATRIE D'ADULTES	Médecine
GOHIER Bénédicte	HEMATOLOGIE ; TRANSFUSION	Médecine
GUARDIOLA Philippe		



GUILLET David	CHIMIE ANALYTIQUE	Pharmacie
HUNAUT-BERGER Mathilde	HEMATOLOGIE ; TRANSFUSION	Médecine
JEANNIN Pascale	IMMUNOLOGIE	Médecine
KAZOUR François	PSYCHIATRIE	Médecine
KEMPF Marie	BACTERIOLOGIE-VIROLOGIE ; HYGIENE HOSPITALIERE	Médecine
KUN-DARBOIS Daniel	CHIRURGIE MAXILLO-FACIALE ET STOMATOLOGIE	Médecine
LACOEUILLE FRANCK	RADIOPHARMACIE	Pharmacie
LACCOURREYE Laurent	OTO-RHINO-LARYNGOLOGIE	Médecine
LAGARCE Frédéric	BIOPHARMACIE	Pharmacie
LANDREAU Anne	BOTANIQUE/ MYCOLOGIE	Pharmacie
LASOCKI Sigismond	ANESTHESIOLOGIE-REANIMATION	Médecine
LEBDAI Souhil	UROLOGIE	Médecine
LEGENDRE Guillaume	GYNECOLOGIE-OBSTETRIQUE	Médecine
LEGRAND Erick	RHUMATOLOGIE	Médecine
LEMEE Jean-Michel	NEUROCHIRURGIE	Médecine
LERMITE Emilie	CHIRURGIE GENERALE	Médecine
LEROLLE Nicolas	REANIMATION	Médecine
LIBOUBAN Hélène	HISTOLOGIE	Médecine
LUQUE PAZ Damien	HEMATOLOGIE BIOLOGIQUE	Médecine
MARCHAIS Véronique	BACTERIOLOGIE-VIROLOGIE	Pharmacie
MARTIN Ludovic	DERMATO-VENEREOLOGIE	Médecine
MAY-PANLOUP Pascale	BIOLOGIE ET MEDECINE DU DEVELOPPEMENT ET DE LA REPRODUCTION	Médecine
MENEI Philippe	NEUROCHIRURGIE	Médecine
MERCAT Alain	REANIMATION	Médecine
ORVAIN Corentin	HEMATOLOGIE ; TRANSFUSION	Médecine
PAISANT Anita	RADIOLOGIE	Médecine
PAPON Nicolas	PARASITOLOGIE ET MYCOLOGIE MEDICALE	Pharmacie
PASSIRANI Catherine	CHIMIE GENERALE	Pharmacie
PELLIER Isabelle	PEDIATRIE	Médecine
PETIT Audrey	MEDECINE ET SANTE AU TRAVAIL	Médecine
PICQUET Jean	CHIRURGIE VASCULAIRE ; MEDECINE VASCULAIRE	Médecine
PODEVIN Guillaume	CHIRURGIE INFANTILE	Médecine
PROCACCIO Vincent	GENETIQUE	Médecine
PRUNIER Delphine	BIOCHIMIE ET BIOLOGIE MOLECULAIRE	Médecine
PRUNIER Fabrice	CARDIOLOGIE	Médecine
PY Thibaut	MEDECINE GENERALE	Médecine
RAMOND-ROQUIN Aline	MEDECINE GENERALE	Médecine
REYNIER Pascal	BIOCHIMIE ET BIOLOGIE MOLECULAIRE	Médecine
RIOU Jérémie	BIOSTATISTIQUE	Pharmacie
RINEAU Emmanuel	ANESTHESIOLOGIE REANIMATION	Médecine
RIQUIN Elise	PEDOPSYCHIATRIE ; ADDICTOLOGIE	Médecine
RODIEN Patrice	ENDOCRINOLOGIE, DIABETE ET MALADIES METABOLIQUES	Médecine
ROQUELAURE Yves	MEDECINE ET SANTE AU TRAVAIL	Médecine
ROUGE-MAILLART Clotilde	MEDECINE LEGALE ET DROIT DE LA SANTE	Médecine
ROUSSEAU Audrey	ANATOMIE ET CYTOLOGIE PATHOLOGIQUES	Médecine
ROUSSEAU Pascal	CHIRURGIE PLASTIQUE, RECONSTRUCTRICE ET ESTHETIQUE	Médecine
ROUSSELET Marie-Christine	ANATOMIE ET CYTOLOGIE PATHOLOGIQUES	Médecine
ROY Pierre-Marie	MEDECINE D'URGENCE	Médecine
SAULNIER Patrick	BIOPHYSIQUE ET BIOSTATISTIQUES	Pharmacie
SERAPHIN Denis	CHIMIE ORGANIQUE	Pharmacie
SCHMIDT Aline	HEMATOLOGIE ; TRANSFUSION	Médecine
TESSIER-CAZENEUVE Christine	MEDECINE GENERALE	Médecine



TRZEPIZUR Wojciech	PNEUMOLOGIE	Médecine
UGO Valérie	HEMATOLOGIE ; TRANSFUSION	Médecine
URBAN Thierry	PNEUMOLOGIE	Médecine
VAN BOGAERT Patrick	PEDIATRIE	Médecine
VENARA Aurélien	CHIRURGIE VISCERALE ET DIGESTIVE	Médecine
VENIER-JULIENNE Marie-Claire	PHARMACOTECHNIE	Pharmacie
VERNY Christophe	NEUROLOGIE	Médecine
WILLOTEAUX Serge	RADIOLOGIE ET IMAGERIE MEDICALE	Médecine

MAÎTRES DE CONFÉRENCES

AMMI Myriam	CHIRURGIE VASCULAIRE ET THORACIQUE	Médecine
BAGLIN Isabelle	CHIMIE THERAPEUTIQUE	Pharmacie
BASTIAT Guillaume	BIOPHYSIQUE ET BIOSTATISTIQUES	Pharmacie
BEAUVILLAIN Céline	IMMUNOLOGIE	Médecine
BEGUE Cyril	MEDECINE GENERALE	Médecine
BELIZNA Cristina	MEDECINE INTERNE	Médecine
BERNARD Florian	ANATOMIE	Médecine
BESSAGUET Flavien	PHYSIOLOGIE PHARMACOLOGIE	Pharmacie
BLANCHET Odile	HEMATOLOGIE ; TRANSFUSION	Médecine
BOISARD Séverine	CHIMIE ANALYTIQUE	Pharmacie
BOUCHER Sophie	ORL	Médecine
BRIET Claire	ENDOCRINOLOGIE, DIABETE ET MALADIES METABOLIQUES	Médecine
BRILLAND Benoit	NEPHROLOGIE	Médecine
BRIS Céline	BIOCHIMIE ET BIOLOGIE MOLECULAIRE	Pharmacie
BRUGUIERE Antoine	PHARMACOGNOSIE	Pharmacie
CAPITAIN Olivier	CANCEROLOGIE ; RADIOTHERAPIE	Médecine
CHABRUN Floris	BIOCHIMIE ET BIOLOGIE MOLECULAIRE	Pharmacie
CHAO DE LA BARCA Juan-Manuel	BIOCHIMIE ET BIOLOGIE MOLECULAIRE	Médecine
CHOPIN Matthieu	MEDECINE GENERALE	
CODRON Philippe	NEUROLOGIE	Médecine
CORVAISIER Mathieu	PHARMACIE CLINIQUE	Pharmacie
DEMAS Josselin	SCIENCES DE LA READAPTATION	Médecine
DESHAYES Caroline	BACTERIOLOGIE VIROLOGIE	Pharmacie
DOUILLET Delphine	MEDECINE D'URGENCE	Médecine
FERRE Marc	BIOLOGIE MOLECULAIRE	Médecine
FORTRAT Jacques-Olivier	PHYSIOLOGIE	Médecine
GHALI Maria	MEDECINE GENERALE	Médecine
GUELFF Jessica	MEDECINE GENERALE	Médecine
HADJ MAHMOUD Dorra	IMMUNOLOGIE	Pharma
HAMEL Jean-François	BIOSTATISTIQUES, INFORMATIQUE MEDICALE	Médicale
HAMON Cédric	MEDECINE GENERALE	Médecine
HELESBEUX Jean-Jacques	CHIMIE ORGANIQUE	Pharmacie
HERIVAUX Anaïs	BIOTECHNOLOGIE	Pharmacie
HINDRE François	BIOPHYSIQUE	Médecine
JOUSSET-THULLIER Nathalie	MEDECINE LEGALE ET DROIT DE LA SANTE	Médecine
JUDALET-ILLAND Ghislaine	MEDECINE GENERALE	Médecine
KHIATI Salim	BIOCHIMIE ET BIOLOGIE MOLECULAIRE	Médecine
LEFEUVRE Caroline	BACTERIOLOGIE ; VIROLOGIE	Médecine
LEGEAY Samuel	PHARMACOCINETIQUE	Pharmacie
LEPELTIER Elise	CHIMIE GENERALE	Pharmacie
LE ROUX Gaël	TOXICOLOGIE	Pharmacie
LETOURNEL Franck	BIOLOGIE CELLULAIRE	Médecine
MABILLEAU Guillaume	HISTOLOGIE, EMBRYOLOGIE ET CYTOGENETIQUE	Médecine
MAROT Agnès	PARASITOLOGIE ET MYCOLOGIE MEDICALE	Pharmacie



MESLIER Nicole	PHYSIOLOGIE	Médecine
MIOT Charline	IMMUNOLOGIE	Médecine
MOUILLIE Jean-Marc	PHILOSOPHIE	Médecine
NAIL BILLAUD Sandrine	IMMUNOLOGIE	Pharmacie
PAILHORIE Hélène	BACTERIOLOGIE-VIROLOGIE	Médecine
PAPON Xavier	ANATOMIE	Médecine
PASCO-PAPON Anne	RADIOLOGIE ET IMAGERIE MEDICALE	Médecine
PENCHAUD Anne-Laurence	SOCIOLOGIE	Médecine
PIHET Marc	PARASITOLOGIE ET MYCOLOGIE	Médecine
PIRAUX Arthur	OFFICINE	Pharmacie
POIROUX Laurent	SCIENCES INFIRMIERES	Médecine
RONY Louis	CHIRURGIE ORTHOPEDIQUE ET TRAUMATOLOGIQUE	Médecine
ROGER Emilie	PHARMACOTECHNIE	Pharmacie
SAVARY Camille	PHARMACOLOGIE-TOXICOLOGIE	Pharmacie
SCHMITT Françoise	CHIRURGIE INFANTILE	Médecine
SCHINKOWITZ Andréas	PHARMACOGNOSIE	Pharmacie
SPIESSER-ROBELET Laurence	PHARMACIE CLINIQUE ET EDUCATION THERAPEUTIQUE	Pharmacie
TEXIER-LEGENDRE Gaëlle	MEDECINE GENERALE	Médecine
VIAULT Guillaume	CHIMIE ORGANIQUE	Pharmacie

AUTRES ENSEIGNANTS

ATER		
BARAKAT Fatima	CHIMIE ANALYTIQUE	Pharmacie
ATCHADE Constantin	GALENIQUE	Pharmacie
ECER		
HASAN Mahmoud	GALENIQUE	Pharmacie
PRCE		
AUTRET Erwan	ANGLAIS	Santé
BARBEROUSSE Michel	INFORMATIQUE	Santé
COYNE Ashley	ANGLAIS	Santé
O'SULLIVAN Kayleigh	ANGLAIS	Santé
RIVEAU Hélène	ANGLAIS	Santé
PAST-MAST		
AUBRUHET Hélène	PHARMACIE DEUST PREPARATEUR	Pharmacie
BEAUVAIS Vincent	OFFICINE	Pharmacie
BRAUD Cathie	PHARMACIE DEUST PREPARATEUR	Pharmacie
CAVAILLON Pascal	PHARMACIE INDUSTRIELLE	Pharmacie
CHAMPAGNE Romain	MEECINE PHYSIQUE ET READAPTATION	Médecine
DILÉ Nathalie	OFFICINE	Pharmacie
GUITTON Christophe	MEDECINE INTENSIVE-REANIMATION	Médecine
KAASSIS Mehdi	GASTRO-ENTEROLOGIE	Médecine
LAVIGNE Christian	MEDECINE INTERNE	Médecine
MARSAN-POIROUX Sylvie	COMMUNICATION	Pharmacie
MOAL Frédéric	PHARMACIE CLINIQUE	Pharmacie
PICCOLI Giorgia	NEPHROLOGIE	Médecine
POMMIER Pascal	CANCEROLOGIE-RADIOTHERAPIE	Médecine
SAVARY Dominique	MEDECINE D'URGENCE	Médecine
PLP		

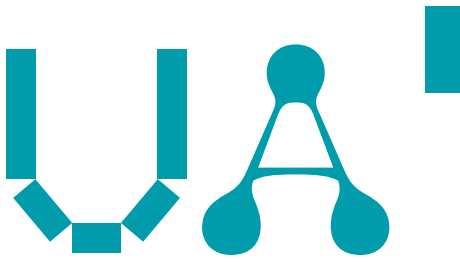
CHIKH Yamina	ECONOMIE-GESTION	Médecine
AHU		
ROBIN Julien	DISPOSITIFS MEDICAUX	Pharmacie

ENGAGEMENT DE NON-PLAGIAT ET DE RESPECT DE LA CHARTE INTELLIGENCE ARTIFICIELLE

Je, soussigné(e) THOMAS GOSSYE.....
déclare être pleinement conscient(e) que le plagiat de documents ou d'une
partie d'un document publiée sur toutes formes de support, y compris l'internet,
constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.
En conséquence, je m'engage à citer toutes les sources que j'ai utilisées
pour écrire ce rapport ou mémoire.

Je déclare également respecter la charte de la Faculté de Santé de l'Université d'Angers sur l'intelligence artificielle
générative dans le cadre de la thèse d'exercice.

signé par l'étudiant(e) le 6 / 12 / 2025



FACULTÉ DE SANTÉ

UNIVERSITÉ D'ANGERS

"La Faculté de Santé déclare que les opinions émises dans les thèses qui lui sont présentées doivent être considérées comme propres à leurs auteurs, et qu'elle entend ne leur donner ni approbation, ni improbation."

REMERCIEMENTS

Aux membres de mon jury,

À mon président de jury, Denis Séraphin, professeur des universités en chimie organique à l'Université d'Angers, merci de me faire l'honneur de présider à mon jury de thèse d'exercice. Merci pour la qualité de vos enseignements au long de mon parcours universitaire.

À mon directeur de thèse, Jean-Jacques Helesbeux, maître de conférences en chimie organique à l'Université d'Angers, merci pour l'encadrement et la liberté dans la réalisation de ces travaux et aussi pour les nombreux conseils et la patience dans la rédaction de cette thèse d'exercice.

À Chau Phi Dinh, merci pour ta participation et ton enseignement en badminton

Aux enseignants de l'université d'Angers,

Plus particulièrement, à mes professeurs de la faculté de pharmacie, merci pour la richesse et la passion de vos enseignements.

À ma famille,

À mes parents, merci pour votre confiance et votre soutien.

À ma sœur, merci pour tes relectures attentives.

À Pierrick, merci pour tes nombreux débogages du fichier latex, certainement que ça ne compilerai pas sans toi.

À mon chat, Toulouse, tu as sans doute passer presque autant de temps que moi sur la rédaction (ou à côté) merci pour ton soutien.

À mes amis,

À Vivie, merci pour ton soutien et tes encouragements.

À Miora, Guillaume et Aurélien, merci pour votre inquisition répétée sur l'avancement de la rédaction, j'espère que votre patience sera récompensée.

À Elise, merci pour de m'avoir montré le chemin.

Aux membres de l'unité de recherche SONAS,

Merci à tous pour votre accueil et votre bienveillance.

À Thomas Charpentier, merci pour ta bonne humeur et tes invitations aux pauses cafés, précieuses.

« L'amour fait comprendre à l'âme » Je respire où tu palpites, V.Hugo.

Table des matières

1 Introduction

- 1.1 L'inflammation, définition, composants et rôles physiologiques
- 1.2 Les maladies liées à l'inflammation

2 État de l'art

- 2.1 Biosynthèse, régulation et activation des eicosanoïdes
- 2.2 Biosynthèse des prostaglandines
- 2.3 La prostaglandine E2
- 2.4 PGE2 récepteurs 1-4
- 2.5 Les médicaments anti-inflammatoires
- 2.6 mPGES-1
- 2.7 Les structures de complexes prostaglandine E synthase 1 microsomale (mPGES-1)-ligand disponibles
- 2.8 les différents inhibiteurs de mPGES-1

3 Méthodes et résultats

- 3.1 Chémo-informatique
- 3.2 Pharmacophores
- 3.3 Amarrage moléculaire
- 3.4 Méthodes

4 Conclusion

5 Bibliographie

Annexes

Table des figures

1	Représentation des différents éléments de l'inflammation
2	Acide arachidonique
3	Métabolisme de l'acide arachidonique
4	Biosynthèse des prostaglandines
5	Expression de EP1 et EP2 par type cellulaire
6	Expression de EP3 et EP4 par type cellulaire
7	Représentation des différents anti-inflammatoires stéroïdiens
8	Quelques anti-inflammatoires non stéroïdiens
9	Représentation des salicyllés et fénamates
10	Représentation des acides arylacétiques
11	Représentation des arylpropioniques
12	Représentation des oxicams
13	Représentation des coxibs
14	Expression de mPGES-1
15	Structure de mPGES-1
16	Surface du site actif de mPGES-1
17	Mesures de validation cristallographie
18	Complexes inhibiteurs-mPGES-1
19	Composés essais cliniques mPGES-1
20	Inhibiteurs de mPGES-1 d'origine naturelle
21	Inhibiteur de mPGES-1 : Arylpyrrolizines
22	Inhibiteur de mPGES-1 : Biarylimidazoles
23	Inhibiteur de mPGES-1 : MK-886
24	Inhibiteur de mPGES-1 : Benzoxazole
25	Inhibiteur de mPGES-1 : Phenanthrene imidazole
26	Inhibiteur de mPGES-1 : Imidazoquinolines
27	Inhibiteur de mPGES-1 : Oxicams
28	Inhibiteur de mPGES-1 : Urée tri-substituée
29	Inhibiteur de mPGES-1 : Acide pirinixique
30	Inhibiteur de mPGES-1 : Acide 2-mercaptohexanoïque
31	Inhibiteur de mPGES-1 : Carbazole benzamide
32	Publications "pharmacophore" et "virtual screening"
33	Représentation des différents attributs de pharmacophore
34	Représentation de différents pharmacophores
35	Courbes ROC

36	Représentation du jeu de données pour la validation de pharmacophore de mPGES-1
37	UMAP du jeu de données de mPGES-1
38	Pharmacophore 4YK5-5T36
39	Pharmacophore 5KOI-5TL9
40	Pharmacophore 4YL0-5BQG
41	Pharmacophore 4YL0-5BQH
42	Pharmacophore 4YL3-5BQG
43	Pharmacophore 4YL3-5BQH
44	Pharmacophore 5BQG-5BGH
45	Capture d'écran du logiciel LigandScout permettant des visualisations 2D et 3D de l'interaction ligand-site actif de mPGES-1
46	Composé 1, représentation 2D
47	Composé 1 issu de la base de données ZINC
48	Composé 2, représentation 2D
49	Composé 2 issu de la base de données ZINC
50	Composé 3, représentation 2D
51	Composé 3 issue de la base de donnée COCONUT
52	Composé 4, représentation 2D
53	Composé 4 issu de la base de données COCONUT
54	Composé 5, représentation 2D
55	Composé 5 issu de la base de données UNDP
56	Composé 6, représentation 2D
57	Composé 6 issu de la base de données UNDP
58	Composé 7, représentation 2D
59	Composé 7 issu de la base de données UNDP
60	Composés vue d'ensemble
61	Logigramme jeu de validation
62	Annexe1 : Représentation du format Molfile
63	Annexe2 : Exemples de pharmacophores non sélectionnés
64	Annexe3 : Exemples de poses non sélectionnées

Liste des tableaux

- 1 Comparaison criblage virtuel
- 2 Tableau règles de filtration
- 3 Tableau packages python

Liste des Annexes

Annexe 1 : Représentation de composés chimiques pour l'informatique

Annexe 2 : Exemples de pharmacophores non sélectionnés

Annexe 3 : Exemples de poses d'amarrage moléculaire non retenues

Annexe 4 : Environnement Python utilisé lors de ces travaux

1 Introduction

L'inflammation est une réaction de défense de l'organisme. Ce mécanisme de défense du système immunitaire inné s'active en réponse à une agression exogène, endogène ou environnementale de l'organisme (microbes, toxines, éléments nécrotiques, radiations...). En réponse à cette agression, des mécanismes s'activent induisant la libération de médiateurs inflammatoires. Cette agression est captée par des récepteurs qui vont libérer des médiateurs inflammatoires et recruter les effecteurs, qui vont adapter l'organisme en réponse à cette agression. Cette réaction inflammatoire, par son intensité et sa durée peut devenir nocive (on parle alors d'inflammation non-résolutive ou d'inflammation chronique contrairement à la phase d'inflammation aïgue). Elle est ainsi associée à plusieurs pathologies (asthme, polyarthrite rhumatoïde, certains cancers...). Cette réaction inflammatoire excessive ou prolongée peut nécessiter un traitement anti-inflammatoire qui visent à limiter la libération de médiateurs pro-inflammatoires. Ces médiateurs dérivent de l'acide arachidonique (ARA) par la cascade enzymatique des cyclooxygénases (COX) et des lipoxygénases (LOX). Ainsi, les traitements anti-inflammatoires ont pour rôle pour une part, de limiter la libération d'ARA et d'autre part, d'inhiber les enzymes responsables de la synthèse de ces médiateurs, pour les traitements anti-inflammatoires stéroïdiens et non stéroïdiens respectivement. Ces traitements sont associés à des effets indésirables importants limitant leurs utilisations au long terme. Au cours de ces dernières décennies, la recherche d'alternatives a permis d'identifier et de valider d'autres cibles thérapeutiques. Ces nouvelles cibles présentent une opportunité de développement de nouveaux agents anti-inflammatoires. Parmi celles-ci, mPGES-1 a fait l'objet de nombreux travaux. Ils ont permis d'aboutir à la découverte de plusieurs inhibiteurs (à l'échelle du laboratoire et parfois sans transposition sur modèle animal) de cette enzyme. Ainsi en associant les données de la littérature et les résultats du SONAS, ce travail de thèse d'exercice a visé à établir plusieurs pharmacophores à partir des complexes co-cristallisés de mPGES-1 avec un inhibiteur. Puis, la validation de certains de ces pharmacophores est effectuée à travers un jeu de données de validation établi à partir des différents inhibiteurs connus. Ensuite, avec chacun des pharmacophores validés, un criblage virtuel est effectué sur différentes bases de données. Les meilleurs résultats ont ensuite été soumis à une étude d'amarrage moléculaire pour permettre de valider ces résultats préliminaires.

1.1 L'inflammation, définition, composants et rôles physiologiques

1.1.1 Définition de l'inflammation

L'inflammation, dont les signes caractéristiques sont décrits dans *De Medicina*, par Celsus selon les quatre termes "*rubor et tumor cum calor et dolor*" (rougeur, gonflement, chaleur et douleur) et complété au deuxième siècle par "*functio laesa*" (perte de fonction)

par Galien [1]], se définit comme la "réponse tissulaire à une infection ou à une lésion qui sert à éliminer ou emmurer l'infection ou la lésion" [2]. L'inflammation est une réponse du système immunitaire indispensable à la vie pour défendre l'organisme. La réponse inflammatoire fait appel à un ensemble d'acteurs dans un réseau d'interactions complexes et dynamiques, en réponse aux inducteurs de l'inflammation, qui entraîne la libération d'un grand nombre de médiateurs par plusieurs acteurs du système immunitaire, notamment du système immunitaire inné. Le système immunitaire inné, dérivé de la lignée cellulaire myéloïde (granulocytes, macrophages et cellules dendritiques) et des lymphocytes cytotoxiques naturels issus de la lignée lymphoïde aussi appelés lymphocytes NK et de l'ensemble des protéines sériques du système du complément, poursuivie dans sa réponse par les anticorps du système humoral et des lymphocytes de l'immunité adaptative. L'inflammation est une réaction dynamique qui se déroule en plusieurs phases. La phase inflammatoire est suivie d'une phase de résolution de l'inflammation, conduisant à la phase de guérison/cicatrisation. Dans cette situation il s'agit d'un processus contrôlé. La notion d'inflammation aiguë est employée. Quand la phase inflammatoire est non résolutive et aboutit sur une inflammation prolongée dans le temps, l'inflammation est chronique. La réponse inflammatoire et immunitaire, bien qu'indispensable à la vie, peut aussi parfois se retrouver disproportionnée mettant en danger la vie de l'hôte, ou se retourner contre l'hôte, formant les maladies auto-immunes.

1.1.2 Les signes et composants de l'inflammation

1.1.2.1 Les signes de l'inflammation

La réponse inflammatoire se caractérise par les 4 signes cardinaux de l'inflammation : chaleur, rougeur, gonflement et douleur, qui s'expliquent aujourd'hui par la vasodilatation (chaleur et rougeur) ralentissant le flux sanguin et augmentant la perméabilité capillaire en réponse aux médiateurs pro-inflammatoires sécrétés permettant l'extravasation de certains acteurs de l'inflammation et une fuite de plasma vers le site nécessitant la réponse inflammatoire.

Comme les signes de l'inflammation, on peut aussi séparer les acteurs de l'inflammation en 4 composants distincts :

- Les inducteurs : ils sont la cause de la réponse inflammatoire
- Les récepteurs : ils détectent les inducteurs de l'inflammation et permettent le déclenchement de la réponse inflammatoire.
- Les médiateurs : ils permettent de recruter les acteurs de l'inflammation et transmettent la réponse inflammatoire
- Les acteurs : ils visent à éliminer ou réguler les inducteurs de l'inflammation.

1.1.2.2 Inducteurs de l'inflammation

Il existe différents éléments stimulant l'inflammation :

- Agents pathogènes : les virus, bactéries, parasites, toxines sont les causes les plus fréquentes d'inflammation et pour lesquelles la réaction inflammatoire est la mieux caractérisée.
- Allergènes : ils entraînent une réaction inflammatoire excessive par rapport à un stimulus ne présentant pas un danger pour l'organisme.
- Corps étrangers : ils regroupent de nombreux éléments différents (suture, prothèse, cristaux d'urée par exemple) provoquant l'inflammation car ils sont non reconnus par le soi ou par les lésions tissulaires induites comme les lésions mécaniques induites par les cristaux.
- Facteurs physiques et chimiques : ces facteurs sont liés à l'environnement. Comme facteur physique, à faible dose, les radiations induisent une réponse inflammatoire. Un exemple de facteur chimique est un choc osmolaire, ou métabolique (ischémie/re-perfusion) qui entraîne aussi une réaction inflammatoire.

Ainsi, les inducteurs de l'inflammation se distinguent par leurs origines exogènes ou endogènes.

Les inducteurs exogènes : Ces inducteurs se catégorisent entre : agents microbiens et non-microbiens (corps étrangers, toxines non microbiennes, corps étrangers, composés toxiques). Parmi les inducteurs exogènes d'origine microbienne, on distingue d'une part les motifs moléculaires associés aux pathogènes (MAMP) et d'autre part les facteurs de virulence. Les premiers sont un ensemble de motifs moléculaires conservés au sein des différentes classes de micro-organismes (commensaux et pathogènes) [3]. Ils sont reconnus par les récepteurs à reconnaissance de motifs (RRM) (cf Récepteurs de l'inflammation). Les facteurs de virulence spécifique d'un pathogène, quant à eux ne sont pas reconnus directement par un type de récepteur particulier, mais augmentent la virulence des pathogènes et contribuent à déclencher une réponse inflammatoire. Par exemple, la pneumolysine de *Streptococcus pneumoniae* est une cytolysine cholestérol dépendante responsable de l'activité cytolytique [4]. Les inducteurs exogènes non microbiens, allergènes, irritants, radiations, composés toxiques et corps étrangers peuvent entraîner une réponse inflammatoire, par exemple, en mimant l'activité protéolytique des helminthes [5], ou les corps étrangers en stimulant l'inflammasome NLRP3 [6]; d'autres sont irritants des épithéliums muqueux et sont éliminés par clairance des épithéliums.

Les inducteurs endogènes Ils sont issus de stress, de tissus endommagés ou dysfonctionnels [7]. Par exemple, la mort cellulaire existe sous

deux mécanismes majeurs. La mort cellulaire programmée qui n'entraîne pas de déclenchement de l'inflammation [8].

À l'inverse, l'autre mécanisme la mort cellulaire non programmée aussi appelée nécrose cellulaire, relarguent des composants qui déclenchent l'inflammation [9]

Parmi les produits relargués, on trouve :

- Des protéines, notamment la protéine du groupe box 1 à haute mobilité (HMGB1), découverte en 1999 [10]] augmente la production de facteur de nécrose tissulaire α ($\text{TNF}\alpha$) par les monocytes quand elle se retrouve dans le milieu extracellulaire [11]], ou de protéines chaperonnes comme la protéine de liaison à l'ARN inductible à froid (e-CIRP), qui active la voie de facteurs nucléaires améliorateurs de chaîne légère kappa de cellules B activées ($\text{NF-}\kappa\text{B}$) par fixation sur son récepteur TLR (MD2) [12].
L' e-CIRP active aussi l'inflammasome NRLP3 (NRLP3). Également, les histones formant habituellement le nucléosome en complexe avec l'ADN entraînent des hauts niveaux de cytokines ($\text{TNF}\alpha$, l'interleukine 6 (IL-6) et l'interleukine 10 (IL-10)). Au niveau de l'endothélium, le collagène active le facteur de Hageman (facteur XII), ce qui entraîne des cascades enzymatiques induisant une réponse inflammatoire [13]
- Des macromolécules et nucléotides, de l'ADN libre tant viral que bactérien ou de l'hôte [14]], de même que pour l'ADN mitochondrial [15]], entraînent la production de $\text{TNF}\alpha$ par activation de la voie $\text{NF-}\kappa\text{B}$ par liaison au récepteur de type Toll (RTT)TLR9.

Les inducteurs endogènes et exogènes sont reconnus par divers récepteurs, qui reconnaissent les MMAP ou les molécules libérées dans des conditions de stress du milieu, motifs moléculaires associés aux dégâts (MMAD). Ces inducteurs de l'inflammation expriment des motifs moléculaires qui sont reconnus par les récepteurs portés par les acteurs de l'inflammation.

1.1.2.3 Récepteurs de l'inflammation

Les MMAP et MMAD sont reconnus par des récepteurs, ce qui entraîne la réponse immunitaire et inflammatoire. Ces récepteurs sont appelés RRM, les mieux connus sont les RTT et les récepteurs de type Nucleotide-binding oligomerization domain (NOD). C'est la détection par ces récepteurs qui est la première étape de la réponse immunitaire. En plus des récepteurs des MMAP et MMAD , des protéines circulantes peuvent aussi reconnaître et/ou détruire les agents infectieux.

- Les récepteurs des MMAP : plusieurs se distinguent :
 - Des récepteurs homologues à Toll : Découverts au début du XX^{eme} siècle chez les drosophiles [16]], plusieurs RTT se trouvent chez l'Homme [17]],

RTT 1 à 10. La partie cytoplasmique de ces récepteurs présente une grande similarité avec les récepteurs à l'interleukine 1 (IL-1). Ils reconnaissent les pathogènes au contact de la cellule ou dans le milieu extracellulaire [18]].

L'activation des RTT conduit à l'activation de NF- κ B médiée par la protéine adaptatrice MyD88 et à la translocation du facteur de transduction NF- κ B qui entraîne la production de cytokines (TNF α , interleukine 12 (IL-12), interféron de type 1 (IFN)) [19]].

- les récepteurs de type Nod (NLR) : contrairement aux RTT, les NLR reconnaissent les bactéries intracellulaires. NOD1 et NOD2 sont des récepteurs qui présentent des structures similaires. Ils s'activent tous deux en présence de peptidoglycanes mais les motifs reconnus sont différents. NOD1 reconnaît spécifiquement un motif porté sur les bactéries Gram négatif, et NOD2 un motif commun à toutes les bactéries. L'activation de NLR conduit au recrutement de Rip2/RICK et mène à l'activation de NF- κ B [14]].

- Les récepteurs des MMAD

- Les récepteurs RTT et NLR reconnaissent majoritairement les MMAP. Certains RTT peuvent aussi reconnaître les MMAD, notamment RTT2 et 4 qui reconnaissent notamment HMGB1 [20]].
- Les récepteurs de lectines de type C (LTC) sont plus communément associés à la reconnaissance des champignons. Certains peuvent reconnaître les MMAD. Notamment, DNGR-1 [21]] qui reconnaît les filaments d'actine libérés lors de la nécrose [22]] et diminue la production d' IL-10[23]].
- Les récepteurs des produits terminaux de glycation (PTG) sont exprimés par plusieurs types cellulaires, monocytes, neutrophiles, cellules endothéliales. Leur expression se trouve augmentée dans plusieurs maladies liées à l'inflammation(fibrose pulmonaire, polyarthrite rhumatoïde, la maladie d'Alzheimer, lupus érythémateux). Ils conduisent à l'activation de NF- κ B et à la production de cytokines(TNF α et IL-1) [24]].
- Les récepteurs couplés aux protéines G (RCPG) et canaux ioniques : Le relargage de nucléotides dans l'espace extracellulaire peut activer les récepteurs purinergiques de surface P2. Ces récepteurs sont divisés entre P2YR et P2X, récepteurs couplés aux protéines G et récepteurs ionotropes respectivement. Par exemple, P2Y2R promeut la migration et l'activation des acteurs de l'inflammation (neutrophiles, éosinophiles, cellules dendritiques et macrophages) mais P2Y2R est aussi associé à l'inflammation incontrôlée et aux maladies inflammatoires chroniques, athérosclérose et asthme [25]]. P2X7 joue un rôle dans de nombreuses maladies inflammatoires par l'activation de plusieurs

- facteurs de transcription dont NF- κ B et la production de cytokines et chemokines de différentes cellules immunitaires [26].
- Les protéines circulantes : Certaines protéines(lectines) présentes dans la circulation peuvent reconnaître les éléments du non-soi portés par les micro-organismes. Par exemple, la lectine liée au mannose active la cascade du complément par la voie des lectines qui est indépendante des anticorps [27].

1.1.2.4 Médiateurs de l'inflammation

Les médiateurs sont des molécules qui intensifient et prolongent l'état inflammatoire, ils sont issus de dérivés protéiques ou sécrétés directement par les cellules (macrophages et mastocytes notamment). Ils peuvent être stockés dans des granules (sérotonine et histamine) ou circuler sous forme inactive dans le plasma. La plupart ont une durée de vie courte. Les médiateurs sont regroupés dans différents groupes : amines vasoactives, peptides vasoactifs, fragments du complément, médiateurs lipidiques, cytokines et chimiokines, et enzymes protéolytiques.

- Amines vasoactives : L'histamine (2-[3H-imidazol-4-yl]ethanamine) et la sérotonine (5-hydroxytryptamine) sont deux amines vasoactives. Stockées dans des granules, elles sont libérées rapidement lors d'une inflammation.
 - L'histamine cause une dilatation artérielle, une augmentation de la perméabilité vasculaire et la contraction des cellules musculaires lisses au niveau gastro-intestinal et broncho-pulmonaire. L'effet pléiotrope de l'histamine est médié par ses différents récepteurs. L'histamine favorise la migration des éosinophiles sur le site d'infection et le recrutement des mastocytes augmentant la réponse inflammatoire et favorisant le passage à un état inflammatoire chronique [28].

L'histamine joue aussi sur d'autres populations cellulaires. Au niveau des macrophages pulmonaires, l'histamine augmente la production de cytokines pro-inflammatoires IL-6 et β -glucuronidase et au niveau cutané, elle favorise le recrutement des lymphocytes "natural killer" (NKL) et la production de chemokines, ce qui fait de l'histamine le principal médiateur dans les dermatites atopiques [29].

 - Sérotonine : La sérotonine (5-hydroxytryptamine) est, pour les cellules immunitaires, localisée majoritairement au niveau des plaquettes et des mastocytes. La sérotonine, en plus de son rôle de vasoconstricteur, favorise le recrutement des cellules immunitaires (mastocytes, cellules dendritiques, neutrophiles et éosinophiles) et augmente la production d'interféron γ par les NKL [30] et la production de cytokines (IL-6) par activation de NF- κ B [31].
- Les peptides vasoactifs : Ils sont stockés sous forme active dans des vésicules (par exemple : la substance P) ou dérivés de précurseurs inactifs sous l'action de protéases(par

exemple : les kinines). La substance P est principalement produite par les neurones, et joue un rôle dans la nociception et l'inflammation. La substance P, en se liant à son récepteur couplé aux protéines G, notamment le récepteur à la neurokinine 1, conduit à l'activation de cible de la rapamycine chez les mammifères (mTOR), une enzyme de la famille des sérines/thréonines kinases qui entraînent la production de cytokines et chimiokines, ainsi qu'à l'activation de facteurs de transcription NF- κ B. La substance P favorise également la synthèse de métabolites de l'ARA (prostaglandines E2 et thromboxanes B2) [32]

Le facteur de Hageman, active le système kinine-kallikréine, des protéases qui clivent le kininogène pour aboutir à la formation de bradykinine. La bradykinine, entraîne une vasodilatation et augmente la perméabilité vasculaire, favorisant le recrutement des cellules immunitaires et entraîne la contraction des cellules musculaires lisses, notamment au niveau pulmonaire qui conduit à l'obstruction des voies respiratoires(asthme) [33]

]. De plus, le facteur de Hageman peut amplifier la réponse inflammatoire en activant le système du complément.

- Les médiateurs lipidiques : Ils représentent des composés importants dans la médiation de l'inflammation et les travaux de cette thèse portent sur une partie d'entre eux. Ils sont traités spécifiquement au chapitre consacré.
- Le système du complément : C'est un ensemble de protéines solubles et de leurs récepteurs membranaires associés, hautement conservés. Il intervient sur l'immunité innée et adaptative. Il existe plusieurs voies d'activation du complément. Parmi les différents fragments du complément, les anaphylatoxines (C3,C4,C5) entraînent la dégranulation des cellules endothéliales, phagocytes et mastocytes ; et favorisent la production de radicaux oxygénés. Plusieurs fragment protéiques du complément ont aussi un rôle de chemoattractant, dans le recrutement des leucocytes [34]
- Les cytokines et chemokines : Ce sont des protéines sécrétées présentant plusieurs fonctions de croissance, de différenciation et de régulation. Elles régulent la réponse immunitaire et la migration cellulaire.

— Les cytokines peuvent avoir un rôle pro-inflammatoire ou anti-inflammatoire. Elles sont produites par plusieurs types cellulaires, macrophages et mastocytes principalement mais aussi par les lymphocytes T dans l'immunité acquise. Les cytokines pro-inflammatoires majoritairement produites par les monocytes sont : le facteur de nécrose tumorale, les interleukines et les interférons.

— Facteur de nécrose tissulaire : Le TNF α [35] est produit par les macrophages, les cellules endothéliales et les mastocytes, pour l'immunité innée. Sa sécrétion est déclenchée par le lipopolysaccharide sur les récepteurs RTT. Il agit à plusieurs niveaux, stimulant l'immunité contre les cellules cancéreuses et favorisant la production de molécules d'adhésion

(ICAM, VCAM et la E-sélectine) permettant la progression de la circulation vers le site de l'inflammation.

- Les interleukines : Elles [35]] représentent l'autre grande famille de cytokines. Elles peuvent présenter une activité pro-inflammatoire (IL-1, IL-6, IL-8, IL-12, IL-15, IL-18, et IL-23) et d'autres anti-inflammatoires (IL-10, IL-4, IL-13, IL-1Ra). L'interleukine 8 est classée à la fois comme cytokine et comme chemokine. L'IL-1 possède plusieurs activités biologiques. Elle active les lymphocytes T mais agit sur le système nerveux central déclenchant la fièvre. Elle contribue également au pic protéique de la phase aigüe avec la production de protéine C-réactive, et de facteurs du complément. Elle contribue aussi à la production de molécules d'adhésion similairement au $\text{TNF}\alpha$
- Les interférons : Ils [35]] limitent la réplication virale, soit en inhibant la réplication des cellules infectées, soit en protégeant les cellules non atteintes et en stimulant l'immunité antivirale.
- Les chemokines : Elles représentent un groupe de plusieurs dizaines de molécules (8-12kD) agissant sur un grand nombre de récepteurs. Originellement caractérisées par leur rôle de chimiotactisme, leur rôle s'avère plus complexe. Elles présentent également un rôle dans l'homéostasie, l'hématopoïèse et la régulation des lymphocytes [36]].
- Les enzymes protéolytiques : Elles ont plusieurs rôles dans l'inflammation, notamment en dégradant la matrice extracellulaire et les protéines membranaires. Par exemple, les métalloprotéinases peuvent réguler l'intégrité des barrières physiques et la migration des leucocytes de la circulation aux tissus. Elles régulent la disponibilité et l'activité des cytokines et chemokines. Certaines métalloprotéinases peuvent activer IL-1 β , d'autres peuvent la dégrader. Elles peuvent aussi générer un gradient de chemokines, favorisant le recrutement des cellules immunitaires, mais également réguler la survie de ces cellules [37].]

1.1.2.5 Acteurs de l'inflammation

Les acteurs de l'inflammation sont les cellules et tissus qui réagissent aux médiateurs de l'inflammation, permettant de maintenir l'homéostasie ou de l'adapter dans des conditions nocives [38].]

Les monocytes/macrophages ont un rôle de supervision, ils peuvent phagocyter et présenter les antigènes comme les cellules dendritiques pour initier la réponse immunitaire adaptative. Ils relarguent également des cytokines et chémokines régulant la réponse inflammatoire et le recrutement des polynucléaires (neutrophiles, éosinophiles, basophiles). Les neutrophiles sont particulièrement nombreux en comparaison des éosinophiles et basophiles ; les éosino-

philes ont un rôle plus important en cas d'infection par des parasites ou de réponse allergique, par exemple la protéine basique majeure (MBP) contenue dans leurs granules tue les larves d'helminthe in vitro [39].

Les mastocytes, comme les éosinophiles, ont plus un rôle sécrétoire dans la réponse inflammatoire. Ils sont particulièrement impliqués dans les cas d'allergies et réactions d'hyper-sensibilités.

L'inflammation comporte donc un ensemble d'acteurs et d'éléments repris dans la figure 1

Représentation des différents éléments de l'inflammation

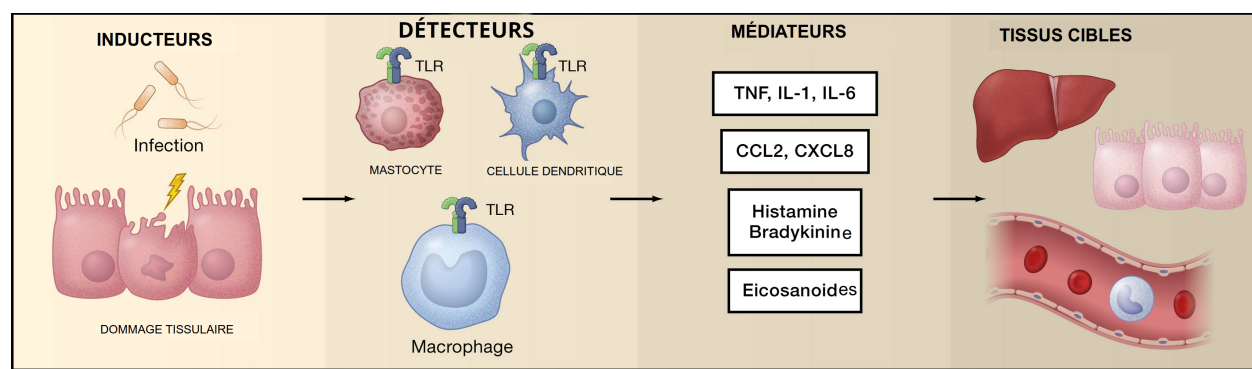


FIGURE 1 – Représentation des éléments de l'inflammation : inducteurs, détecteurs, médiateurs et tissus cibles, adapté des travaux de Medzhitov R. [40]

1.1.3 Résolution de l'inflammation et inflammation chronique

L'inflammation est normalement suivie d'une phase de résolution permettant un retour à une fonction normale du lieu d'inflammation. L'inflammation est alors caractérisée d'aiguë, c'est une inflammation rapide d'une durée de quelques minutes ou heures, présentant des lésions tissulaires ou une fibrose (permettant de réparer les tissus) limitées, accompagnée de symptômes importants comparée à une inflammation chronique qui est plus à bas bruit. Dans l'inflammation aiguë les détecteurs activent les macrophages et les mastocytes, qui relarguent des médiateurs pro-inflammatoires. Ces derniers recrutent les neutrophiles circulant vers le site inflammatoire. À cet endroit, ils sont activés par les cytokines ou par contact avec un agent pathogène et ils libèrent le contenu toxique de leurs granules pour éliminer le pathogène (dérivés réactifs de l'oxygène, protéinase 3, cathepsine G, élastase) [41]. Cette action, en plus de neutraliser le pathogène, endommage aussi par ses composés toxiques le milieu environnant

[42]], ne distinguant pas spécifiquement l'agent pathogène. Ce relargage des médiateurs pro-inflammatoires, notamment de la prostaglandine E2 (PGE2), entraîne à la suite un changement de production des médiateurs lipidiques pro-inflammatoires pour ceux contribuant à la résolution de l'inflammation (lipoxines, résolvines, protectines et marésines) [43]] et à une phase de réparation menée par les macrophages [44].]. Ces médiateurs lipidiques anti-inflammatoires limitent le recrutement des neutrophiles et promeuvent le recrutement des monocytes pour nettoyer le site inflammatoire [44]. Si cette étape n'est pas déclenchée et que le phénomène inflammatoire persiste, ce dernier acquiert de nouvelles caractéristiques. L'inflammation bascule dans un état chronique. L'exemple le plus connu d'inflammation chronique est l'inflammation autour de *Mycobacterium tuberculosis* qui peut persister plusieurs années [45].]. Ainsi, de multiples facteurs peuvent contribuer à l'inflammation chronique, une stimulation persistante ou une stimulation répétée à une toxine mais aussi une réaction inflammatoire excessive ou à l'inverse insuffisante [45].].

1.2 Les maladies liées à l'inflammation

L'inflammation permet d'activer le système immunitaire et aide l'organisme à se défendre contre une infection. Elle est aussi associée, notamment lors qu'elle devient chronique, au développement de plusieurs pathologies.

1.2.1 L'athérosclérose

L'athérosclérose se caractérise par une accumulation de lipides et de débris tissulaires formant une plaque dans l'intima vasculaire [46]]. L'athérosclérose est liée à plusieurs complications comme l'infarctus du myocarde ou l'accident vasculaire cérébral [47].]. Le lien entre athérosclérose et inflammation a été posé par Russel Ross en 1999 [48]. En effet, les macrophages ont un rôle important dans le développement et la régulation de la pathologie [49].]. De plus, le célécoxib, un inhibiteur spécifique des COX2 pourrait avoir un intérêt dans la prise en charge de cette pathologie [50].]. Plus précisément, diminuer l'activité des métalloprotéinase matricielles éviterait la rupture de plaques en limitant la dégradation du collagène mais en pratique le célécoxib déséquilibre l'homéostasie vasculaire en bloquant la production de prostacycline déconseillant son utilisation prolongée dans cette pathologie (essai APC).

1.2.2 L'asthme

L'asthme est une maladie multifactorielle qui associe des facteurs environnementaux avec de l'inflammation et une hypersensibilité bronchique qui conduit à une obstruction des voies respiratoires où NF- κ B et RTT jouent un rôle important [51].

L'asthme est une maladie respiratoire commune qui touche environ 7% de la population adulte en France et 10% des enfants de plus de 10 ans [52].

L'inflammation joue un rôle important dans l'asthme, à ce titre plusieurs prostaglandines (PG) sont impliquées ; principalement la prostaglandine D2 qui provoque une bronchoconstriction forte [53].

La PGE2 semble avoir un rôle nuancé au niveau des voies pulmonaires, l'inhalation de PGE2 inhibe la réponse asthmatique [54].

1.2.3 La polyarthrite rhumatoïde

La polyarthrite rhumatoïde est une maladie où plusieurs facteurs entrent en jeu. Des facteurs génétiques et environnementaux (la cigarette, l'obésité, l'infection à *Porphyromonas gingivalis* et l'inflammation, qui conduisent à la modification d'auto-antigènes [55].

Cette pathologie touche environ 1% de la population mondiale, avec une prédominance pour les femmes et la population âgée [56].

Elle se caractérise par une inflammation persistante de la membrane synoviale qui finit par endommager le cartilage et les os et entraîne une déformation des articulations [57].

Les anti-inflammatoires non stéroïdiens (AINS) réduisent la douleur et les gonflements associés à la pathologie mais ne suffisent pas à soigner la maladie et sont associés à plusieurs effets indésirables (hépatique, rénal, gastro-intestinal et cardiovasculaire) [58].

1.2.4 Les cancers

L'hypothèse d'un lien entre inflammation et cancer a été avancée depuis longtemps par Virchow [59].

Par exemple, l'infection à *Helicobacter pylori* et l'inflammation chronique associée sont liées avec le développement du cancer de l'estomac [60].

Parmi les molécules impliquées dans les cancers liés à l'inflammation, plusieurs molécules majeures ont été identifiées, notamment le facteur de transcription NF- κ B et les cytokines inflammatoires [61].

Plus généralement 15% des cancers sont associés aux infections et à l'inflammation [60].

Ainsi le traitement par AINS montre une amélioration de la qualité de vie du patient dans plusieurs études [62]

et réduit l'incidence des tumeurs pour certains cancers [63].

Par exemple, ces études montrent que le sulindac réduit la formation de polypes chez des patients atteints du syndrome de Garner ou de polypose adénomateuse familiale.

1.2.5 Les troubles métaboliques

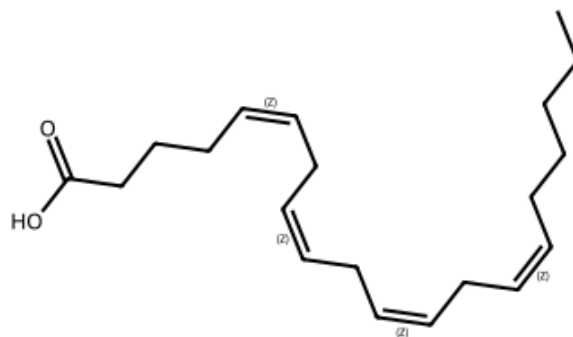
Plusieurs troubles métaboliques impliquent une inflammation chronique et/ou sont liés à d'autres maladies inflammatoires. Par exemple le diabète de type 2 favorise le développement et la progression de l'athérosclérose [64].]

Le $\text{TNF}\alpha$ pro-inflammatoire est aussi impliqué dans la régulation du métabolisme lipidique, son expression est augmentée dans le cadre de l'obésité. L'obésité favorise l'inflammation par l'expression de ce facteur [65].]

2 État de l'art

2.1 Biosynthèse, régulation et activation des eicosanoïdes

Les eicosanoïdes sont une famille de médiateurs lipidiques impliqués dans l'inflammation (initiation et résolution) et dans plusieurs autres processus physiologiques tels que la régulation de la pression artérielle, la perception de la douleur, la survie cellulaire, le recrutement et la fonctionnalisation des cellules immunitaires [66].
Ils se composent de dérivés oxydés de l'ARA, acide gras polyinsaturé à 20 atomes de carbone, présentant 4 doubles liaisons. Les travaux de Burr *et al.* [67]



Acide arachidonique

FIGURE 2 – Représentation de l'acide arachidonique (acide (5Z,8Z,11Z,14Z)-icosa-5,8,11,14-tétraénoïque), dessiné à l'aide de RDkit.

ont montré l'importance dans l'alimentation de l'apport des acides gras insaturés essentiels. Notamment la famille des ω -6 à laquelle appartient l'acide linoléique dont dérive l'ARA [68].
[69] Les travaux, à la fois, de Van Dorp *et al.* [70]

ont montré le lien entre l'acide arachidonique et les médiateurs lipidiques de l'inflammation, en démontrant que l'acide arachidonique est précurseur de la synthèse des PG. Plusieurs voies de biosynthèse existent au sein des eicosanoïdes :

- La voie des COX (prostaglandine-endoperoxyde synthases) pour la formation des PG et des thromboxanes (TX).

- La voie des LOX (HPETE) pour la formation des leucotriènes (LT) et des lipoxines.
- La voie du cytochrome P450(EET, HETE, DHET) conduisant à la formation d'acide époxyeicosaénoïque qui après l'action d'une hydrolase donnera de l'acide dihydroxyeicosaénoïque. Cette voie permet, comme les deux précédentes, aussi la formation d'acide hydroxyeicosatétraénoïque [71].

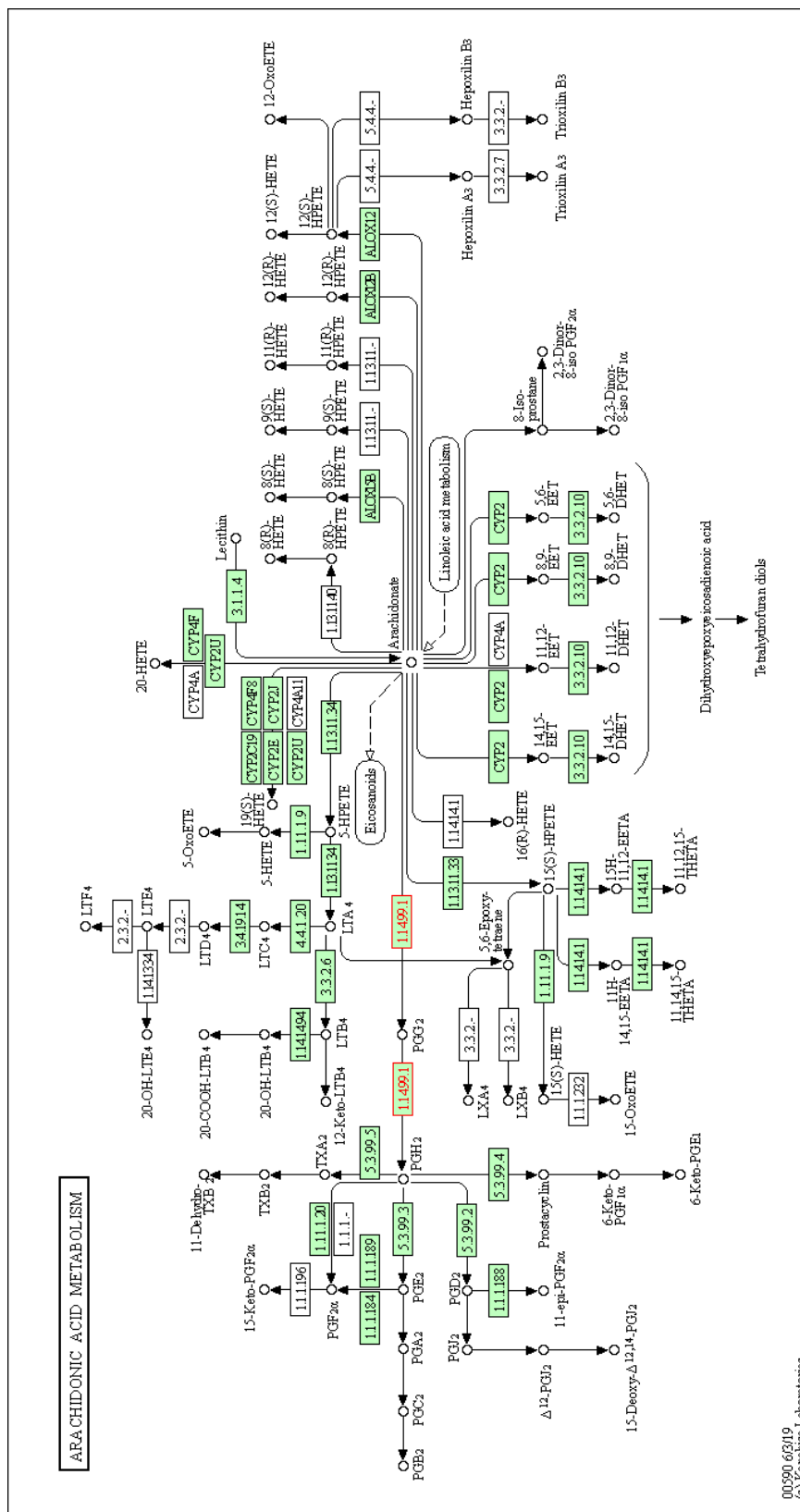


FIGURE 3 – Métabolisme de l'acide arachidonique chez *Homo sapiens*, issue de *Kyoto Encyclopedia of Genes and Genomes* [kanehisakeggky].

Comme le montre la figure 3, le métabolisme complet de ARA est complexe. L'enzyme mPGES-1 est une PGE2 synthase, c'est le métabolisme des prostaglandines qui sera détaillé par la suite, les travaux de cette thèse d'exercice portant sur celle-ci.

2.2 Biosynthèse des prostaglandines

Les PG sont synthétisées *de novo* car elles présentent une demi-vie courte et ne peuvent être stockées. L'ARA est produit par hydrolyse des glycérophospholipides, par la phospholipase A2, phospholipase du groupe IV, spécifique de l'ARA [73].

À la suite de la production d'ARA, les COX, 2 isoformes de la prostaglandine H2 synthase, transforment l'ARA pour former la PG H2. Cette prostaglandine intermédiaire peut être modifiée par un ensemble d'enzymes aboutissant à des fonctions physiologiques différentes (cf figure 4). Paradoxalement, ces différentes enzymes ne présentent de spécificités communes pour la liaison de la PGH2 [74].

La PGI2 et le TXA2, respectivement formés par la prostacycline synthase et la thromboxane-A synthase, jouent un rôle important dans l'homéostasie vasculaire [75, 76].

La PGD2, formée par la prostaglandine D2 synthase, est un médiateur de l'inflammation et de l'allergie. Elle agit principalement sur les mastocytes et les lymphocytes T helper 2 [77].

La PGF2 α , agit comme une hormone, participant à de nombreuses fonctions physiologiques, notamment sur la contraction et relaxation des muscles lisses. Elle peut aussi être administrée pour l'accouchement et l'avortement [78].

Elle est synthétisée à partir de PGH2, PGE2 ou PGD2 [78].
par les PGF2 synthases.

Enfin la PGE2 est transformée à partir de glspgH2 par les PGE2 synthases. Ces enzymes présentent des activités différentes suivant le statut inflammatoire de l'hôte.

2.2.1 Les PGE2 synthases

Elle sont trois :

- prostaglandine E2 synthase cytosolique (cPGES) : C'est une enzyme de 26 kDa constitutive et non affectée par les stimulus inflammatoires. L'activité de cPGES est augmentée par la présence de glutathion (GSH). L'étude de cotransfection des gènes de cPGES et de COX1 et COX2 dans des cellules HEK293 montre une augmentation de la production de PGE2 quand cPGES est associée à COX1, une production requise pour l'homéostasie cellulaire [79].
- prostaglandine E synthase 2 microsomale (mPGES-2) : Une protéine isolée en 1999 par Watanabe *et al.* qui possède une activité PGE2 synthase indépendante du GSH

[80].]. Purifiée par Tanikawa *et al.*, elle a un poids moléculaire de 33 kDa [81].].

La mPGES-2 est produite au niveau de l'appareil de Golgi avant que la partie N terminale hydrophobe soit clivée et que la protéine soit relarguée dans le cytoplasme. La mPGES-2 peut être couplée avec COX1 ou COX2 pour produire PGE2 en réponse à une inflammation aiguë ou chronique mais présente une activité légèrement préférentielle pour COX2 [82].].

- mPGES-1 : Identifiée pour la première fois en 1999 par Jacobson *et al.* [83].

La mPGES-1 est d'abord rapprochée de la glutathion S-transférase microsomale 1 (MGST1) avec laquelle mPGES-1 partage 38% d'identité au niveau de la séquence en acides aminés [84].]. Toutes

deux appartiennent à la superfamille des MAPEG ("Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism"), les autres protéines humaines appartenant à cette superfamille sont : FLAP, LTC4S, MGST2 et MGST3 [85].

La mPGES-1 est une isomérase, GSH dépendante qui produit la PGE2.

2.3 La prostaglandine E2

2.3.1 PGE2, inflammation et tumorigénèse

Dans l'inflammation, PG E2 est associée à l'inflammation, la douleur et la fièvre.

La douleur est un symptôme lié à l'inflammation périphérique qui entraîne une hypersensibilité à la douleur. Également par la PGE2 spinale, par l'action sur ses récepteurs EP1 et EP3, est un médiateur de l'allodynie et de l'hyperalgésie [86].].

La fièvre est un autre symptôme associé à l'inflammation. Elle est contrôlée par le système nerveux central et la PG E2 par sa taille et son caractère lipophile peut passer la barrière hémato-encéphalique. L'augmentation de la production de PG E2 par l'action des cytokines inflammatoires ou des lipopolysaccharides pyrogènes induit de la fièvre par l'action sur son récepteur EP3 [87].].

L'inflammation est associée principalement avec la PG E2 causant l'augmentation de la perméabilité membranaire, douleur, fièvre, dilatation et contraction musculaire. La production de PG E2 est associée avec l'augmentation de l'expression des COX, de même que l'expression de mPGES-1 est augmentée avec l'inflammation [88].

La PG E2 est aussi impliquée dans la tumorigénèse. Dans le carcinome colorectal, la PG E2 est associée avec la survie cellulaire, la croissance, la motilité et l'adhésion cellulaire [89].].

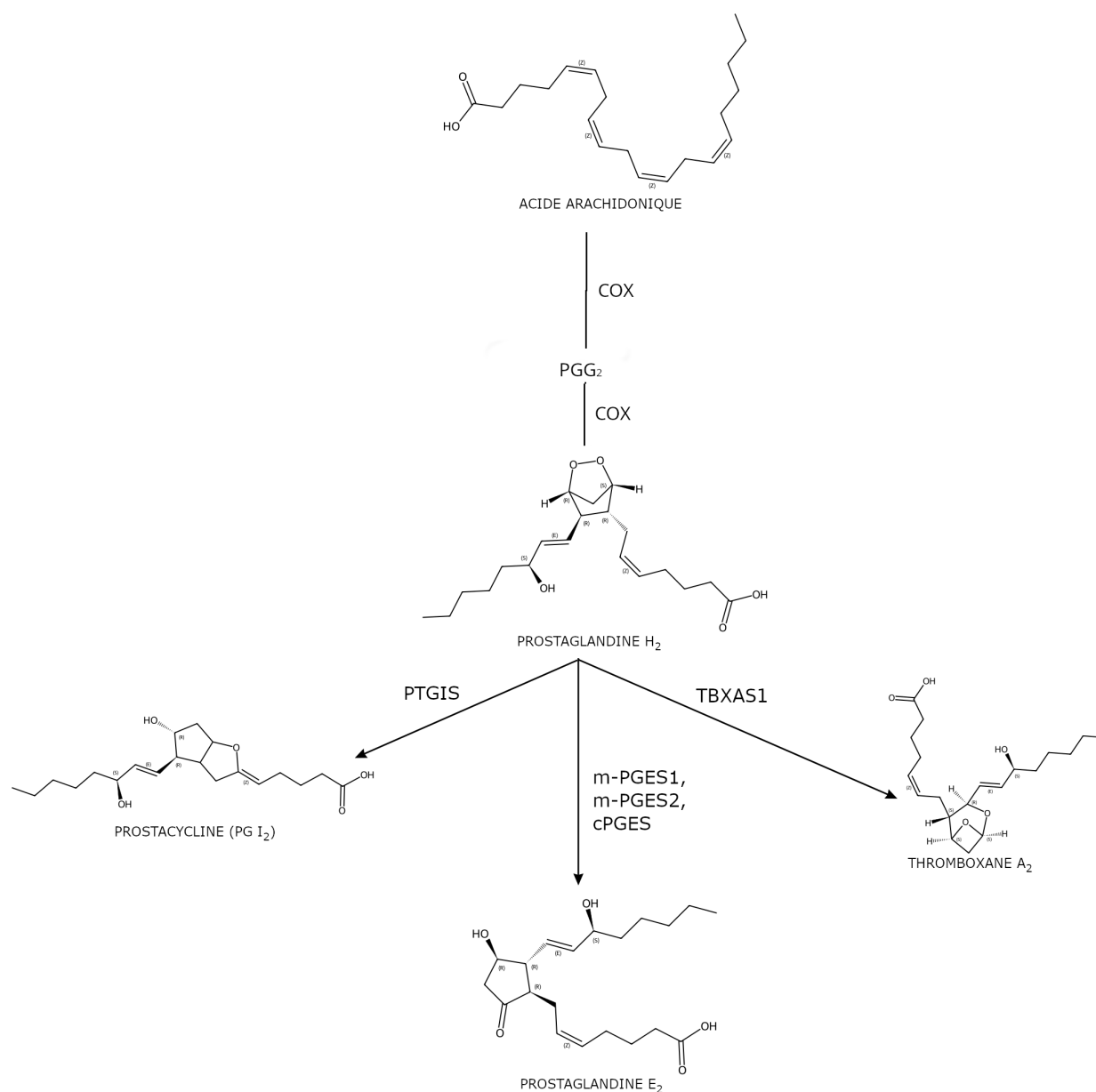


FIGURE 4 – *Biosynthèse des prostaglandines, à partir de l'acide arachidonique aboutissant à la formation des différentes prostaglandines, dessiné à l'aide de RDKit.*

2.3.2 PGE2 et autres implications physiologiques

PG E₂, en plus de son rôle dans l'inflammation, possède aussi plusieurs fonctions physiologiques. La PG E₂, facilite la formation d'ostéoclastes participant à la résorption osseuse induite dans l'inflammation [90].

Dans l'ovulation, les PG E₂ jouent un rôle de médiateur des gonadotropines [91].

En plus, PG E₂ a aussi un rôle dans la fonction rénale. Dans la fonction rénale, mPGES-1 est exprimée constitutivement au niveau des épithéliums, des tubules contournés distaux, des canaux collecteurs médullaires, et aux niveaux des cellules épithéliales transitionnelles de la

vessie et de l'uretère [92].
namique glomérulaire.

]. Elle influe sur l'hémody-

2.4 PGE2 récepteurs 1-4

La PG E2 agit suite à la fixation sur son récepteur. Plusieurs sous-types de récepteurs à la PG E2 existent, au nombre de quatre, EP1, EP2, EP3 et EP4; ils présentent aussi une constante de dissociation différente entre récepteurs et présente une similarité d'environ 30% entre eux [93].

Les 4 récepteurs EP1-4(codés par PTGER1-4) sont tous des récepteurs couplés aux protéines G. Cependant, comme ils présentent une expression différente suivant le récepteur et le type cellulaire (figure 5, 6), ils sont aussi couplés à des protéines G différentes, médiant des effets biologiques différents.

Le récepteur EP1 est couplé à une protéine G_q, active une phospholipase C qui entraîne la production d'inositol-triphosphate et de diglycéride et conduit à l'augmentation du calcium intracellulaire.

Le récepteur EP2 est couplé à une protéine G_s, active une adénylate cyclase et augmente la concentration de l'adénosine monophosphate cyclique.

Le récepteur EP3 est couplé à une protéine G_i, qui inhibe l'adénylate cyclase et diminue la concentration de l'adénosine monophosphate cyclique.

Le récepteur EP4 est couplé, comme le récepteur EP2, à une protéine G_s qui conduit à une augmentation de l'adénosine monophosphate cyclique.

Expression des récepteurs EP1 et EP2 à la prostaglandine E2 suivant les différents types cellulaires

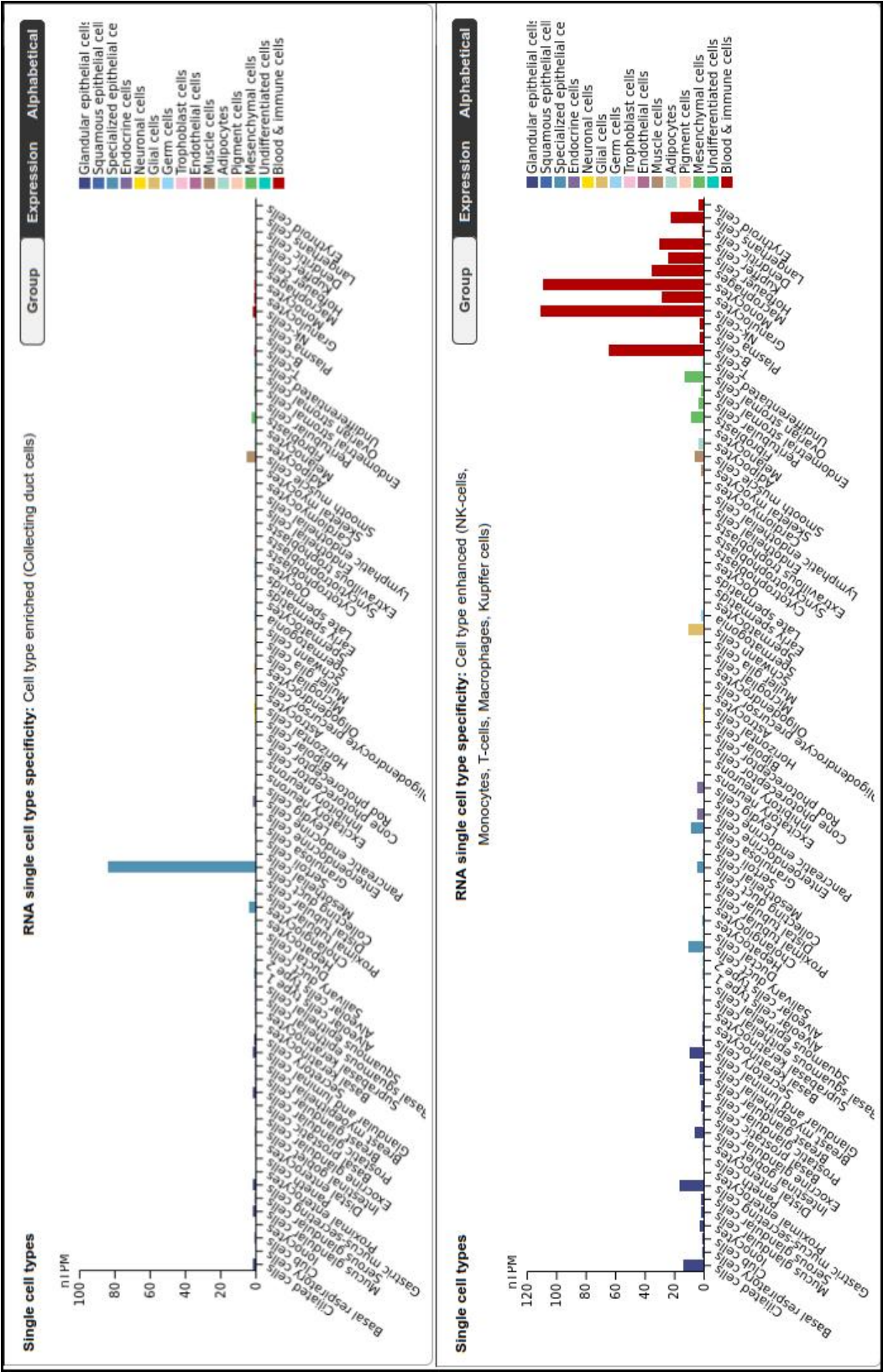


FIGURE 5 – Distribution et expression de PTGER1 et PTGER2, d'après "The Human Protein Atlas"/[pontenGlobalViewProtein2009/].

Expression des récepteurs EP3 et EP4 à la prostaglandine E2 suivant les différents types cellulaires

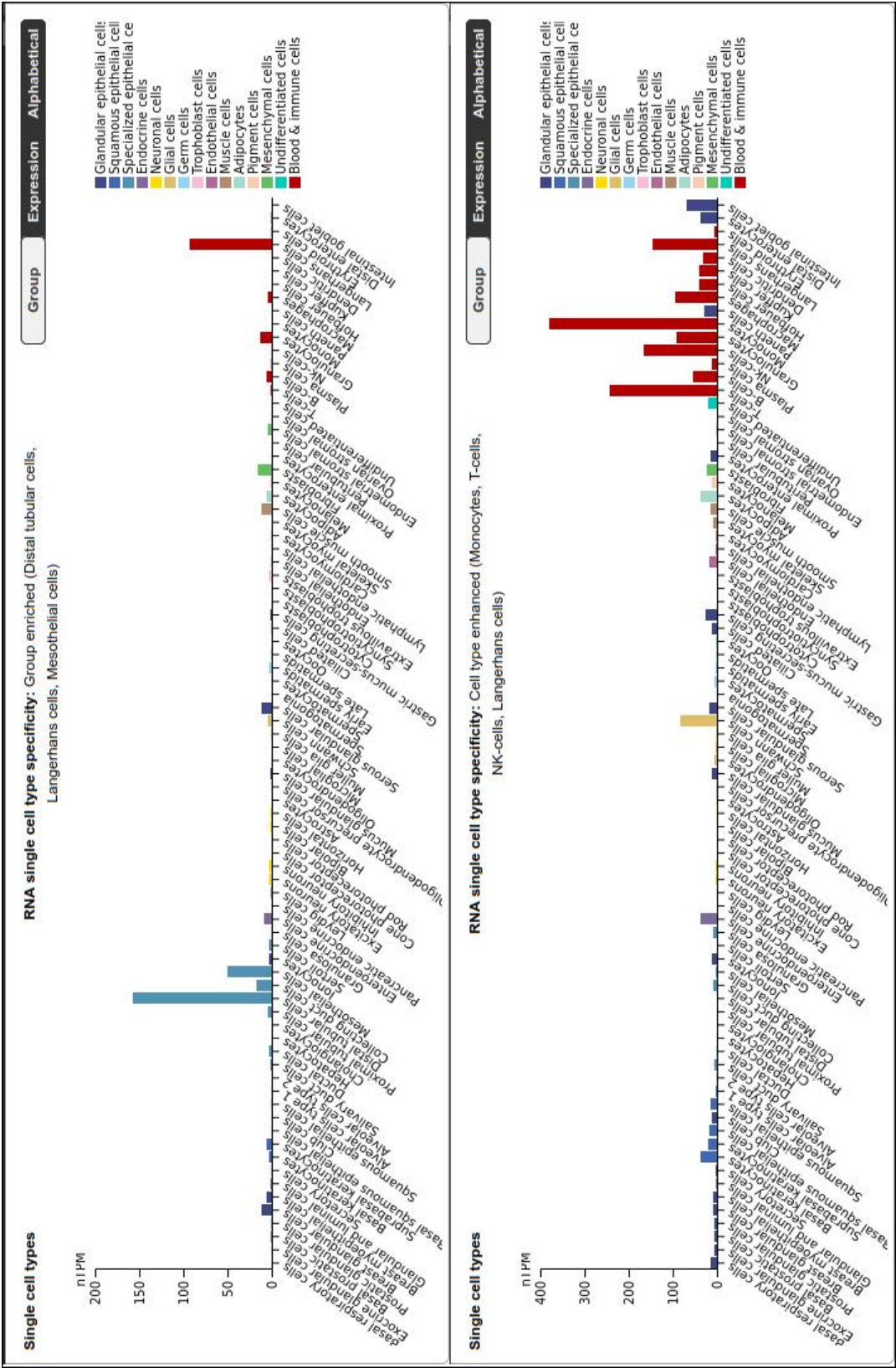


FIGURE 6 – Distribution et expression de PTGER3 et PTGER4, d'après "The Human Protein Atlas "[pontenGlobalViewProtein2009].

2.5 Les médicaments anti-inflammatoires

Plusieurs classes de médicaments anti-inflammatoires se distinguent suivant leur mode d'action. Une première classe, les AINS, visent à diminuer la synthèse de PG diminuant ainsi la réponse inflammatoire. L'autre classe, les anti-inflammatoires stéroïdiens (AIS), limite la réponse inflammatoire de l'organisme en agissant sur la transcription et la réduction de la réponse inflammatoire (notamment par la production de lipocortine qui inhibe la phospholipase A₂).

2.5.1 Les médicaments anti-inflammatoires stéroïdiens

Les médicaments anti-inflammatoires stéroïdiens sont la première grande classe des médicaments anti-inflammatoires. Les médicaments anti-inflammatoires stéroïdiens miment l'action de l'hormone glucocorticoïde, l'hydrocortisone issue de la réduction enzymatique de la cortisone. Plusieurs médicaments existent (figure 7), présentant comparativement à l'hydrocortisone une affinité moindre pour les récepteurs des minéralocorticoïdes (diminuant la réabsorption de sodium) et une demi-vie augmentée [95].

2.5.2 Mécanismes d'actions et effets indésirables

Les glucocorticoïdes exercent un spectre large d'activités, notamment en activant ou inhibant des facteurs de transcription [96]. Ils entraînent :

- la répression de la transcription des cytokines pro-inflammatoires (TNF α , IL1) interférant avec les facteurs de transcription NF- κ B et AP-1.
- la transcription de plusieurs gènes, contribuant à la production d'IL2. Promouvant ainsi l'action des lymphocytes T régulateurs.

Cette activité transcriptionnelle importante peut modifier la transcription estimée à environ 1% du génome humain [97], avec des différences suivant le tissu considéré et le glucocorticoïde employé [98].

Les glucocorticoïdes sont aussi associés avec un grand nombre d'effets secondaires. Ils entraînent une hyperglycémie et augmentent le risque de diabète ; ils augmentent aussi la prévalence de l'hypertension [99]. La prise sur le long terme est associée au risque d'athérosclérose et d'ostéoporose ainsi que la destruction des articulations [95]. Les effets indésirables d'apparitions à court et long terme sont également la psychose et les lésions oculaires [97].

Représentation de différents anti-inflammatoires stéroïdiens

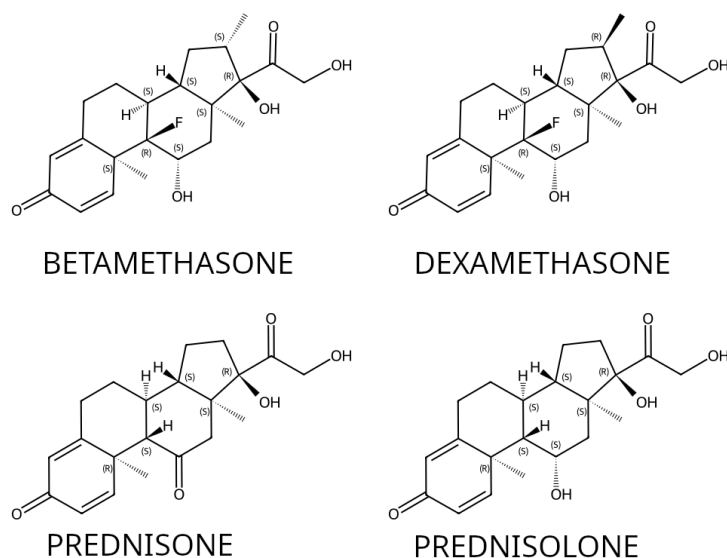


FIGURE 7 – Représentation de différents médicaments anti-inflammatoires stéroïdiens, dessiné avec RDKit.

2.5.3 Les anti-inflammatoires non stéroïdiens

Les AINS sont l'autre grande classe des actifs anti-inflammatoires. Ils visent à limiter la réponse inflammatoire par la diminution de la synthèse des PG pro-inflammatoires notamment les PGE2. Les AINS ciblent les PG synthases H2 aussi appelées COX situées en amont de la voie de synthèse des PGE2. Plusieurs isoformes des COX existent, les COX1 constitutives et les COX2 constitutives et inductibles, stimulées par l'inflammation. Les isoformes des COX présentent des structures cristallines très similaires, seul un acide aminé (AA) est différent entre COX1 et COX2 au niveau du site actif, l'isoleucine en position 523 est remplacée par la valine. L'homologie entre les deux isoformes est grande. Cette différence d'un AA a permis la mise au point d'AINS spécifiques de COX2. Les AINS se distinguent suivant qu'ils ciblent spécifiquement COX2 ou bien l'ensemble des COX. Au sein des AINS qui ciblent COX1 et COX2 plusieurs familles de médicaments se retrouvent suivant leurs structures chimiques, comme le montre la figure 8.

Représentation des anti-inflammatoires non stéroïdiens

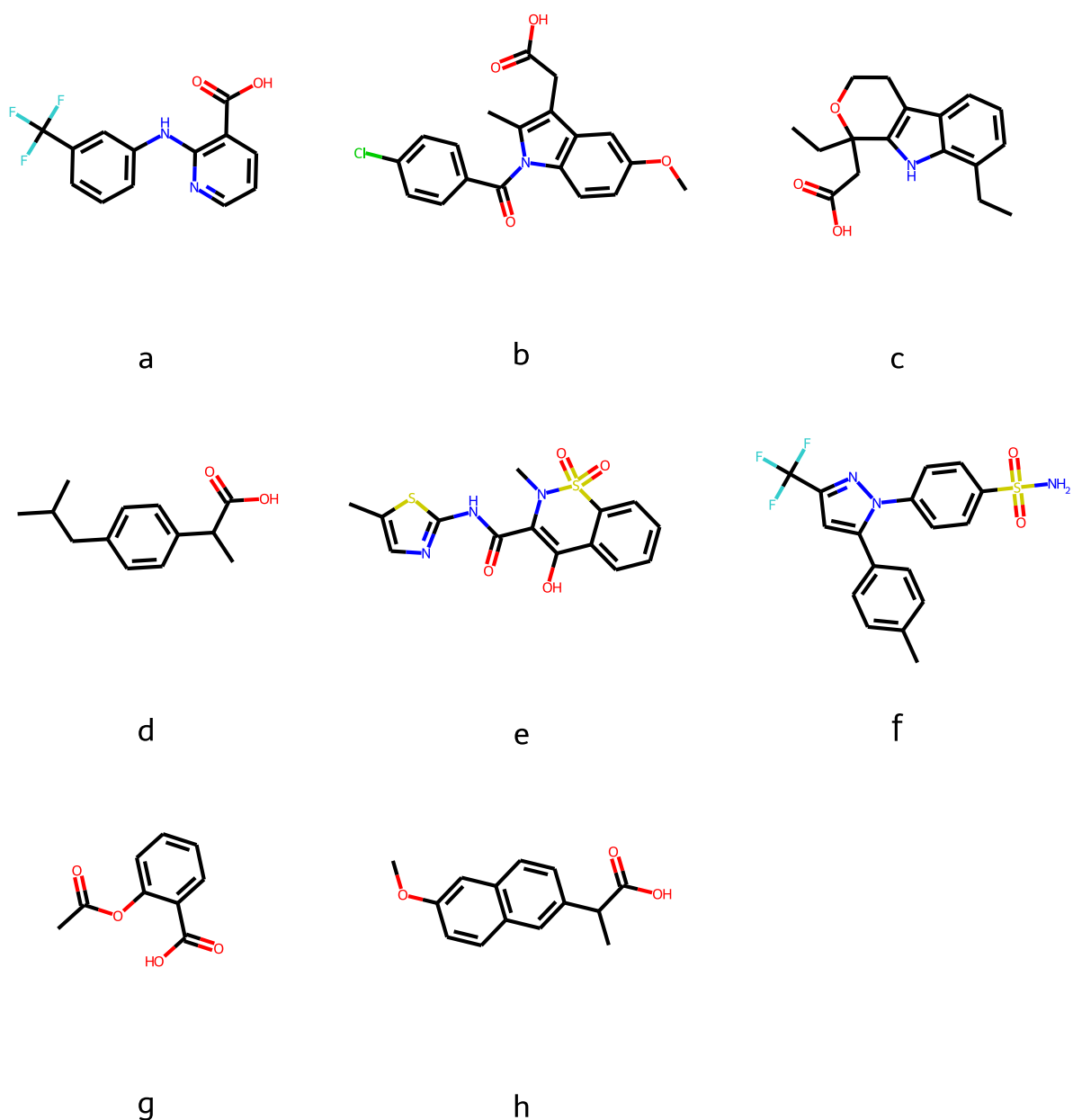


FIGURE 8 – Quelques AINS illustrant les différentes familles ; a) acide niflumique, b) indo-méthacine, c) étodolac, d) ibuprofène, e) méloxicam, f) célécoxib, g) acide acétylsalicylique, h) naproxène. Dessiné à l'aide de RDKit.

2.5.4 Les différentes familles d'AINS

Les AINS sont classés suivant leurs structures chimiques, leur fonction acide et leur caractère inhibiteur sélectif ou non de COX2.

- Les salicylés et les fénamates : ils présentent un cycle aromatique ou hétéro-aromatique, une fonction acide carboxylique portée par le cycle aromatique pour les salicylés. Les

fénamates (figure 2.5.4), dérivent de l'acide anthranilique (acide 2-aminobenzoïque).

Représentation des salicylés et fénamates

cale=2] ./salicyles_acide_mefenamique2_gimp.png

FIGURE 9 – Représentation de la structure des salicylés et de l'acide méfénamique (un fénamate), dessiné avec RDKit.

- Les dérivés acides arylacétiques : présentent un cycle aromatique ou un cycle hétéro-aromatique, une fonction acide carboxylique sur le carbone 2 de la chaîne à deux unités, comme l'illustre la figure 10.

Représentation des acides arylacétiques

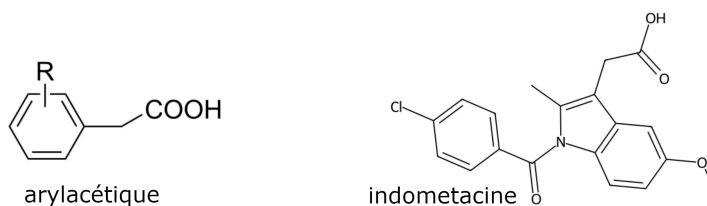


FIGURE 10 – Représentation de la structure structure générales des acides arylacétiques et de l'indométacine à droite comme exemple, dessiné avec RDKit.

- Les dérivés acides arylpropioniques : présentent un cycle aromatique ou hétéro-aromatique, porté par le carbone 2 en alpha de la fonction acide carboxylique comme le montre la figure 11.

Représentation des acides arylpropioniques

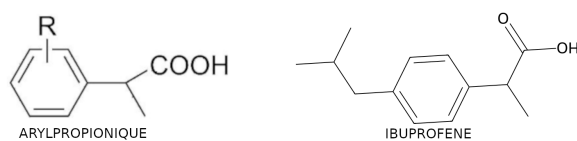


FIGURE 11 – Représentation de la structure générales des acides arylpropioniques et de l'ibuprofène à droite comme exemple, dessiné avec RDKit.

- Les oxicams : comportent un cycle à 6 avec une fonction type sulfamide en 1,2 et une fonction émol en 3,4 (figure 12).

Représentation des oxicams

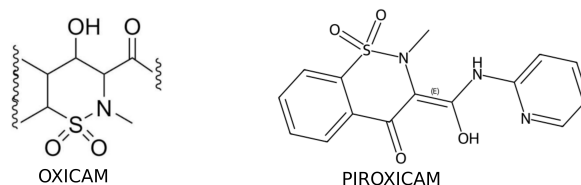


FIGURE 12 – Représentation du motif structural caractéristique des oxicams et du piroxicam en exemple à droite, dessiné avec RDKit.

- Les coxibs : COX2 spécifiques, ils présentent un cycle central substitué par deux cycles benzènes et par un groupement sulfone.

Représentation des coxibs

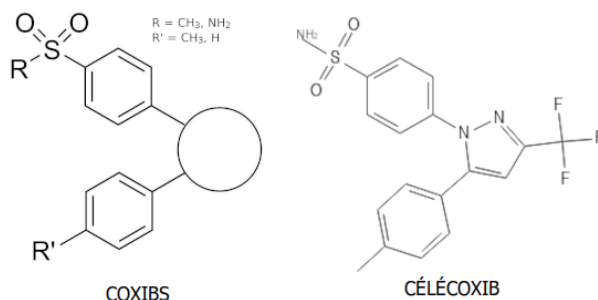


FIGURE 13 – Représentation du motif structural caractéristique des coxibs et du célécoxib en exemple de composé à droite, dessiné avec RDKit.

2.5.5 AINS et effets secondaires

Les COX sont présentes de manière ubiquitaire dans le corps et exercent de nombreuses fonctions physiologiques, en plus de leurs fonctions impliquées dans l'inflammation [100].

Les AINS, en produisant l'effet anti-inflammatoire, entraînent également plusieurs effets secondaires, suivant l'action spécifique ou non des AINS sur COX2, avec notamment des atteintes aux niveaux du système digestif, des reins, du foie et du système cardiovasculaire [101]

- Effets secondaires sur le système digestif : L'effet indésirable le plus connu et aussi le plus fréquent des AINS est l'atteinte du système gastro-intestinal. Diverses pathologies peuvent se retrouver, dont la sévérité se mesure à la dose et la durée de traitement par les AINS. Parmi celles-ci se trouvent les ulcères gastriques et intestinaux. De plus, le risque de complication gastro-intestinal est augmenté en cas de d'événements indésirables précédents [102].

]. Bien que les effets indésirables liés aux AINS soient bien connus, la cause est encore en discussion. Historiquement c'est l'inhibition de la production des PGI₂ et PGE₂, protectrices de la muqueuse gastrique qui était la cause des effets indésirables. Ceci a conduit au développement des inhibiteurs COX2 spécifiques. Cependant, il apparaît que COX2, joue aussi un rôle dans la protection de la barrière de la muqueuse gastrique, notamment dans la guérison de l'ulcère [103].

La plupart des AINS sont des acides faibles. Ils peuvent interagir avec le film hydrophobe, exposant ainsi la muqueuse ou se retrouvant en concentration importante par "piégeage d'ions" dans les cellules de l'épithélium [104].

Brièvement, les AINS sont sous forme non ionisée compte tenu du pH gastrique et peuvent diffuser librement à travers la membrane cellulaire, puis suite au changement de pH dans l'environnement cellulaire, se retrouvent sous forme ionisée, piégés dans la cellule. La prise chronique d'AINS est aussi très associée avec des complications au niveau entérique [105].

]. Dans ce cas, comme dans les complications gastriques, les AINS peuvent conduire à un stress oxydatif des mitochondries ou être impliqués dans une modification du microbiote [106].

- Effets secondaires sur le système rénal : L'impact des AINS peut inclure de multiples atteintes rénales : insuffisance rénale aiguë et chronique, glomérulonéphrite associée avec hypertension, hyponatrémie et hyperkaliémie [107, 108].

]. L'utilisation d' AINS au long court est associée à une augmentation importante du risque d' insuffisance rénale chronique [109].

]. Les AINS bloquent les COX, qui contrôlent la filtration glomérulaire et la rétention de sodium. De plus, en bloquant la production de PG ils bloquent aussi le rôle des PGE₂ et PGI₂ qui ont une fonction respectivement au niveau tubulaire et vasculaire. La PGE₂ prévient le transport du sodium au niveau de l'anse de Henlé et la PGI₂ induit une vasodilatation, augmentant la perfusion rénale, maintenant la filtration glomérulaire [107].

Par exemple, l'hyponatrémie associée aux AINS s'explique par l'inhibition de la régulation de l'hormone anti-diurétique par les PG, forçant le néphron à réabsorber l'eau.

- Effets secondaires sur le système cardiovasculaire : La consommation d'AINS est associée à une augmentation de la pression artérielle, des crises cardiaques et des ac-

cidents vasculaires cérébraux [110, 111].

- Effets secondaires hépatiques : 10 % des médicaments hépatotoxiques sont des AINS. Plusieurs ont été retirés du marché pour cette raison [112, 113].

]. Ce problème est un événement rare dans la prise d'AINS (1.4–9/100000 par personne-année [113] liée a une hypersensibilité ou un métabolite toxique causé par un polymorphisme génétique. Plusieurs mécanismes expliquent cette toxicité suivant l'AINS ciblé [114] Par exemple, l'antagonisme des AINS avec le récepteur farnésioïde X, impliqué dans la régénération et réparation du foie ainsi que dans la réponse inflammatoire, pourrait expliquer cette atteinte hépatique [115].]. Dans le cas du diclofénac, l'hépatotoxicité est causée par l'altération des mitochondries entraînant la production d'espèces réactives de l'oxygène [116].].

2.5.6 AINS et innovation

L'amélioration des AINS pour limiter ces effets indésirables reste un sujet de recherche actif.

Pour compenser l'effet indésirable gastro-intestinal, des prodrogues des AINS ont été développées [117, 118].

Ainsi, associer la prise d'AINS avec la libération de monoxyde d'azote sous forme de prodrogue permet de compenser les quantités limitées de pour la maintenance de la muqueuse gastrique [119, 120].

Plusieurs molécules ont été développées et associées à une diminution de la toxicité gastrique, selon plusieurs stratégies [121].

]. Par exemple en associant un AINS avec la libération de NO ou d'équivalent permettant de limiter les effets de l'inhibition des PG ou des AINS liés à des agents mucoprotecteurs [122].].

Cependant des effets mutagènes ont rapidement été mis en évidence et ont stoppé le développement clinique de certains de ces dérivés [121].

]. Diverses associations avec des AINS ont été envisagées :

- avec un anticholinergique pour bloquer les récepteurs muscarinique M1 et diminuer la sécrétion d'acide gastrique,
- avec un phosphate dialkyl pour limiter la formation de espèces réactives de l'oxygène et de l'azote (RONS) en inhibant la thiorédoxine.
- avec un donneur de H₂S pour prévenir la réduction du flux sanguin et favoriser la protection de la muqueuse gastrique. Certains composés développés sont à la fois donneurs de NO et de H₂S [117].].

2.5.7 Effets secondaires et variabilité individuelle

L'effet secondaire le plus fréquent est l'atteinte digestive par inhibition des COX1 constitutives et atteinte de la muqueuse gastrique. Cependant la toxicité des AINS est modulée par la mutabilité du Cytochrome P450 2C7, la vitesse du métabolisme est modifiée et, avec elle, les effets secondaires des AINS [123].]. Cela met en exergue la nécessité d'alternatives ou d'amélioration pour limiter les effets secondaires des AINS et faciliter le traitement anti-inflammatoire de long terme.

2.6 mPGES-1

Les effets indésirables des anti-inflammatoires montrent la nécessité de développer de nouvelles solutions thérapeuthiques et ainsi les PGE2 synthases sont intéressantes vu l'implication de PGE2 dans l'inflammation. Parmi celles-ci, mPGES-1 apparaît la plus intéressante car son expression est induite par l'inflammation [124].]. De plus elle intervient reès en aval dans la voie du métabolisme de , permettant ainsi de maintenir la production de médiateurs lipidiques important pour l'homéostasie.

2.6.1 mPGES-1 expression et régulation

L'expression de mPGES-1 est constitutive(Figure 14) dans de nombreux types cellulaires mais peut aussi être induite suite à un événement inflammatoire. Par exemple les travaux de Murakami *et al.* [124].] montrent l'absence de son expression dans le cœur et le foie humain dans des conditions normales mais a été observée après un infarctus ou une hépatite. La mPGES-1 est à la fois constitutive et inductible.

La mPGES-1 est codée par PTGES. Il est composé de trois exons, mesure 14,8 kb et est localisé sur le chromosome 9q34.3 [125].].

mPGES-1 présente plusieurs activités enzymatiques. Premièrement, elle est la plus efficace parmi les prostaglandines E2 synthases pour catalyser l'oxydoréduction de PGH2 en PGE2 avec un Km apparent de 310 mM⁻¹s⁻¹. mPGES-1 catalyse aussi la transformation des endocannabinoïdes en ester glycérolique de prostaglandines et la transformation de PGG2 en 15-hydroperoxy-PGE2. mPGES-1 présente aussi une petite activité enzymatique de peroxydase liée au GSH [126].].

La première projection structurale de mPGES-1 [126].] avec une résolution de 10Å, révèle sa forme trimérique (fig 15). D'autres études avec une meilleure résolution ont suivies, permettant de préciser la structure d'homotrimère de mPGES-1. Le monomère de 152 acides aminés présente 4 hélices transmembranaires. Dans le trimère, l'extrémité N terminale d'un monomère est proche de l'extrémité C terminale de son voisin [127].].

Aussi, mPGES-1 apparaît assez rigide, stabilisée par plusieurs liaisons hydrogènes interhélices [128, 129], contrairement à la plupart des protéines membranaires et des récepteurs couplés aux protéines G [130, 131].

Plusieurs cytokines (IL-1 β , TNF α) induisent l'expression de mPGES-1 [132-134], stamatakis Prostaglandins Induce Early 2015, par l'intermédiaire du facteur de transcription Egr-1. PTGES ne présente pas dans sa région promotrice de boîte TATA mais présente une boîte GC qui permet la liaison de Egr-1 [135-136].

2.6.2 La mPGES-1 et l'inflammation

Plusieurs éléments induisent l'expression de mPGES-1, notamment les cytokines IL-1 β et TNF α [137].

Ces travaux sur des rongeurs montrent une augmentation de la synthèse de mPGES-1 dans une situation inflammatoire ou de stress oxydant. Chez des souris mPGES-1 -/- comparative-ment avec des souris mPGES-1 +/+, la déplétion de mPGES-1 prévient au niveau du système vasculaire, la raideur vasculaire et l'altération de la structure de l'élastine; l'augmentation de l'expression de marqueurs pro-inflammatoires et des marqueurs de fibrose, l'augmentation de la vasoconstriction, l'augmentation de la NADPH oxydase, l'augmentation d'espèces oxygénées activées. De plus, ces effets sont aussi abolis lors de l'utilisation d'antagonistes des récepteurs EP1 et EP3 ainsi que par l'inhibition de JNK et ERK1/2.

2.6.3 Le site actif

Le site actif de mPGES-1 se situe à l'interface de deux sous-unités, proche du GSH (figure 20). Plusieurs études ont permis de montrer les éléments importants pour l'interaction de l'enzyme avec le substrat endogène (PGH₂) ou avec les inhibiteurs de mPGES-1.

T.Hammarberg *et al.* ont construit un modèle par homologie basé sur l'glsPGA leucotriène C₄ synthase humaine et ont montré que l'Arg126 est essentielle à l'activité de mPGES-1. Les études de mutagenèse montrent une diminution importante de l'activité ou un changement d'activité enzymatique en cas de mutation de ce résidu [138].

T.Sjögren *et al.* par la détermination de la structure cristallographique de mPGES-1 avec une meilleure résolution (1,2 Å), présentent le site actif comme une cavité de 15 Å, avec une ouverture de 12 par 9 Å. Le GSH établit des liaisons hydrogènes avec les chaînes latérales de Arg73, Asn74, Glu77, His113, Tyr117, Arg126, Ser127 et de l'Arg38. En plus des liaisons hydrogènes, le groupement phénol de la Tyr130 est aussi impliqué dans une interaction de π stacking avec le GSH [128]. Ces travaux montrent aussi que l'interaction entre l'Arg126 et l'Asp49 est importante pour la réaction de catalyse dans les différents modèles putatifs de réaction enzymatique.

Plusieurs études visent à étudier le site actif de mPGES-1 mais toutes ne s'accordent pas sur le mécanisme d'action. Sans se vouloir exhaustif les similitudes et différences sont relevées ci-après.

S. Di Micco *et al.* à travers une expérimentation de dynamique moléculaire corrobore les résultats de T. Sjögren. Dans ces travaux, les acides aminés suivants contribuent de manière significative à la liaison du ligand : Ala31, Thr34, Arg38, Arg73, Asn74, His113, Tyr117, Arg126, Ser127 et la Tyr130. Cela est confirmé également par l'évaluation de l'énergie de liaisons après mutation de l'acide aminé en alanine. L'Arg38 et Arg73 établissent une liaison ionique et deux liaisons hydrogènes avec le GSH. Asn74, His113, Tyr117, Arg126 et Ser127 établissent une liaison hydrogène avec le GSH. Ala31, Thr34 et Tyr130 sont impliqués dans des interactions de Van der Waals avec le GSH. De plus, cette étude se focalise sur Arg73 qui présente 2 conformations distinctes (4AL0), et montre que Arg73 pourrait être en interaction avec une molécule d'eau impliquée dans la biosynthèse de PGE2 [129].

].

Expression de mPGES-1 dans "The Human Protein Atlas"

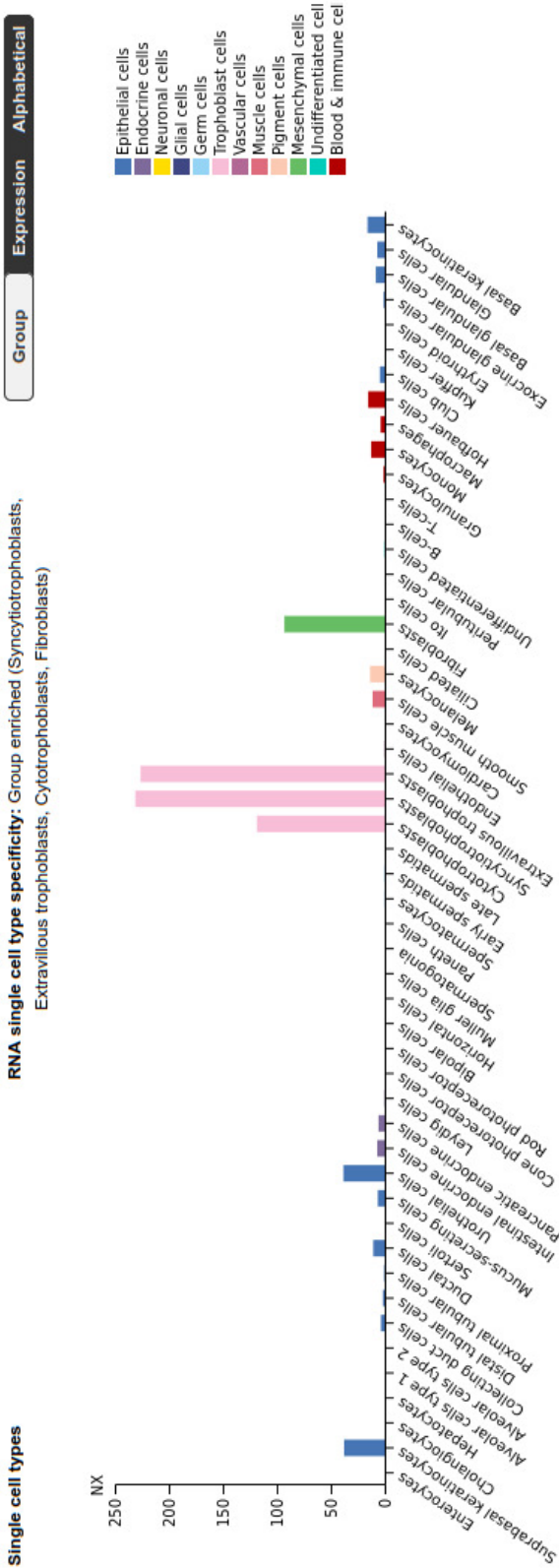


FIGURE 14 – Expression de mPGES-1 en fonction de différents types cellulaires, d'après "The Human Protein Atlas" /pontenGlobalViewProtein2009/.

Structure homotrimérique de mPGES-1

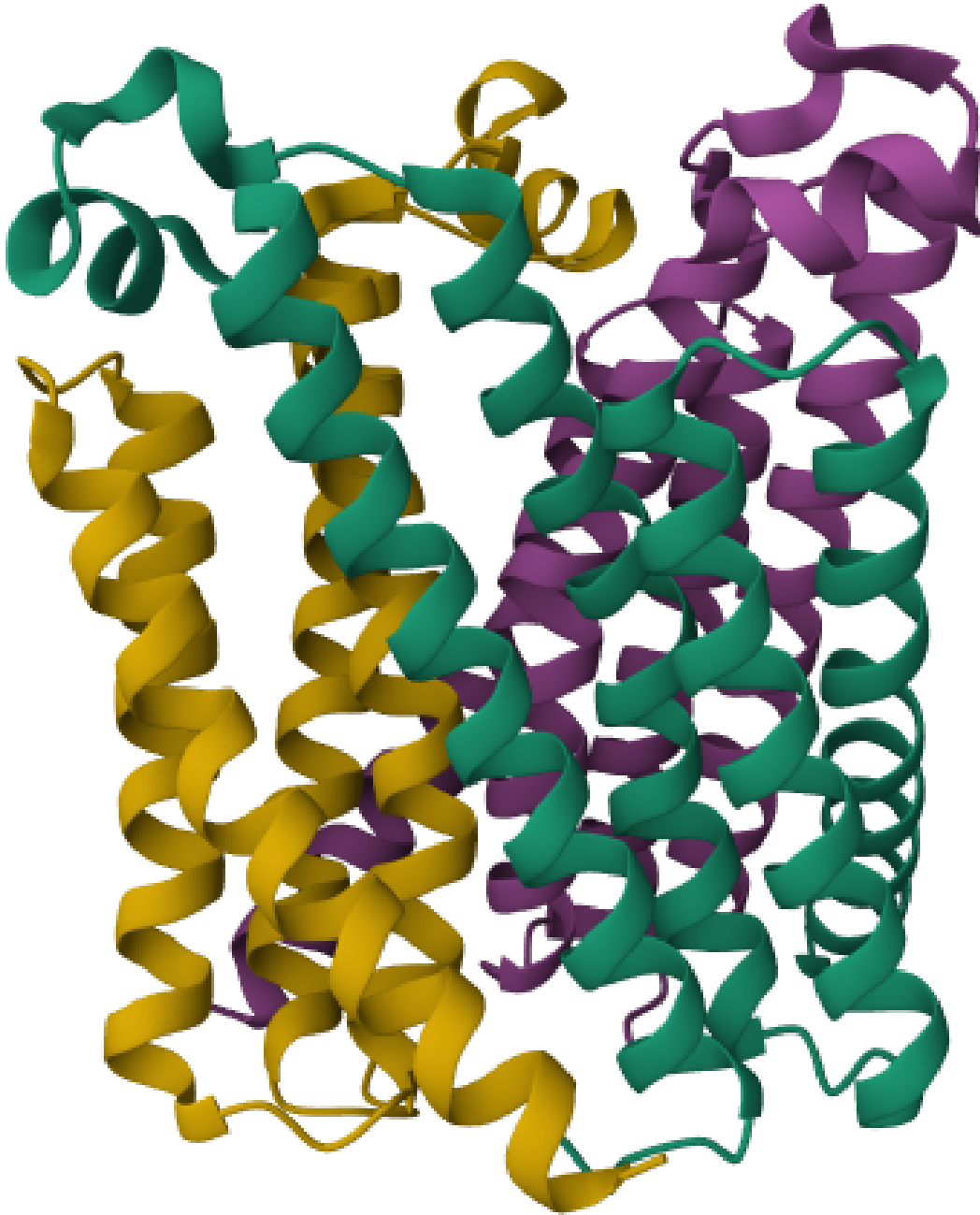


FIGURE 15 – *Structure de mPGES-1 issue de la PDB (4YL0), colorisée par homotrimère.*

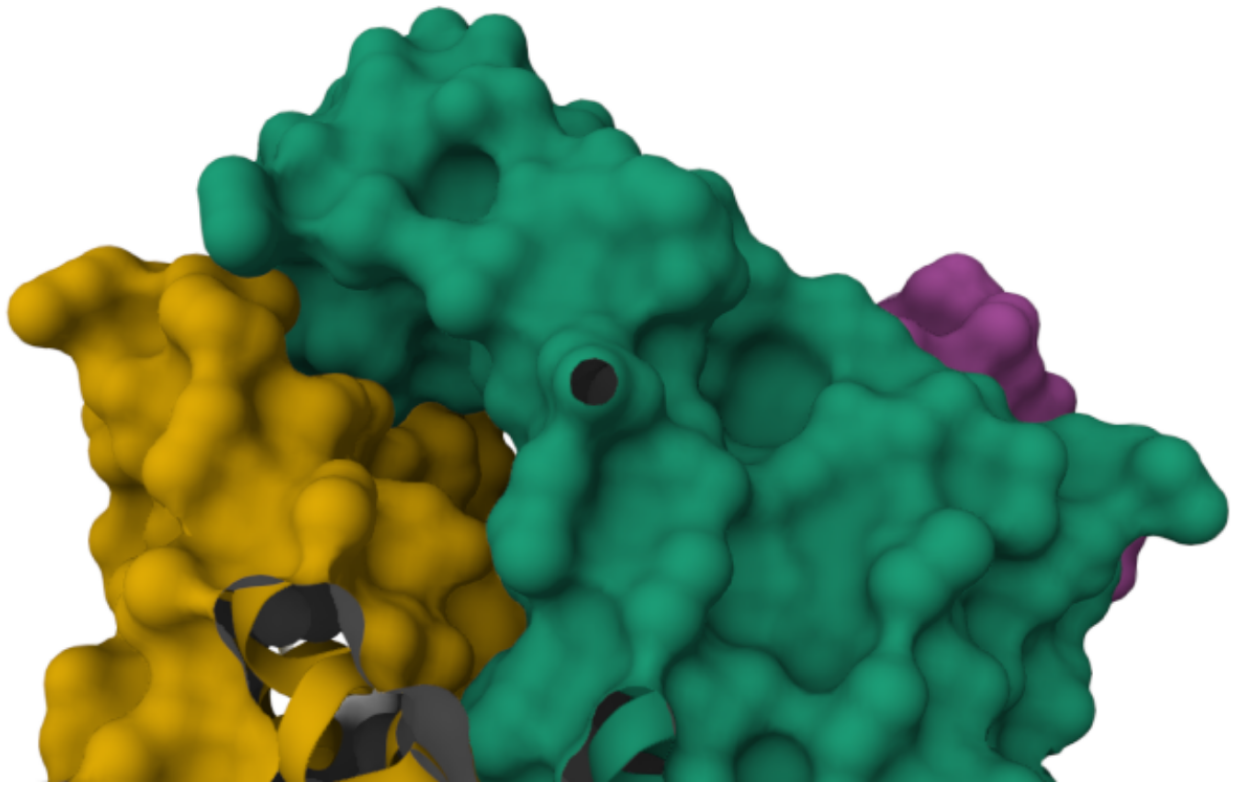


FIGURE 16 – *Surface de mPGES-1 issue de la PDB, colorisée par homotrimère.*

J. Brock *et al.* implique aussi le solvant dans la réaction catalytique, notamment par l'interaction entre Asp49 et Asp126, mais ces travaux remettent en cause l'interaction de la Ser127 avec le GSH [139].

mPGES-1 présente un site actif à l'interface entre deux homotrimères et donc trois sites potentiels par protéine, cependant chacun d'eux présente une affinité différente avec le GSH. Il en résulte qu'un seul site présente une activité catalytique comme le montre les travaux de J. Alexander *et al.* [140].

De plus, mPGES-1 présente plusieurs conformations. En présence de GSH, la conformation fermée est prédominante. Cependant, la barrière énergétique entre les deux conformations est assez faible et la conformation ouverte présente un niveau énergétique relativement plus faible que la conformation fermée. En l'absence de GSH, la co-existence des deux conformations est probable. Ces travaux sur les conformations de mPGES-1 montrent que les principaux changements d'interactions des acides aminés concernent notamment l'Asp49, qui dans la conformation ouverte n'est plus en interaction avec l'Asp126 mais forme une liaison hydrogène avec l'Asn46. L'ensemble des changements de conformations entre l'Asn46 et l'Asn66 semble jouer le rôle de porte d'entrée pour le GSH entre les différentes conformations. Ce rôle est invalidé pour l'Arg73 dans l'étude de S. Di Micco *et al.* [141].

mPGES-1 présente également des différences inter-espèces ce qui complique le développement d'inhibiteurs de mPGES-1, notamment les modèles murins ne sont pas adaptés pour valider *in vivo* les meilleurs candidats identifiés *in vitro* sur enzyme humaine.

2.7 Les structures de complexes mPGES-1-ligand disponibles

Plusieurs structures de mPGES-1 sont disponibles au sein de la PDB et certaines cristallisées en complexe avec un ligand. Ce sont ces structures qui ont été utilisées pour la réalisation des pharmacophores.

Ces complexes ont une résolution de qualité suffisante, entre 1,56Å à 1,20Å. En plus de ce facteur clé, d'autres facteurs mesurent la qualité de la protéine cristallisée comme cela peut se voir par l'exemple donné par la figure 17 qui illustre les mesures de validation du complexe 4YL0 au sein de la PDB

Exemple des mesures de validation de 4YL0

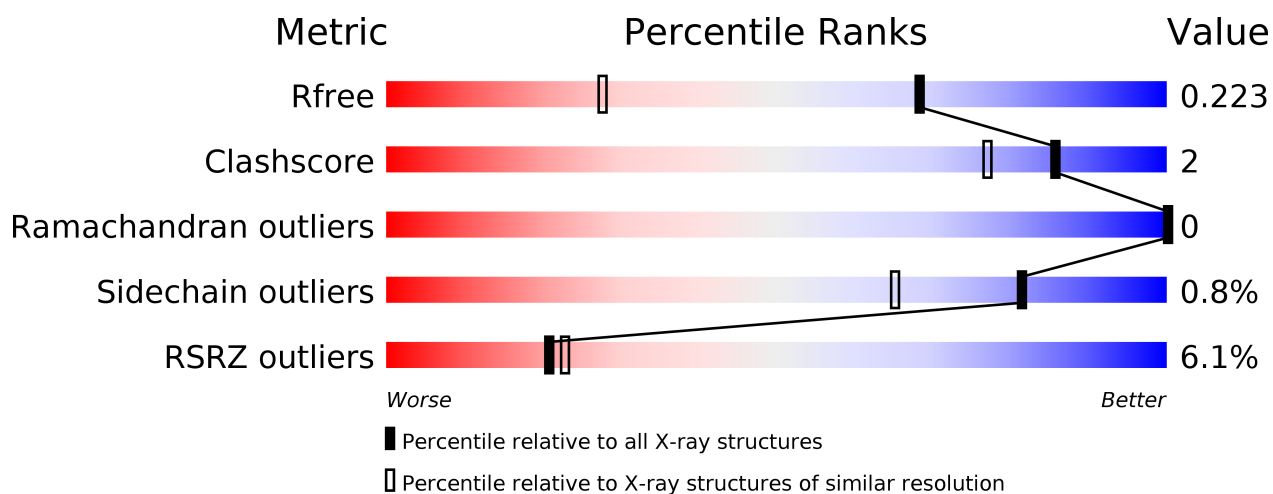


FIGURE 17 – Exemple des mesures de validation de 4YL0 issu de la structure cristallisée de mPGES-1 déposée au sein de la PDB.

Deux modes de liaison sont possibles au sein du site actif. Un premier identique à la pose du ligand naturel et un deuxième plus excentré, tout en restant proche du glutathion comme l'illustre la figure 20 présentant différents complexes.

2.8 les différents inhibiteurs de mPGES-1

De nombreux composés naturels ou de synthèses ciblent mPGES-1, et plus d'un millier de composés ont été rapportés depuis la découverte de la structure de mPGES-1 en 1998 [83]. Différents inhibiteurs de mPGES-1 ont atteint différents stades de développement mais sans aboutir pour l'instant à la mise sur le marché.

Exemples de complexes cristallographiques inhibiteurs-mPGES-1

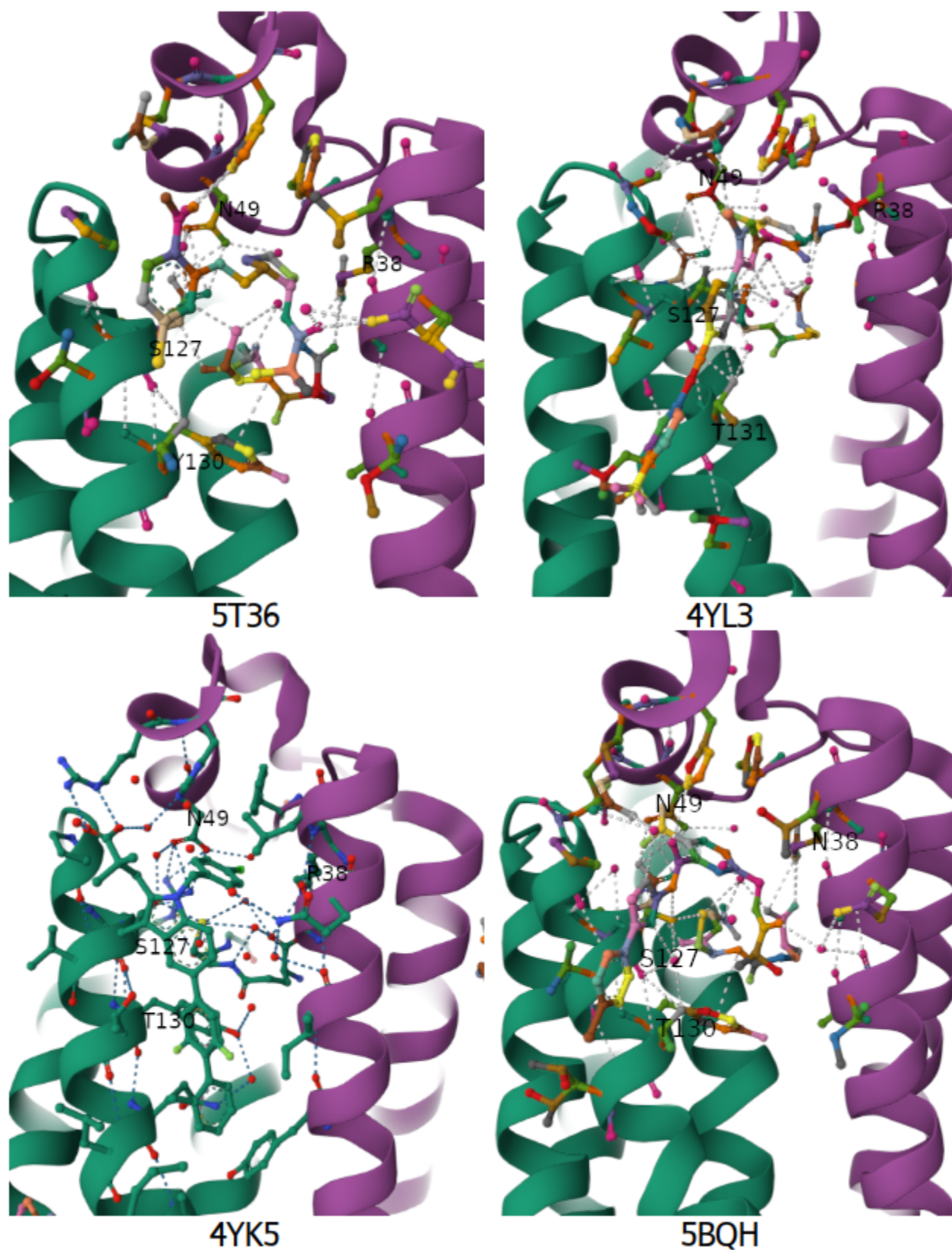


FIGURE 18 – Exemples de complexes inhibiteurs-mPGES-1 représentant la diversité des modes de liaison au sein du site actif de mPGES-1.

Plusieurs inhibiteurs de mPGES-1 sont en cours d'essai clinique.

Inhibiteurs de mPGES-1, ayant atteint le stade d'essai clinique

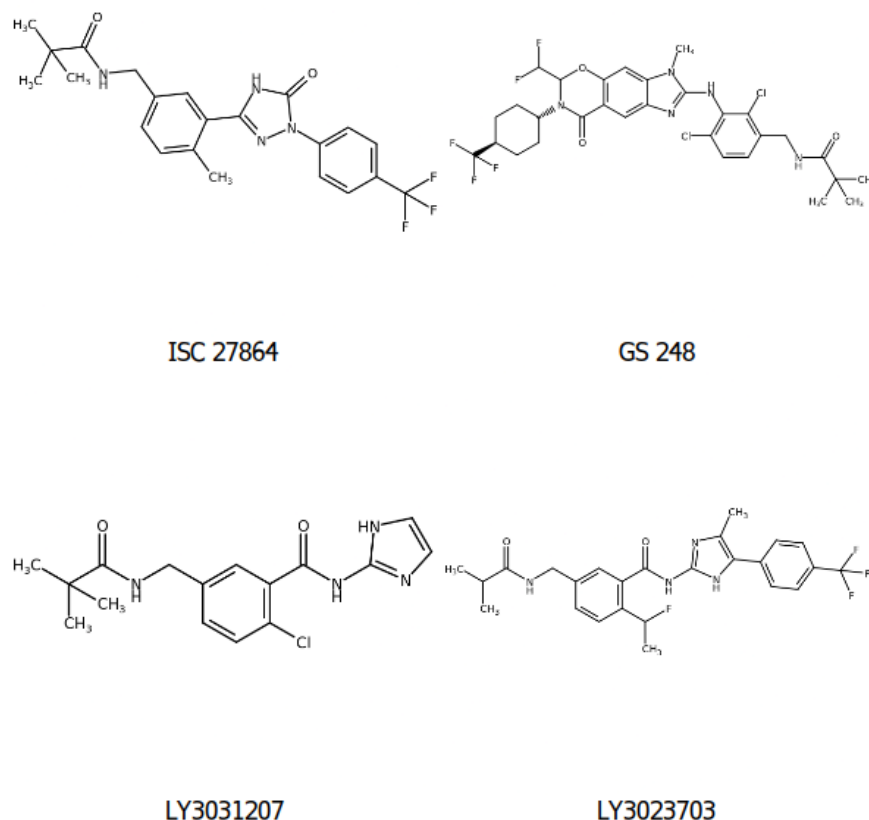


FIGURE 19 – Exemples d'inhibiteurs de mPGES-1 ayant atteint différents stades d'essais cliniques.

- ISC 27864 aussi appelé Zaloglanstat montre une diminution de la production de PGE₂, de manière dose dépendante dans une étude de phase 1 comparée avec le célécoxib. Ces travaux ont été complétés par un essai de phase 2 en double aveugle randomisé, en Inde au cours de l'année 2020 [142].
Cependant il n'apparaît plus dans les pipelines de R&D de Glenmark Pharmaceuticals Ltd qui en détient la licence.
- Vipoglanstat est un inhibiteur de mPGES-1 en cours d'essai de clinique de phase 2 pour la prise en charge du syndrome de Raynaud. Il est associé à une baisse dose dépendante de la production de PGE₂ et une augmentation de PGI₂ [143].
- LY3031207 et LY3023703 [144, 145] sont des inhibiteurs qui ont été arrêtés au cours d'essai de phase 1 à cause de métabolites réactifs hépatotoxiques [146].

Hormis ces composés ayant atteint le stade de l'essai clinique, beaucoup de dérivés synthétiques ou naturels sont capables d'interagir fortement avec mPGES-1.

Plusieurs produits de la cascade enzymatique de l'ARA se comportent comme des inhibiteurs endogènes de mPGES-1. Ceux sont le leucotriène C4 (LTC₄) et le 15-deoxy- $\Delta^{13,14}$ PGJ₂ qui présentent une meilleure affinité que le leucotriène. Cette interaction peut s'expliquer par l'homologie structurale des membres de la superfamille MAPEG (FLAP,LTC₄S,MGST1,MG

D'autres composés d'origine naturelle inhibent mPGES-1. La plupart sont des polyphénols, mais on trouve aussi des composés terpéniques.

Tous ne sont pas des inhibiteurs directs de mPGES-1. Par exemple la curcumine issue du rhizome séché de *Curcuma longa* L. supprime l'expression de Egr-1, facteur de transcription de mPGES-1 [147].

Un exemple parmi les polyphénols est le gallate d'épigallocatechine [148].

La myrtucommulone A, une acylphoroglucinol non prénylé est un exemple issue des lactones [149]. Elle inhibe mPGES-1 sans

inhiber les COX et inhibe aussi potentiellement la synthèse des LT. L'hyperforine, au contraire, un acylphoroglucynol polyprénylé, extrait de *Hypericum perforatum* L. [150].

présente une CI₅₀ de l'ordre du micromolaire mais présente un profil d'activités similaires en inhibant mPGES-1 et la synthèse des LT sans inhiber les COX.

Inhibiteurs de mPGES-1 d'origine naturelle

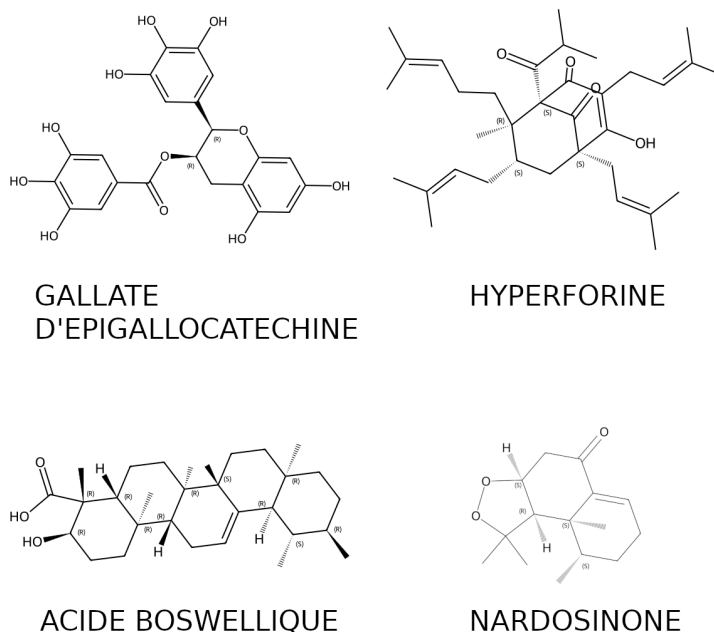


FIGURE 20 – Exemples Inhibiteurs de mPGES-1 d'origine naturelle. Dessiné avec RDkit.

En plus des polyphénols, des composés terpéniques sont aussi inhibiteurs de mPGES-1.

L'acide boswellique [151].] issu de *Boswellia serrata* R. et la nordanisone N [152].] issue de *Nardostachys chinensis* B. sont des composés terpéniques qui ont des CI_{50} du même ordre que les composés précédents.

Le travail de recherche d'inhibiteurs de mPGES-1 a contribué à la découverte de plusieurs familles de composés :

1. Arylpyrrolizines : Plusieurs composés sont dérivés de la licofélone (figur e21) ,u ninhibi
teur double des COX et de la 5-LOX[153].]

Les modifications structurales ont visé notamment la fonction acide carboxylique pour diminuer l'interaction avec les COX et augmenter celle avec mPGES-1[154].

Notamment le remplacement par un acide propanoïque augmente la stabilité du composé.

Représentation de la licofélone

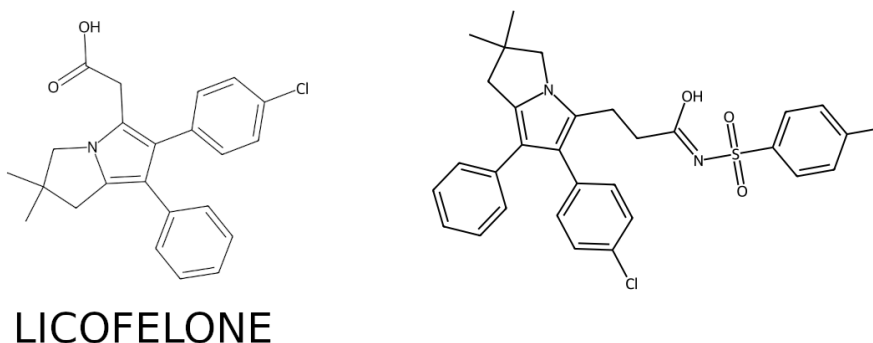


FIGURE 21 – *Licofélone, structure chef de file des arylpyrrolizines. Dessiné avec RDkit.*

2. Biarylimidazoles : Ils dérivent de MF-66, le groupe imidazole est disubstitué en position 2 et 4, soit R1 et R2 respectivement. R1 peut être substitué avec un groupe hydrophobe à travers un "linker" rigide maximise l'inhibition biologique ; R2 est substitué par un groupement électronégatif [155]] comme le montre la figure 22.

Représentation des biarylimidazoles

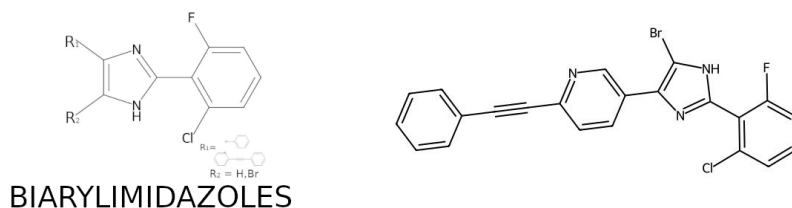


FIGURE 22 – *noyau des biarylimidazoles, avec un exemple de composé où R1 et R2, varient impactant la CI₅₀. Dessiné avec RDkit.*

3. MK-886 : Mk-886(figur e23)es tu ncompos écomprenan tu ngroup eindol ee tu acide carboxylique, inhibiteur de FLAP (protéine activatrice de la 5-LO) inhibant aussi mPGES-1. Plusieurs inhibiteurs sélectifs de la mPGES-1 ont été développés [156]. Ces composés fortement lipophiles ont montré une liaison forte aux protéines plasmatiques dans le sang total, supprimant l'activité, bien que la CI₅₀ est de l'ordre du nanomolaire pour ces inhibiteurs dérivés de MK-886.

Représentation de MK-886

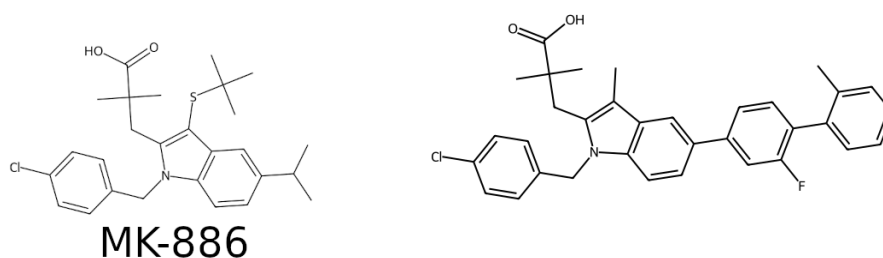


FIGURE 23 – *mk-886, structure "lead" des dérivés de MK-886. Dessiné avec RDkit.*

4. Benzoxazole : Le noyau benzoxazole, identifié au cours d'un criblage à haut débit, a conduit à l'identification d'un inhibiteur, aboutissant à un benzoxazole piperidinecarboxamides avec une CI₅₀ de l'ordre du nanomolaire [157].]

Représentation des Benzoxazoles

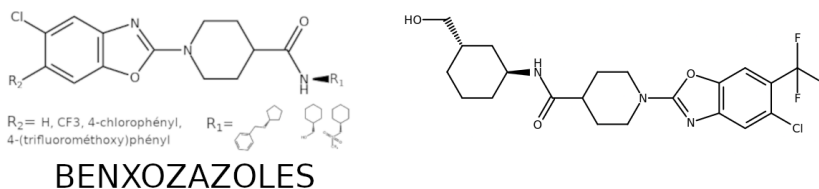


FIGURE 24 – Noyau des benzoxazoles, avec le groupement R qui varie, modifiant l'activité inhibitrice. Dessiné avec RDkit.

5. Phénanthrène imidazole : Ils dérivent de MF-63, inhibiteur sélectif de mPGES-1 [158]. MF63 présente une CI_{50} meilleure que MK-866. Les positions 6 et 9 du phénanthrène imidazole sont critiques pour l'activité.

Représentation des Phenanthrenes imidazoles

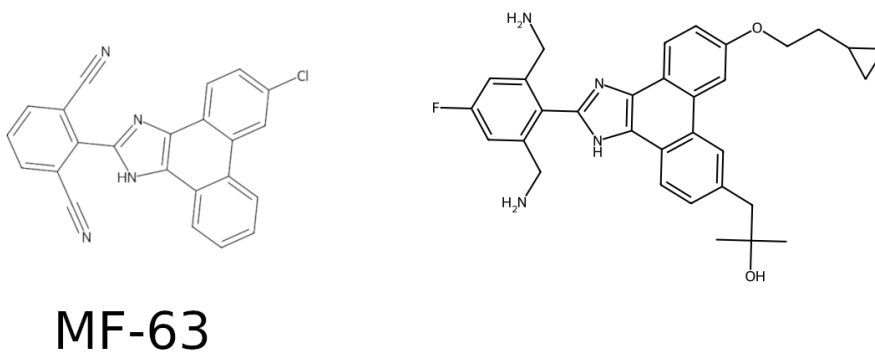


FIGURE 25 – Noyau des phenanthrenes imidazoles, avec les groupements $R1, R2$ et $R3$ qui varient, modifiant la CI_{50} des composés respectifs. Dessiné avec RDkit.

6. Imidazoquinolines : Le vhef de file de cette structure (26) est issu d'un criblage haut débit [159]. Il montre une inhibition modérée de mPGES-1. L'optimisation de cette structure conduit à un inhibiteur spécifique de mPGES-1 comparé aux COX [160], présentant des résultats ADME in vitro favorables [161].

Représentation des Imidazoquinolines

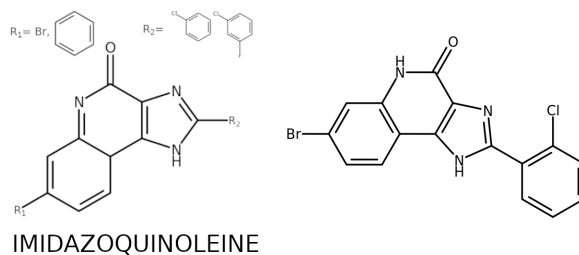


FIGURE 26 – Noyau des imidazoquinolines, avec le groupement R_1 qui varie, modifiant la CI_{50} des composés respectifs. Dessiné avec *RDkit*.

7. Oxicams : Les Oxicams sont une famille AINS, ils inhibent donc les COX. À partir de ce noyau oxicam, un composé présentant une sélectivité plus de 200 fois plus importante pour mPGES-1 que pour COX2, avec une CI_{50} de $0.42 \mu\text{M}$ [162].

Représentation des Oxicams

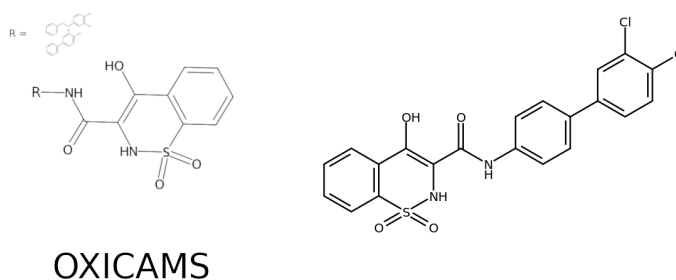


FIGURE 27 – Noyau des oxicams avec le groupement R qui varie, impactant la CI_{50} . Dessiné avec *RDkit*.

8. Urées trisubstituées : Merck Frosst, à travers une campagne de criblage haut débit, identifie une nouvelle classe d'inhibiteurs de mPGES-1, les urées trisubstituées [163]. L'optimisation a consisté à rallonger les substituants par un lien rigide, un alkyle ; permettant d'atteindre une CI_{50} de 1 nanomolaire. Cependant lors d'essais sur sang total humain, la forte liaison aux protéines plasmatiques empêche l'inhibition de l'activité de mPGES-1.

Représentation des urées tri-substituées

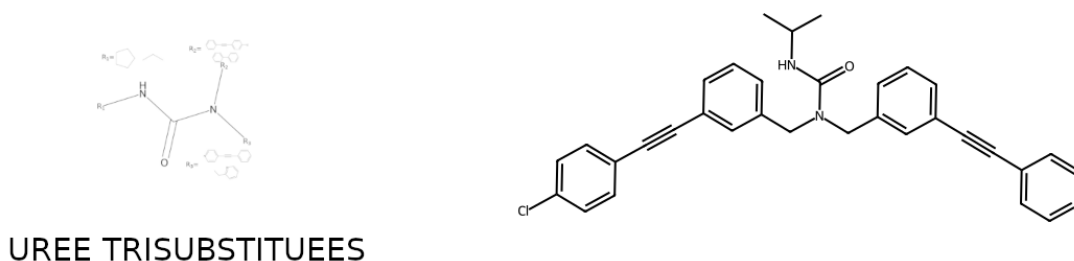


FIGURE 28 – Noyau des urées tri-substituées avec le groupement $R1, R2$ et $R3$ qui varient, entraînant une modification de la CI_{50} ; avec un exemple de composés à droite Dessiné avec *RDkit*.

9. Acide pirinixique : L'acide pirinixique est le chef de file de cette série. Conduisant au double inhibiteur de la 5-lipooxygénase (5-LOX) et de mPGES-1 [164].

C'est la substitution par un groupement biphényle qui donne l'inhibiteur dual avec la plus petite CI_{50} . La substitution de ce groupement par un α -n-hexyl conduit à YS121 [165].] qui présente une CI_{50} semblable sur sang total humain à celle des tests invitro acellulaires et limite l'infiltration de leucocytes dans un modèle inflammatoire du rat.

Représentation des dérivés de l'acide pirinixique

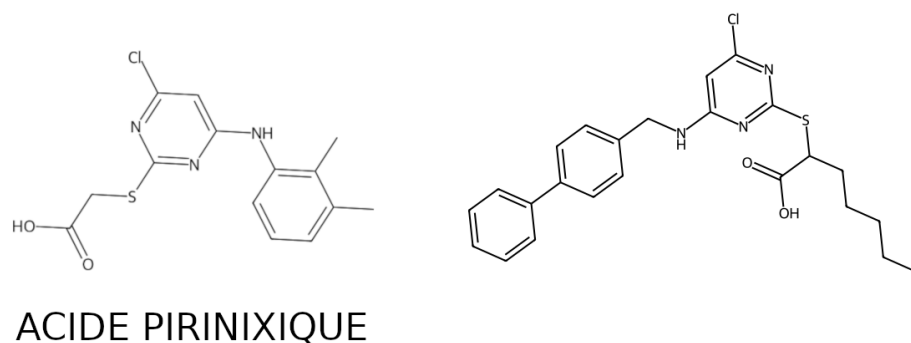


FIGURE 29 – Noyau de l'acide pirinixique avec les groupements $R1$ et $R2$ qui changent, impactant la CI_{50} ; et un exemple de molécule à droite. Dessiné avec RDkit.

10. Acide 2-mercaptohexanoïque : D'abord identifiés comme inhibiteurs des PPAR α et γ , ils inhibent dualement mPGES-1 et la 5-LOX, mimant la structure de l'acide arachidonique en allongeant l'acide 2-mercaptohexanoïque par un fragment aryloxy [166].

Représentation des dérivés de l'acide 2-mercaptohexanoïque

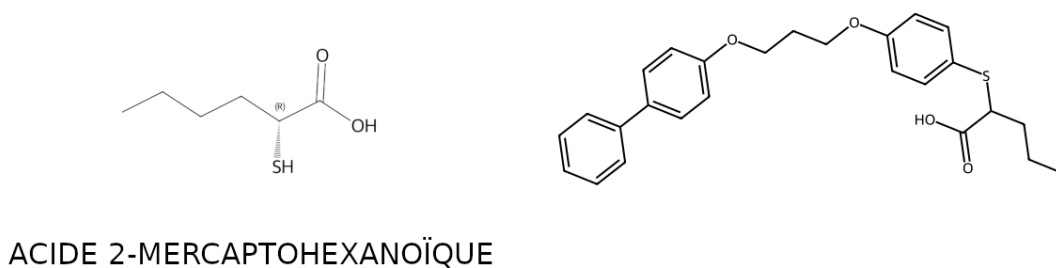


FIGURE 30 – Noyau de l'acide 2-mercaptohexanoïque avec un représentant de cette famille d'inhibiteurs. Dessiné avec RDkit.

11. Carbazole benzamide : Parmi une série de carbazoles benzamides, un inhibiteur potentiel a été identifié par Bruno *et al.*, l' AF3442 [167].

Il inhibe la production de PG E2, sans affecter les COX ,de manière dose dépendant dans un modèle inflammatoire de monocytes induit par le lipopolysaccharide.

Représentation du carbazole benzamide

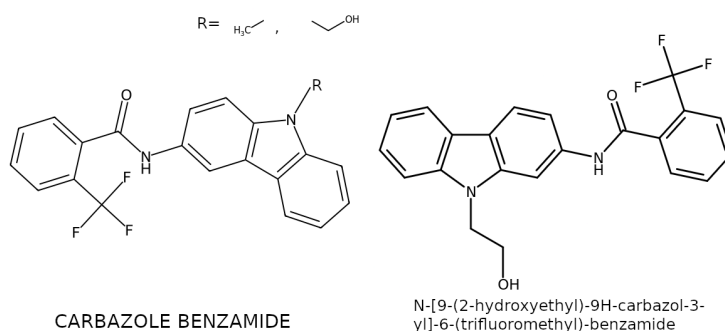


FIGURE 31 – *Inhibiteur de mPGES-1 représentant des carbazoles benzamides. Dessiné avec RDkit.*

3 Méthodes et résultats

3.1 Chémo-informatique

3.1.1 Stratégie de criblage

La découverte et le développement de médicaments est un processus coûteux et long [168].

Une approche classique consiste en la synthèse et le criblage d'un grand nombre de composés pour identifier de potentiels candidats. Depuis de nombreuses années, une approche computationnelle est de plus en plus utilisée pour limiter les coûts de découverte de médicaments, en rationalisant les travaux de découverte et mise au point de candidats médicaments. Dans cette approche de conception assistée par ordinateur, plusieurs stratégies existent. Elles se distinguent suivant leurs méthodes, centrées sur le ligand ou basées sur la structure d'une cible.

"Structure based drug design" (SBDD) : fait référence à l'utilisation de données structurales obtenues soit par l'expérimentation (cristallographie à rayon X, spectroscopie par résonance magnétique nucléaire et cryo-microscopie électronique) soit par modélisation informatique, par homologie dans le but de concevoir des ligands de haute affinité en tenant compte des caractéristiques chimiques et structurales de la cible.

Parmi les différentes stratégies de SBDD, les plus employées sont : le criblage virtuel basé sur la structure, l'amarrage moléculaire et la dynamique moléculaire.

Le criblage virtuel, similairement au criblage à haut débit, vise à déterminer les molécules actives à partir d'une librairie de composés sans détermination expérimentale de l'activité de l'ensemble de la librairie [169].

La dynamique moléculaire prédit l'évolution dans le temps de l'ensemble des atomes d'un système moléculaire (composants, interaction et environnement). Elle permet de prédire la structure 3D d'une molécule ou d'étudier l'effet d'une mutation et, principalement, d'étudier la dynamique d'un processus et ainsi estimer de façon plus juste l'affinité d'un ligand avec sa cible [170].

L'amarrage moléculaire peut être rigide, semi-rigide ou flexible. Il repose sur l'appariement spatiale et énergétique entre un ligand et une cible. Plusieurs algorithmes et fonctions de score existent. Ils sont les paramètres importants de cette technique.

"ligand based drug design" (LBDD) fait référence à l'utilisation de l'information issue d'un ensemble de ligands préexistants connus, que les ligands soient des inhibiteurs actifs ou au contraire ne présentant pas d'activité importante contre la cible.

Trois stratégies différentes de LBDD existent. La relation quantitative structure à activité cherche à établir un modèle probabiliste qui corrèle l'activité biologique aux variations structurales d'un ensemble de molécules [171].

La modélisation de pharmacophores vise à déterminer les éléments essentiels de la liaison

avec la cible (décrit plus explicitement à la section modélisation des pharmacophores (3.2)). La recherche de similarités s'appuie sur un grand nombre de descripteurs issus d'un ensemble de ligands pour identifier des molécules similaires susceptibles de présenter aussi une activité biologique contre la cible.

LBDD et SBDD présentent chacun des avantages et inconvénients résumés dans le tableau ci-après [172].

Techniques	Avantages	Inconvénients
Pharmacophore basé sur le ligand "ligand-based"	Possible sans informations structurales sur les protéines	Structure des protéines non prises en compte, ensembles d'entraînements difficiles à construire si peu de ligands connus, surajustement (capture de la spécificités des données plutôt que la tendance générale) possible si jeu de données trop peu diversifié
Pharmacophore basé sur la structure "structure-based"	Possible sans information sur le ligand, prise en compte de toute la capacité de la poche protéique, prédiction des modes de liaison	Sélection de fonctionnalités essentielles plus difficiles si absence de ligand, besoin de tenir compte des différentes conformations protéiques
Amarrage moléculaire	Possible lorsque les informations sur les ligands sont rares, absence de biais envers les ligands existants, prédiction des modes de liaison	Simplification par les fonctions de score, coût de calcul élevé, différentes conformations protéiques et flexibilité des protéines

TABLE 1 – *comparaison "ligand-based" et "structure-based" pharmacophore et l'amarrage moléculaire*

Pour tirer partie des avantages complémentaires des différentes approches et compenser les limitations individuelles, des approches combinant LBDD et SBDD sont de plus en plus développées [173].

Les différentes approches stratégiques combinant SBDD et LBDD, définies par Drwal et Griffith [172]

— séquentielle : différentes méthodes à partir des informations issues du ligand ou basées

sur la structure sont utilisées pour s'intégrer dans le schéma d'un criblage virtuel en plusieurs étapes. De manière générale, les méthodes de plus en plus consommatrices en ressource de calculs sont utilisées au fur et à mesure que le nombre de composés à tester est réduit. Par exemple un criblage par pharmacophore peut être réalisé sur des bibliothèques de plusieurs milliers de composés et être suivi d'une évaluation par amarrage moléculaire au niveau du site de liaison de la protéine sur un ensemble de quelques centaines de candidats filtrés à travers le criblage. C'est le cas des travaux de Khan *et al.* portant sur l'identification de modulateurs de récepteurs 1 aux œstrogènes couplés aux protéines G [174].

- parallèle : utilisation de méthodes SBDD et LBDD en parallèle et indépendamment l'une de l'autre. Les meilleurs résultats de chaque méthode sont alors sélectionnés suivant un algorithme de fusion avant évaluation biologique. Par exemple, les travaux de Vucicevic *et al.* [175] portant sur la recherche de nouveaux ligands des récepteurs à l'imidazoline I₁ d'une part à l'aide d'une approche LBDD qui mesure la similarité du composé avec la rimelnidine et d'autre part une étude SBDD à partir du site actif de la nischarine. À la suite de ces méthodes, un score global a été attribué aux composés à partir de leurs scores issus des méthodes LBDD et SBDD.
- hybride : L'approche hybride repose sur une méthode à part entière combinant SBDD et LBDD. Par exemple dans l'amarrage moléculaire, la fonction de score qui permet de classer les différentes poses est un élément essentiel et critique de la méthode puisqu'elle permet de discriminer les éléments susceptibles de présenter un intérêt biologique (cf partie amarrage moléculaire). Incorporer à la fonction de score des éléments issus de méthode LBDD par une pondération du score de la fonction basée par exemple sur la similarité avec des ligands connus aboutit à une méthode hybride [176]. C'est le cas pour les travaux de Anighoro et Bajorath portant sur la recherche d'antagonistes au récepteur à l'adénosine A_{2A} par l'intégration d'un calcul de similarité des poses issues d'amarrages moléculaires à partir d'antagonistes cristallisés. Cette fonction de score présente une performance supérieure aux autres fonctions de score [177].

C'est la stratégie de criblage séquentielle, aussi dénommée hiérarchique, qui a été choisie pour les travaux présentés ici. C'est la stratégie la plus communément employée. Elle permet de faire appel aux méthodes les plus consommatrices en ressources computationnelles à la fin du criblage et donc sur un nombre limité de dérivés [173, 178].

3.2 Pharmacophores

Le développement de pharmacophores tout comme l'amarrage moléculaire s'appuie sur la spécificité de reconnaissance moléculaire des processus biologiques. Cette notion de spécificité a été développée au cours du XIX^{me} siècle. D'abord Langley développe l'idée que certains médicaments agissent de manière particulière sur des récepteurs spécifiques [179].

Théorisée par Emil Fisher en 1894 avec son analogie célèbre de "clé-serrure" [180]

, la reconnaissance moléculaire est nécessaire à de nombreux processus biologiques, par exemple dans la transduction de signaux ou de réactions enzymatiques et permet de comprendre l'effet pharmacologique d'un composé. La paternité du concept de pharmacophore est disputée car il a évolué au fil du temps [181]

], cependant les travaux fondateurs de Paul Ehrlich font date, bien qu'il n'utilise pas le terme de pharmacophore [182].

Il en définit le concept où des "groupes chimiques" ou fonctions sont responsables de l'effet biologique. Schueler, en 1960, emploie le terme pharmacophore associé à une définition plus moderne où il fait référence à un modèle de caractéristiques abstraites [183, 184].

Enfin, l'Union Internationale de Chimie Pure et Appliquée définit ainsi le pharmacophore : "Un pharmacophore est l'ensemble des caractéristiques stériques et électroniques nécessaires pour assurer l'interaction supramoléculaire optimale avec la cible biologique spécifique pour déclencher ou bloquer une réponse biologique" [185]

]. Un pharmacophore ne représente pas une molécule réelle ou une association de groupements fonctionnels, mais est un concept abstrait de caractéristiques qui regroupent les éléments communs à l'interaction moléculaire avec une cible. Les pharmacophores ont été utilisés avant l'avènement de l'outil informatique mais c'est avec celui-ci que leur utilisation se démocratise, en parallèle de l'évolution du concept de pharmacophore. Le premier logiciel commercial permettant la recherche de composés à partir d'un pharmacophore 3D est diffusé à partir de 1989 (CATALYST) [186]

] à la suite du développement des bases de données 3D [187].

]. C'est aussi à partir des années 2000 que le nombre de publications faisant référence aux termes "pharmacophore" et "virtual screening" augmente fortement comme l'indique la figure 32.

LigandScout® a permis les travaux décrits dans ce mémoire de thèse [188].

Ce logiciel permet notamment la création de pharmacophores visuels à partir de données structurales, notamment de complexes ligands-protéines.

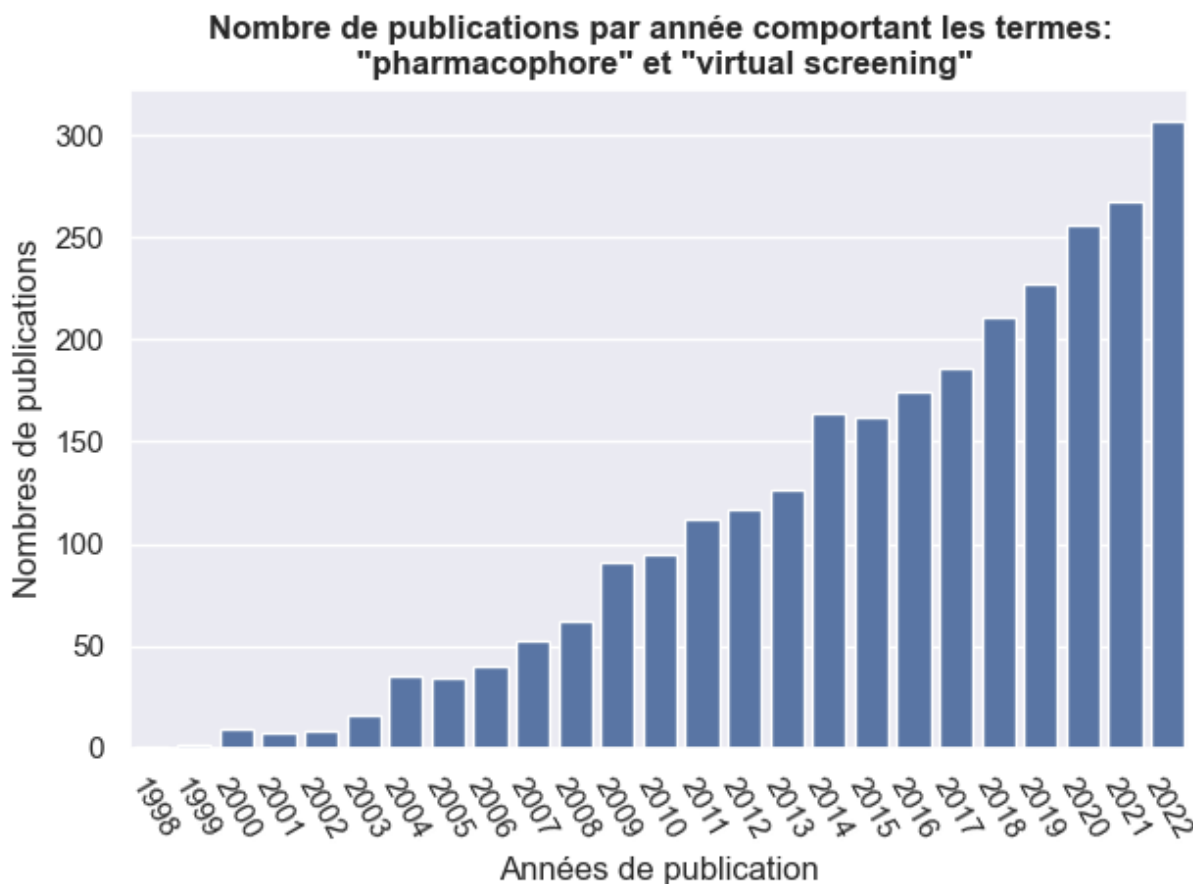


FIGURE 32 – Évolution du nombre de publications en fonction des années dans la base de données PubMed, en retour de la requête des mots clés "pharmacophore" et "virtual screening", produit à l'aide du paquet seaborn

Le pharmacophore obtenu se caractérise par différentes propriétés d'un ensemble de caractéristiques. Les caractéristiques chimiques représentées dans LigandScout® sont :

- HBD : Donneur de liaison hydrogène
- HBA : Accepteur de liaison hydrogène
- PI : Positivement ionisable
- NI : Négativement ionisable
- H : Hydrophobe
- aromatic feature : Aromatique (interaction π - π et cations- π)
- metal binding : Liaison métallique
- xvol : Volume d'exclusion

Ces caractéristiques se présentent visuellement selon la figure 33.

Caractéristique	Visuels
HBD	
HBA	
PI	
NI	
H	
Aromatic	
Metal	
Xvol	

FIGURE 33 – Représentation des différents attributs de pharmacophore dans LigandScout®

3.2.1 Pharmacophores de mPGES-1

Plusieurs structures cristallisées de mPGES-1 sont disponibles sur la PDB. Parmi celles-ci, plusieurs sont co-cristallisées en interaction avec un inhibiteur. Les structures utilisées dans le cadre des travaux présentés ici, sont : 4YK5, 4YL0, 4YL1, 4YL3, 5T36, 5BQG, 5BQH, 5BQI, 5KOI, 5TL6, 5T37, 5TL9. Les pharmacophores sont établis à partir de ces structures alignées deux à deux (confère Méthodes) et en modifiant le nombre de caractéristiques de ces pharmacophores.

Plusieurs pharmacophores ont été élaborés, puis évalués avant d'être modifiés de nouveau puis réévalués dans un cycle d'essais-erreurs afin d'obtenir plusieurs modèles de pharmacophores capables de distinguer au mieux les inhibiteurs des "decoys" (des composés putativement inactifs) du jeu de données de validation (jeu de données qui permet d'évaluer les pharmacophores créés). Quelques exemples de ces pharmacophores sont présentés en figure 34.

Représentation des différents pharmacophores

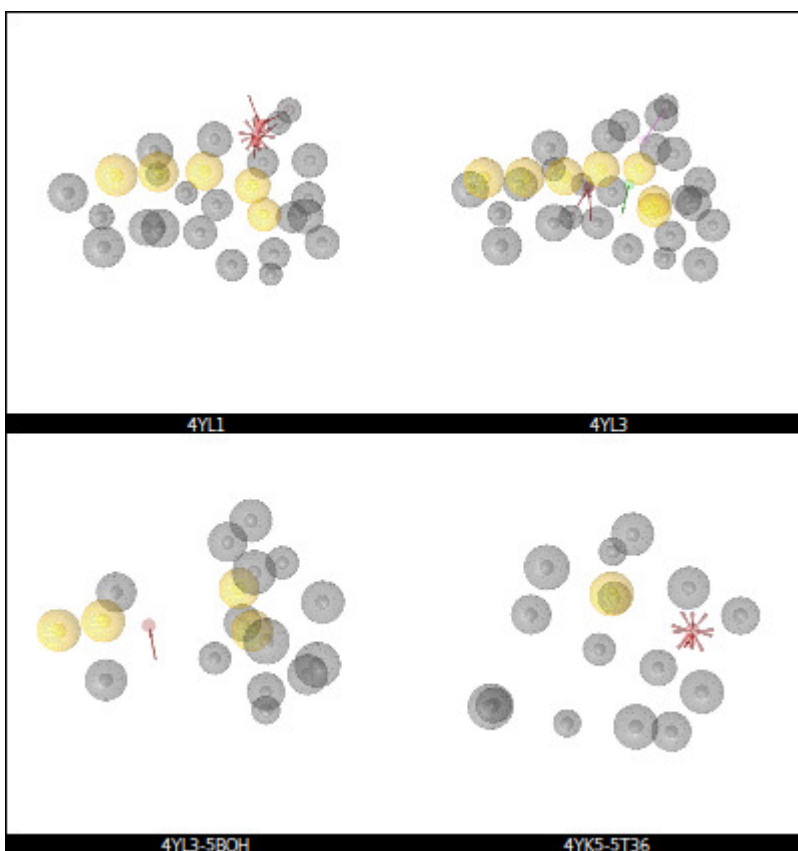


FIGURE 34 – Représentation de pharmacophores dans LigandScout®

3.2.2 Validation de pharmacophores

Le pharmacophore est un modèle dont l'objectif est de discriminer les molécules actives des molécules inactives. Dans ce sens, il est nécessaire d'évaluer sa propension à le faire. À partir d'un jeu de données de validation du modèle, plusieurs critères d'évaluation existent :

- Sensibilité [189] : La sensibilité représente la capacité du pharmacophore à sélectionner les composés actifs dans le jeu de données. C'est un ratio entre les composés actifs trouvés par le pharmacophore (Vrais positifs) par rapport à l'ensemble des composés actifs (vrais positifs et faux négatifs) où les faux négatifs sont les composés actifs non détectés par le pharmacophore. Ainsi, la sensibilité se décrit par la formule :

$$\text{Sensibilité} = \frac{\text{VraisPositifs}}{\text{VraisPositifs} + \text{FauxNégatifs}}$$

- Spécificité [189] : La spécificité est la propriété de rejeter les composés inactifs comme inactifs. C'est un ratio, entre les composés

inactifs retrouvés par le pharmacophore (Vrais Négatifs) par rapport à l'ensemble des composés inactifs (Vrais Négatifs et Faux Positifs) où les faux positifs sont les composés inactifs acceptés par le pharmacophore. Ainsi, la spécificité se décrit par la formule :

$$Spécificité = \frac{Vrais\ Négatifs}{Faux\ Positifs + Vrais\ Négatifs}$$

- Précision [190] : La précision est le pourcentage de composés correctement identifiés par le pharmacophore. C'est à dire les vrais positifs et vrais négatifs comparé à l'ensemble du jeu de données. La précision se formule :

$$Précision = \frac{Vrais\ Positifs + Vrais\ Négatifs}{\mathcal{P}(Jeu\ de\ données\ de\ validation)}$$

- Facteur d'enrichissement [191] : Le facteur d'enrichissement mesure combien de composés actifs sont trouvés dans les premiers x% de la base de données par rapport aux composés actifs présents dans le même pourcentage d'une liste randomisée (liste assemblée en tirant les éléments aléatoirement) de la même base.

$$Facteur\ d'enrichissement = \frac{\frac{Vrais\ Positifs^{x\%}}{\sum (Positifs + Négatifs)^{x\%}}}{\frac{Vrais\ Positifs}{\sum (Positifs + Négatifs)}}$$

Le facteur d'enrichissement présente par sa formule une sensibilité à la reconnaissance précoce, c'est à dire aux composés actifs présentant le meilleur score. Cependant, il est insensible au classement relatif des composés dans ces premiers x% [192]

- La courbe ROC [193] : La courbe ROC, contrairement aux précédents indicateurs, n'est pas seulement une valeur numérique mais aussi une représentation graphique. La courbe ROC présente en abscisse le ratio de faux positifs (1 - Spécificité) qui mesure combien de composés inactifs ont été classés comme actifs, et en ordonnée, la sensibilité. Ainsi une courbe ROC idéale part de l'origine et monte verticalement au dessus de l'origine jusqu'au coin supérieur gauche. Cela signifie que tous les composés actifs ont été trouvés. Puis, elle arrive jusqu'au coin supérieur droit, puisque seul les "decoys" restent à trouver. Au contraire, dans une liste randomisée entre "decoys" et composés actifs, la sensibilité égale le ratio de faux positifs et la courbe ROC dessinée est une diagonale de l'origine au coin supérieur droit, avec l'augmentation concomitante des "decoys" et composés actifs trouvés.

Jeu de données de validation des pharmacophores de mPGES-1

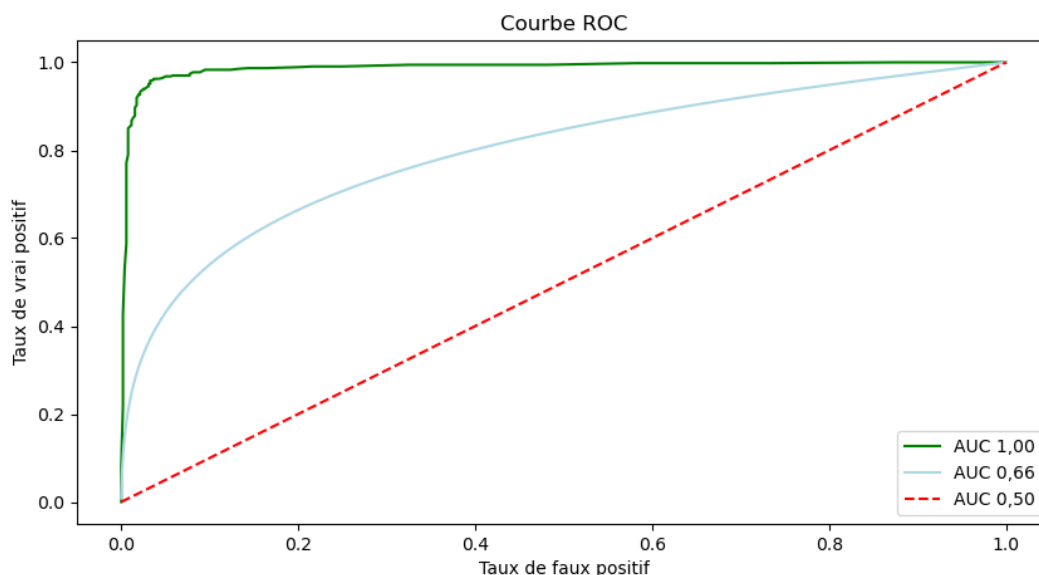


FIGURE 35 – Représentation de courbes ROC avec différentes aires sous la courbe

- L'aire sous la courbe ROC : la mesure de l'aire sous la courbe est un classificateur compris entre 0 et 1 qui se compare à une valeur de classification aléatoire de 0,5. Elle permet de distinguer visuellement et numériquement la capacité d'un pharmacophore à mieux classer les molécules actives que l'aléatoire. Un score de 1 correspond à une discrimination parfaite entre les actifs et inactifs, un score entre 1 et 0.5 représente une discrimination de moins en moins efficace entre une discrimination excellente et une aléatoire.

En plus de la propension à distinguer les composés actifs des "decoys", comme dans un criblage classique, il est important que les composés actifs soient mieux classés que les composés "decoys", c'est le problème de reconnaissance précoce. Seulement, les composés les mieux classés sont testés car l'expérimentation est plus coûteuse que la recherche computationnelle. Par exemple, sur un criblage d'un million de composés si 0,1% des composés sont testés cela représente néanmoins un millier de composés [194].]

3.2.3 Jeu de données test des pharmacophores

Pour la validation des pharmacophores, un jeu de données comportant des inhibiteurs connus de mPGES-1 et des composés testés comme non-inhibiteurs de mPGES-1 (comme le recommande les travaux de M.Réau *et al.*[195]) complété par des composés "decoys" pour obtenir un ratio d'inhibiteurs actifs/inactifs similaires à ceux trouvés dans un criblage haut débit. Par exemple, M. Lindh *et al.* dans leurs travaux pour préparer plusieurs jeux de données de validation présentent un ratio actif/inactif

proche de 1% (19/59405 à 369/337634) [196]. La sélection des "decoys" est importante car elle impacte directement les résultats comme l'ont montrés les travaux de L.Chaput *et al.* en comparant les performances de plusieurs programmes à travers différents jeux de données actifs/inactifs de référence [197]. Cependant, au moins pour certains travaux, une plage de ratio actifs/inactifs n'impacte pas les résultats; J.Xia *et al.* ont fait varier un ratio actifs/inactifs de 1/30 à 1/100 sans impact sur les résultats de leurs travaux ciblant les récepteurs couplés aux protéines G [198]. Pour évaluer la diversité structurale de ces composés et l'homogénéité structurale entre ces deux groupes, une représentation à partir du logP et du poids moléculaire est effectuée, ainsi qu'une projection après réduction de dimension privilégiant la variabilité intergroupe (UMAP) à partir de plusieurs propriétés physico-chimiques (logP, poids moléculaire, nombre de donneurs de liaison hydrogène, nombre d'accepteurs de liaison hydrogène, surface polaire topologique et le nombre de liaisons permettant une rotation libre du ligand) calculées à l'aide de RDkit.

Jeu de données de validation des pharmacophores de mPGES-1

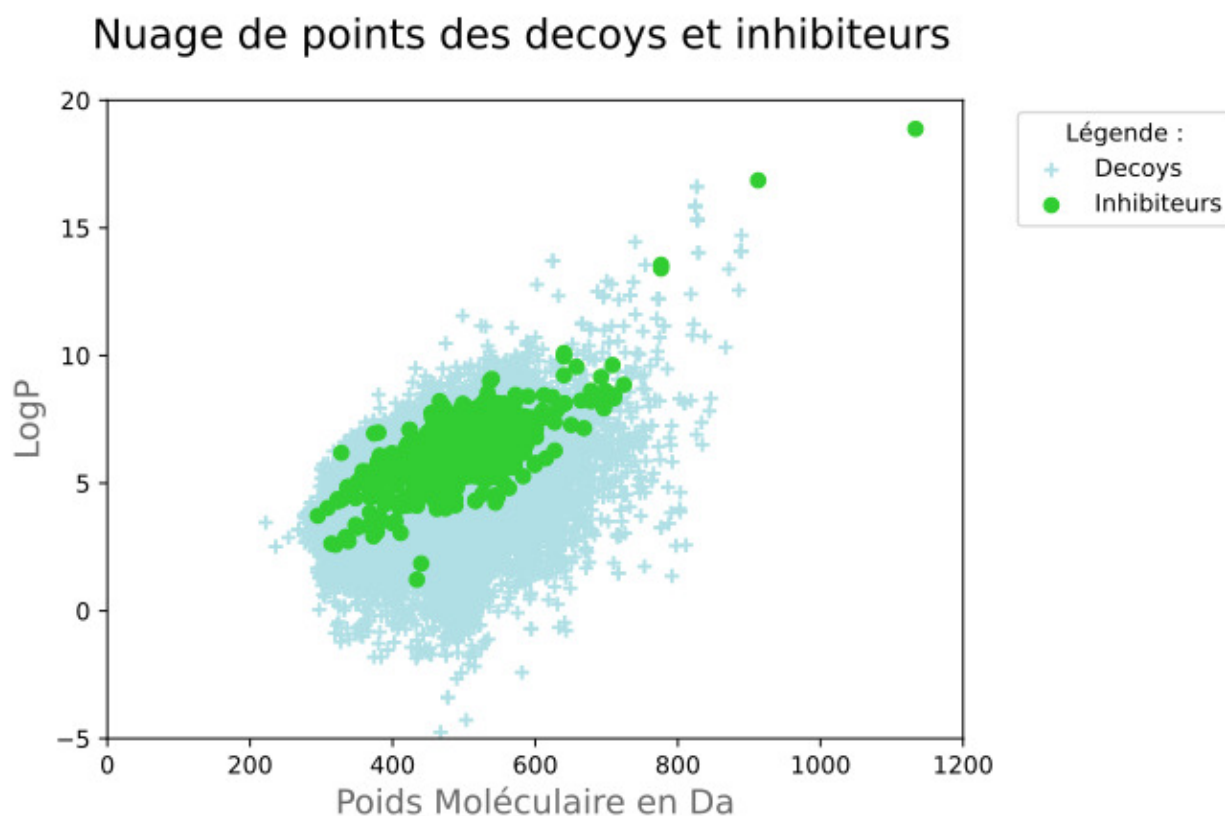


FIGURE 36 – Représentation du jeu de données pour la validation de pharmacophores de mPGES-1 à partir du LogP et du poids moléculaire

La figure 36 présente le jeu de données en fonction du logP et du poids moléculaire. Elle montre une cohérence entre les nuages de points des decoys et des inhibiteurs de mPGES-1

Projection des propriétés physico-chimiques du jeu de données de validation des pharmacophores de mPGES-1

Nuage de points des decoys et inhibiteurs projeté après réduction de dimensions par UMAP

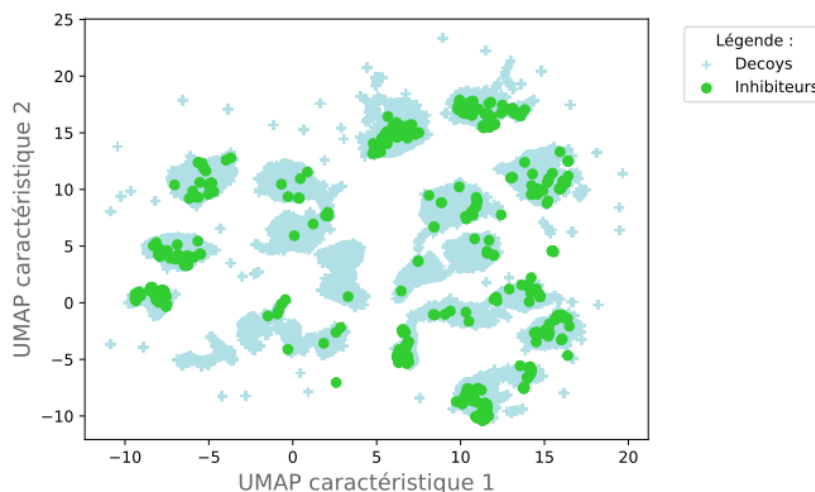


FIGURE 37 – Représentation du jeu de données pour la validation de pharmacophores de mPGES-1 à partir de plusieurs propriétés physico-chimiques calculées à l’aide de rdkit ($\log P$ calculé, poids moléculaire, nombre de donneurs de liaison hydrogène, nombre d’accepteurs de liaison hydrogène, surface polaire topologique et le nombre de liaisons permettant une rotation libre du ligand)

avec un poids moléculaire similaire entre les nuages de points et une distribution sur une plage du $\log P$ plus importante pour les decoys.

Sur la figure 37 plusieurs groupes sont visibles. Ces groupes comprennent des inhibiteurs et decoys, hormis un regroupement de decoys seuls. Il est raisonnable de spéculer d’un noyau structural commun au sein d’un groupe. Ceci permet de valider la diversité structurale recherchée entre composés actifs et inactifs. Cependant le ratio inhibiteurs/decoys varie d’un groupe à l’autre. Ce ratio est un critère important pouvant influencer les résultats si le nombre de decoys n’est pas assez important. Cependant sans méthode d’apprentissage artificiel de regroupement, il n’est pas possible de conclure si un regroupement sur le graphique correspond effectivement à un regroupement présentant un noyau structurale homologue ou à l’aggrégation de plusieurs groupes. Cette méthode aurait pu permettre d’établir, en plus des groupes, le ratio actif/decoys pour chaque groupe. Cependant, la méthode de construction du jeu de données (cf Méthodes) permet de supposer d’un nombre suffisant de decoys par inhibiteur.

Ce jeu de validation est ensuite utilisé pour les pharmacophores établis. Les résultats sont visibles sur les figures de la section suivante.

3.2.4 Résultats : évaluation des pharmacophores

Les pharmacophores ont été établis et évalués à l’aide de LigandScout®(cf Méthodes) et du jeu de données de test.

- 4YK5-5T36 Ce pharmacophore présente une aire sous la courbe à respectivement 1 ;5 et 10% de 0,99 ; 0,97 et 0,94. et le facteur d'enrichissement sur le premier pourcent de la base de données est aussi important. Il est de 7,3 pour le premier pourcent mais diminue rapidement et atteint un score bas à partir de 5% à seulement 2,2. Ce pharmacophore permet de retrouver 14% des inhibiteurs connus de mPGES-1 soit 101 inhibiteurs parmi les 724 du jeu de données pour un total de 6903 correspondances. Cela correspond à une sensibilité de 14%, une spécificité de 84% et une précision de 83%. Il présente donc une spécificité et une précision importante mais manque de sensibilité puisqu'il trouve en proportion sensiblement le même nombre de composés actifs que d'inactifs. Cependant, vu le facteur d'enrichissement et l'aire sous la courbe, les inhibiteurs retrouvés sont mieux classés que les faux positifs.

Pharmacophore 4YK5-5T36 et courbe ROC

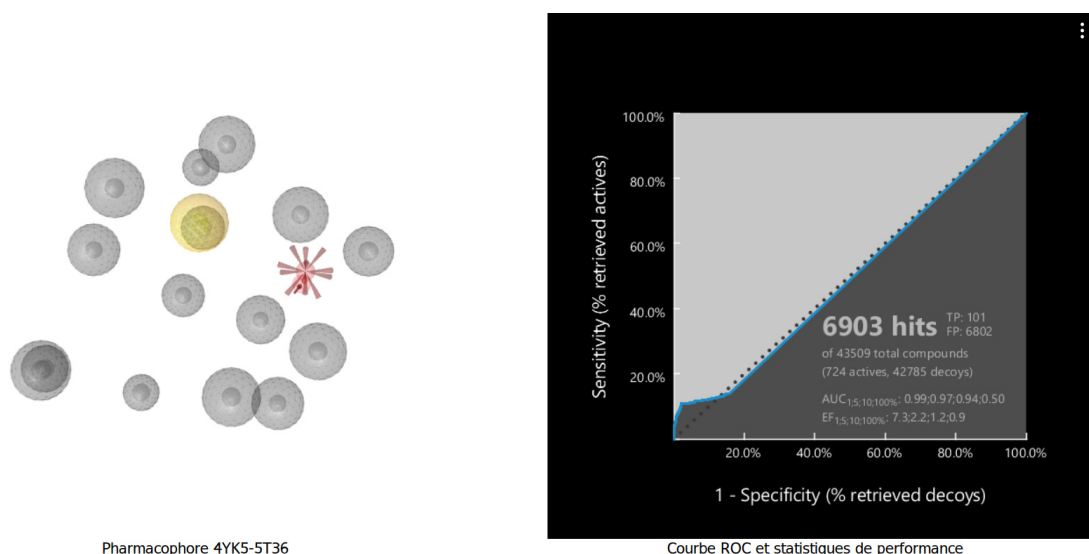


FIGURE 38 – *Pharmacophore 4YK5-5T36 et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

- 5KOI-5TL9 Ce pharmacophore permet de retrouver plus de 44% des inhibiteurs de mPGES-1 mais est aussi moins restrictif sur les faux positifs trouvés. L'aire sous la courbe est jusqu'au 10% supérieure ou égale à 0,9. Le facteur d'enrichissement est moins élevé que pour le pharmacophore précédent (5,9 à 1%) et se dégrade ensuite mais moins rapidement que le précédent (2,7 à 10%) même si cela reste un facteur d'enrichissement faible. Il présente une précision de 80%, une sensibilité de 44% et une spécificité de 81%. Ce pharmacophore est moins spécifique que le précédent ce qui lui permet de retrouver plus d'inhibiteurs de mPGES-1. Comme pour le précédent, le facteur d'enrichissement plus élevé et l'aire sous la courbe montre que plus d'inhibiteurs sont mieux classés au début du jeu de données.

Pharmacophore 5KOI-5TL9 et courbe ROC

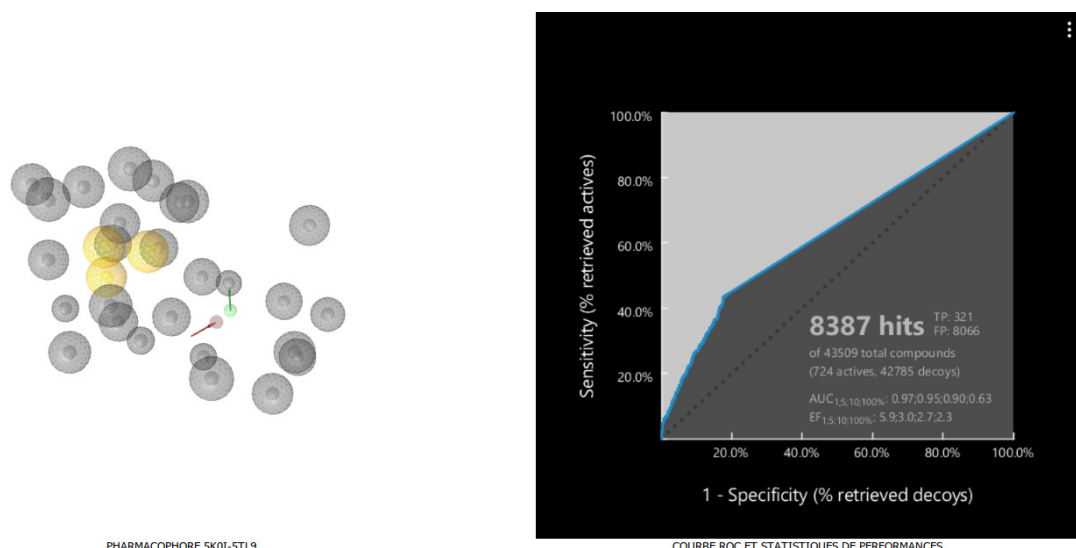


FIGURE 39 – *Pharmacophore 5KOI-5TL9 et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

- 4YL0-5BQG Ce pharmacophore retrouve plus de la moitié (57%) des inhibiteurs et seulement 20% des composés "decoys", cependant l'aire sous la courbe de ce pharmacophore est moins bonne que les pharmacophores précédents comme les autres critères de performance de ce pharmacophore, particulièrement lorsque l'on considère les premiers pourcentages de la base de données. Il présente une sensibilité de 57%; une spécificité de 77%; et une précision de 77%. Cependant, l'aire sous la courbe à 1;5 et 10% montre que la discrimination entre composés inactifs et inhibiteurs est moins discriminante (0,25; 0,73; 0,76) donc au moins une partie des composés inactifs sont mieux classés que les inhibiteurs de mPGES-1; de même que le montre le facteur d'enrichissement qui est bas.
- 4YL0-5BQH L'évaluation de ce pharmacophore sur le jeu de données de validation présente une aire sous la courbe de plus de 0.9 jusqu'à 10% de la base de données. Un facteur d'enrichissement est aussi important de presque 5 à 5% de la base de données et d'encre 4 à 10% du jeu de validation. De plus il permet de trouver plus de 50% des inhibiteurs de mPGES-1 inclus dans la base d'évaluation. Il présente ainsi une sensibilité de 50%; une spécificité de 82% et une précision de 82%. Ce pharmacophore présente une sensibilité augmentée par rapport aux autres pharmacophores (hormis 4YL0-5BQG) cependant la spécificité est un peu plus basse que les autres pharmacophores.
- 4YL3-5BQG Ce pharmacophore est assez sélectif, peu de composés sont sélectionnés avec ce pharmacophore. Il présente une sensibilité de 22%; une spécificité de 93% et une précision de 92%. Le facteur d'enrichissement est correct, respectivement de 4,7 à

Pharmacophore 4YL0-5BQG et courbe ROC

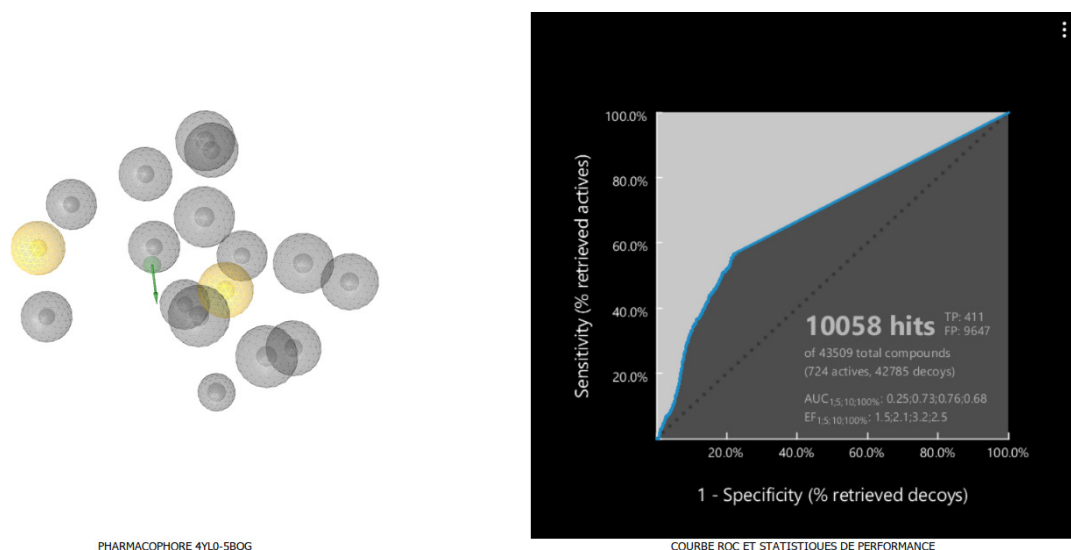


FIGURE 40 – *Pharmacophore 4YL0-5BQG et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

Pharmacophore 4YL0-5BQH et courbe ROC

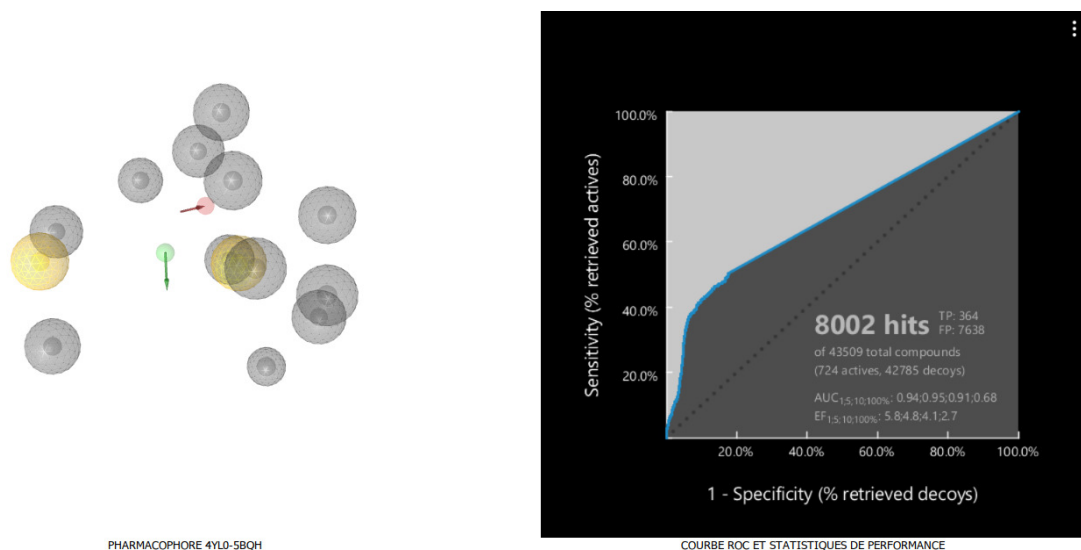


FIGURE 41 – *Pharmacophore 4YL0-5BQH et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

1% et de plus de 3 à 10%. Cependant l'aire sous la courbe à 1% est seulement de 0,6 ce qui montre qu'une partie importante des composés les mieux classés sont inactifs.

Pharmacophore 4YL3-5BQG et courbe ROC

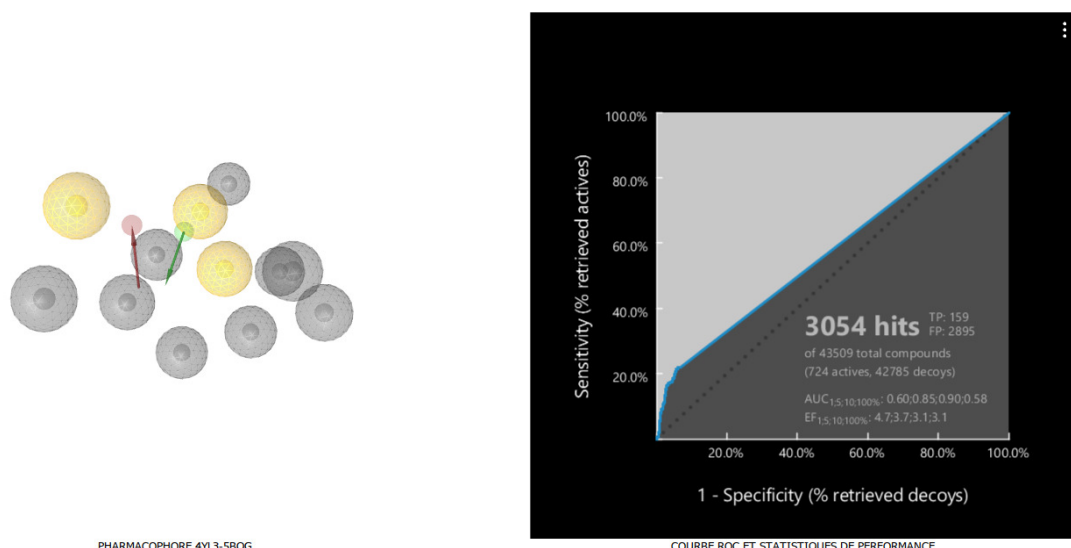


FIGURE 42 – *Pharmacophore 4YL3-5BQG et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

- 4YL3-5BQH L'évaluation de ce pharmacophore présente une aire sous la courbe de 0,86 à 1% et de 0,74 à 10% et un facteur d'enrichissement faible (2,5 à 1%). Il présente une sensibilité de 14,50% ; une spécificité de 94% et une précision de 93%. Ce pharmacophore présente une performance moyenne avec un facteur d'enrichissement faible et une aire sous la courbe moyenne aussi comparée aux autres pharmacophores.
- 5BQG-5BGH Ce pharmacophore est peu sélectif puisqu'il retrouve 63% des inhibiteurs du jeu de données mais aussi 32% des composés "decoys" du même ensemble. L'aire sous la courbe montre aussi cette sélection large avec une aire sous la courbe qui ne dépasse pas les 0,75. Le facteur d'enrichissement est aussi très faible de 2,5 à 1% de l'ensemble de test puis proche de 1 ensuite. Il présente une sensibilité de 63% ; une spécificité de 67% et une précision de 67%.

Tous les pharmacophores établis n'ont pas été utilisés pour la suite du criblage virtuel(annexe 1). Les pharmacophores utilisés ne présentent pas toujours des performances excellentes, cela peut s'expliquer au moins partiellement par la diversité des sites de liaison au sein du site catalytique de l'enzyme. Deux sites se distinguent dans le site en forme d'épingle de nourrice et la diversité du jeu de données comportant des composés pour les deux sites distincts. Les pharmacophores présentant les meilleurs critères sont utilisés pour cribler les différentes bases de données construites ou existantes au sein du laboratoire(confère Méthodes). Tous les pharmacophores présentés ont été utilisés (hormis ceux en annexe) pour la suite des travaux. Chacun représente un consensus entre deux inhibiteurs propres de et

Pharmacophore 4YL3-5BQH et courbe ROC

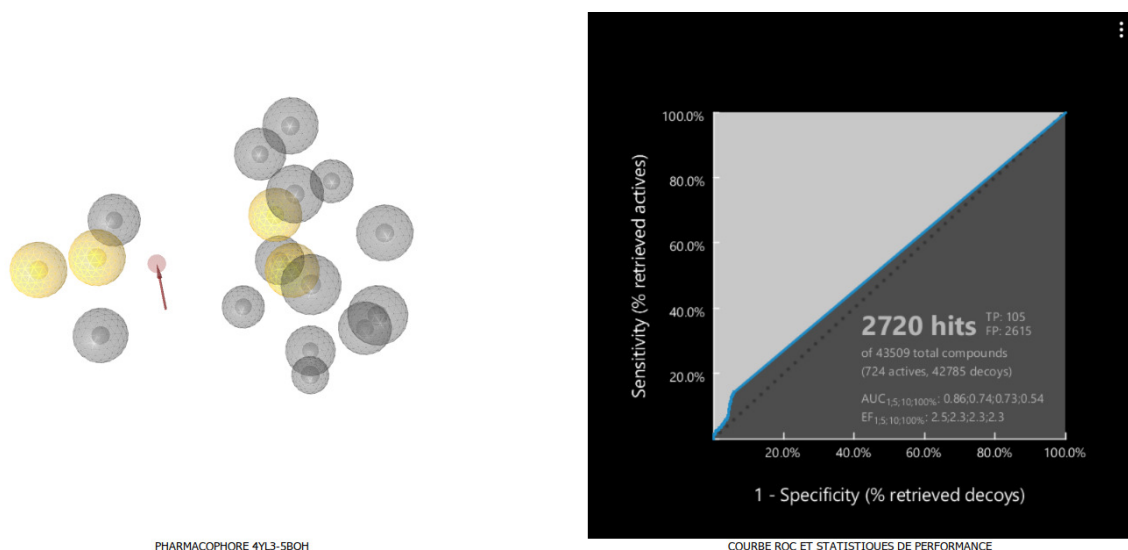


FIGURE 43 – *Pharmacophore 4YL3-5BQH et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

Pharmacophore 5BQG-5BGH et courbe ROC

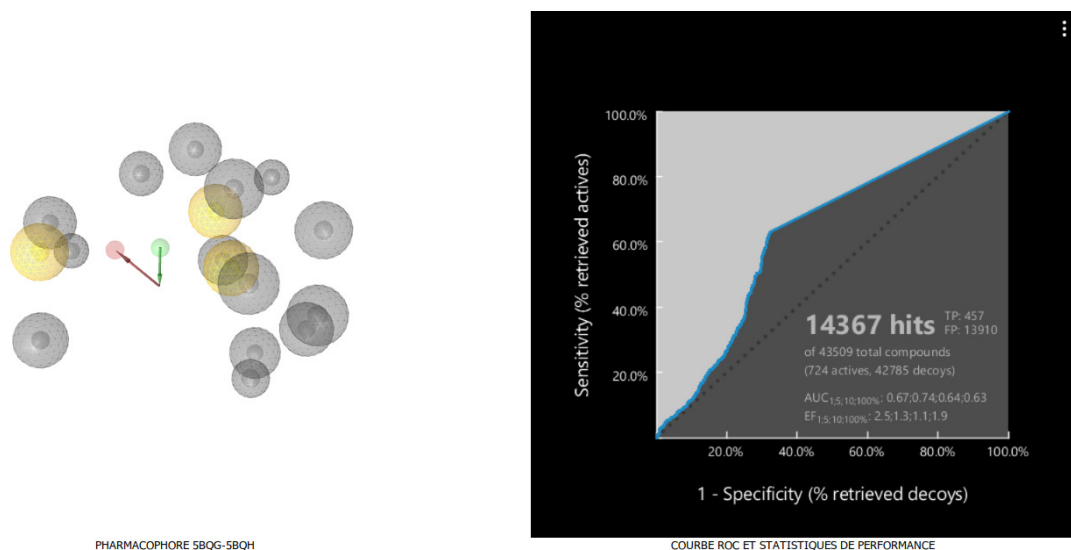


FIGURE 44 – *Pharmacophore 5BQG-5BGH et sa courbe ROC ainsi que des statistiques de performances par rapport au jeu de données de validation*

présentant des critères de performance suffisant particulièrement pour les composés les mieux classés. Globalement, une différence de performance entre les pharmacophores apparaît suivant la poche qu'ils ciblent.

3.2.5 Filtration des résultats sur les propriétés physico-chimiques

Les molécules issues du criblage à partir des pharmacophores ont été filtrées sur plusieurs caractéristiques. Au cours des dernières années plusieurs motifs fonctionnels ont été identifiés comme présentant une réaction positive sur de nombreux tests biologiques (Pan-Assay Interference Compounds (PAINS)), cette réaction n'étant pas spécifique ils sont considérés comme des faux positifs [199]. Cependant une utilisation trop rigoriste ou simpliste peut mener à éliminer un composé potentiellement d'intérêt ou ne pas étiqueter un composé qui n'a pas d'intérêt pour un développement [200]. Parmi les bibliothèques de composés criblés, plusieurs étaient déjà disponibles au sein de l'unité de recherche SONAS (ZINC et UNDP). Pour la bibliothèque de composés "COCONUT" [201] la filtration des "PAINS" n'a pas été effectuée.

Dans la recherche de composés chef de file, similairement aux règles des 5 de Lipinsky [202], un poids moléculaire inférieur à 500 Daltons, un logP calculé inférieur à 5, un nombre d'accepteurs de liaisons hydrogènes inférieur à 10 et à 5 pour les donneurs de liaison H ; pour une biodisponibilité correcte par voie orale des candidats médicaments ; Oprea T.I. *et al.* à partir de l'effort développement pour les bases de données en ligne [203] déterminent ces règles pour les composés chef de file, un poids médian plus bas de 290 Daltons et une distribution à 90% de 436 Daltons, un nombre de donneurs de liaisons hydrogène et d'accepteurs de liaisons hydrogène moins grand (3 et 8.6 à 90% respectivement), un logP calculé inférieur à 5 [204].

Ainsi les composés, après le criblage virtuel, ont été filtrés avec les conditions suivantes :

Masse Moléculaire	inférieure à 400Da
LogP	supérieur ou égal à 1 et inférieur ou égal à 3,1
nombre de donneurs de liaisons hydrogènes	7 maximum
nombre d'accepteurs de liaisons hydrogènes	12 maximum

TABLE 2 – Règles de filtration des composés criblés

3.3 Amarrage moléculaire

Les résultats issus du criblage virtuel à partir des pharmacophores de mPGES-1 ont été confirmés par d'autres méthodes computationnelles. Plusieurs méthodes existent, principalement l'amarrage moléculaire et la simulation de dynamique moléculaire. C'est une étude d'amarrage moléculaire qui a été effectuée ici, en rapport avec les ressources computationnelles.

nelles disponibles au sein de l'unité SONAS, de l'expertise de l'équipe de recherche et du temps nécessaire à l'exploration d'un nombre de composés importants. L'amarrage moléculaire, comme l'approche pharmacophorique basée sur la structure, s'appuie sur le principe de reconnaissance moléculaire. Le déclenchement d'un processus biologique ou l'effet pharmacologique d'un composé se fait selon cette reconnaissance moléculaire, qui peut aussi concerner plusieurs cibles [205]. L'amarrage moléculaire consiste donc, à partir de la structure d'une cible, le plus souvent une protéine, et de la structure moléculaire d'un ligand, à établir leur interaction au sein d'un complexe ligand-protéine, c'est à dire la position, l'orientation et la configuration du ligand en considérant également les échanges énergétiques entre entités. Depuis ses débuts dans les années 80 [206], l'amarrage moléculaire a connu plusieurs succès et évolutions. Bien que d'apparence cela soit semblable à trouver la complémentarité 3D entre deux entités rigides, en réalité la flexibilité des entités appelle un grand nombre de degrés de liberté ou d'interactions entre les entités et le solvant, avec des contributions d'enthalpie et d'entropie complexes. L'amarrage moléculaire consiste donc à rechercher des minima énergétiques locaux dans cet espace complexe à partir d'heuristiques approxi- mant les calculs, mais aussi le temps nécessaire à une hypothèse probable des interactions protéine-ligands.

Les entités en interaction dans une problématique d'amarrage sont donc d'une part la protéine, qui est bien plus grande que le ligand, celui-ci est en interaction avec une portion seulement de celle-ci qu'on appelle plus communément une "poche". C'est cette partie de la protéine (proche du ligand) à qui on accorde plus de flexibilité. Pour le ligand, non seulement la structure moléculaire doit être connue, mais aussi sa configuration absolue.

L'amarrage se divise en 2 étapes principales : une phase exploratoire "le searching", une phase quantitative et qualitative "le scoring".

3.3.1 Phase exploratoire

La phase de "searching", permet d'explorer l'espace énergétique et de configuration de l'interaction protéine-ligand pour trouver les minima locaux d'énergie, de l'énergie globale de liaison. Chacun de ces minima d'énergie est une solution probable de l'interaction protéine-ligand et s'appelle une "pose", mais c'est le minima global qui est la conformation protéine-ligand la plus probable sous le critère énergétique. Pour explorer cet espace de configuration, différents algorithmes existent [207]. Les algorithmes se divisent en deux catégories, les algorithmes systématiques et les algorithmes stochastiques :

- Les algorithmes systématiques : Ils explorent l'espace des configurations possibles à travers de faibles variations de la conformation du ligand [208]. C'est une méthode efficace mais qui peut être bloquée dans un minimum local d'énergie.

Cela peut être surmonté en répétant la méthode à partir de différents niveaux d'énergie du ligand [209]. Cependant, ces algorithmes se heurtent au problème d'explosion combinatoire [210], c'est à dire que selon le volume de données la solution peut devenir difficilement atteignable ou impossible. Parmi les algorithmes systématiques se trouvent l'algorithme de recherche exhaustive, l'algorithme incrémental par fragmentation, l'algorithme par méthode d'ensemble de conformations [211].

- Les algorithmes stochastiques : Ils explorent l'espace des configurations par des variations aléatoires. Ces algorithmes perforent généralement à partir d'un ensemble (ou population) de configurations initiales différentes. Cette méthode échappe mieux au piège du minimum local que les méthodes systématiques partant d'une population initiale variée [212]. Parmi les algorithmes stochastiques se trouvent l'algorithme de Monte-Carlo, l'algorithme génétique et l'algorithme de recherche avec tabou [213]. Brielvement, Monte-Carlo utilisent des échantillonnages aléatoire pour estimer des solution. Cela le rend flexible et robuste aux espaces de recherche complexes. Cependant la convergence est lente et sujette au bruit statistique et l'absence de mémoire (contrairement aux algorithmes génétiques et par recherche avec tabou) ne le rend pas efficace pour l'optimisation pure et le reserve plutôt à l'estimation, l'exploration. La recherche avec tabou au contraire de Monte-Carlo est une méthode d'optimisation avec mémoire qui evite de revisiter les solutions récentes (tabou). Il est flexible, peut eviter les optimum locaux grâce à la memoire. Il est sensible aux paramètres et difficile à parraléliser.

Au sein du logiciel GOLD [214] utilisé pour ce travail de thèse, un algorithme génétique est implémenté pour la phase exploratoire. C'est un exemple de méthode stochastique. Une population de "chromosomes" est générée, où chaque chromosome code pour les paramètres d'une solution candidate (la position, l'orientation et la conformation des variables de translation, de rotation et de torsion) sous la forme d'un vecteur. Puis ce chromosome est soumis à des mutations et croisements génétiques, où chaque génération, comme dans la théorie de l'évolution, transmet majoritairement les caractères les mieux adaptés jusqu'à ce qu'un nombre limite de générations(paramètre à régler dans ligandscout®par défaut=30) soit atteint ou qu'un critère de terminaison précoce, comme une amélioration marginale de l'écart quadratique à la moyenne(écart-type), soit atteint.

3.3.2 Le "scoring"

Le "scoring" est une étape essentielle du processus de docking car elle permet d'évaluer et classer les différentes poses obtenues lors de l'étape précédente. C'est le "scoring" qui permet de différencier la pose la plus réaliste du ligand des autres poses proposées et qui permet de

retrouver les vrais positifs sur la base de l'évaluation de l'énergie libre de liaison du complexe ligand-protéine. Cependant, bien que l'énergie libre de liaison est la plus à même de classer les poses, elle n'est pas facilement accessible. Elle s'évalue à partir des éléments impliqués dans la liaison ligand-récepteur, incluant les interactions intermoléculaires, la désolvatation et l'entropie [215].

]. Des méthodes de simulation de l'énergie libre de liaisons ont été développées [216]] mais cette fonction ne peut pas être appliquée facilement sur un grand nombre de composés. En effet, plus le nombre de paramètres physico-chimiques pris en compte est important, plus l'on peut espérer de justesse dans la fonction. Cependant le coût de calcul combinatoire augmente en proportion du nombre de paramètres inclus dans la fonction. C'est pourquoi d'autres fonctions de "scoring" existent basées sur des approximations mais permettant de garder une précision importante.

Il existe principalement 3 types de fonction de score : les fonctions basées sur les champs de force, les fonctions empiriques et les fonctions basées sur la connaissance.

- Les fonctions de score basées sur le champ de force sont développées à partir des interactions physiques à un niveau atomique, incluant les interactions de Van der Waals, les interactions électrostatiques et les forces d'étirement, de torsion et de flexion des liaisons. Les fonctions de champs de force sont développées à la fois à partir de données expérimentales et *ab initio* à partir de la mécanique quantique [217].

Un exemple de champ de force est Amber [218]] pour les protéines et acides nucléiques, qui a plus tard été élargi aux molécules organiques [219].

]. D'autres champs de forces existent suivant l'objet de recherche (CHARMM, GROMOS, LFMM, ANI, SYRAH, MARTINI).

- Les fonctions empiriques estiment l'affinité de liaison à partir d'un ensemble pondéré de contraintes énergétiques. Ces contraintes énergétiques peuvent inclure Van der Waals, les liaisons hydrogènes, la désolvatation, l'entropie, l'hydrophobicité. La pondération de chaque interaction est obtenue à partir d'un ensemble d'entraînements de structures tridimensionnelles de complexes protéines-ligands. Un exemple de fonction empirique est ChemScore, développée par Eldridge MD *et al.* et calibrée à partir d'un ensemble de 82 complexes ligand-récepteurs. Elle prend en compte les liaisons hydrogènes, les atomes métalliques, la lipophilie et le nombre de liaisons permettant une rotation libre du ligand [220].

]. Les fonctions empiriques sont plus rapides que les fonctions de champ de force pour le calcul de l'affinité de liaison.

- Les fonctions basées sur la connaissance aussi dénommées fonctions basées sur le potentiel statistique. Ce sont des fonctions de score qui définissent paire à paire les potentiels énergétiques qui sont établis à partir d'un ensemble de complexes protéines-ligands pour construire une fonction générale. Ces potentiels sont construits en prenant en compte la fréquence à laquelle 2 atomes sont retrouvés à une distance propre et pondé-

rée en fonction de la fréquence d’occurrence des interactions observées. Le score final est obtenu en sommant l’ensemble de ces potentiels énergétiques. Un avantage de ces fonctions est qu’elles permettent par leur simplicité computationnelle de cribler un grand nombre de composés. Une limite cependant est qu’elles sont bornées par l’information de l’ensemble du complexe protéine-ligand encodé par la fonction.

Les molécules d’eau sont aussi importantes dans la construction d’un modèle d’amarrage moléculaire, puisque environ 65% des structures cristallographiques de complexe protéine-ligand comportent au moins une molécule d’eau impliquée dans l’interaction [221] comme c’est le cas pour mPGES-1. La structure de la protéine pour l’amarrage moléculaire ayant déjà été préparée au sein de l’unité SONAS en collaboration avec l’équipe de Daniela SCHUSTER (PMU, Salzbourg) lors de travaux précédents, la discussion des molécules d’eau retenues ne sera pas traitée ici.

3.3.3 Résultats des études d’amarrage moléculaire

3.3.3.1 Préparation des ligands

Avant les études d’amarrage moléculaire, en plus de la préparation de la protéine qui peut influencer fortement sur les résultats des études d’amarrage moléculaire, la préparation des ligands est aussi importante [222].

La documentation de CCDC Gold recommande un logiciel (CSD Conformer Generator) pour la minimisation énergétique. C’est celui-ci qui a été utilisé pour les ligands après s’être assuré de l’ajout des hydrogènes à l’aide de RDkit.

3.3.3.2 Choix de la fonction de score

Plusieurs fonctions de score existent. La fonction choisie pour cette étude d’amarrage moléculaire est ChemPLP. Elle est plus rapide comparée à d’autres fonctions de score tel que ChemScore ou ASP [223].

3.3.3.3 Calcul de l’étude d’amarrage et curage manuel

CCDC Gold a été utilisé pour l’étude d’amarrage moléculaire à partir de la protéine mPGES-1 préalablement échafaudée et des ligands sélectionnés à partir du criblage virtuel utilisant les pharmacophores précédents. Ces ligands ont ensuite été préparés pour l’étude d’amarrage moléculaire.

Les ligands présentant le meilleur score calculé à partir de l’étude d’amarrage par la fonction de score ChemPLP ont ensuite été inspectés à l’aide de Ligandscout® pour un curage manuel. Ainsi, la visualisation des résultats de l’étude d’amarrage permet d’exclure les résultats où la pose du ligand semble aberrante par rapport au site d’interaction avec mPGES-1.

Ce curage a permis de sélectionner plusieurs molécules susceptibles de présenter un intérêt comme inhibiteur potentiel de mPGES-1.

Capture d'écran du logiciel LigandScout permettant des visualisations 2D et 3D de l'interaction ligand-site actif de mPGES-1

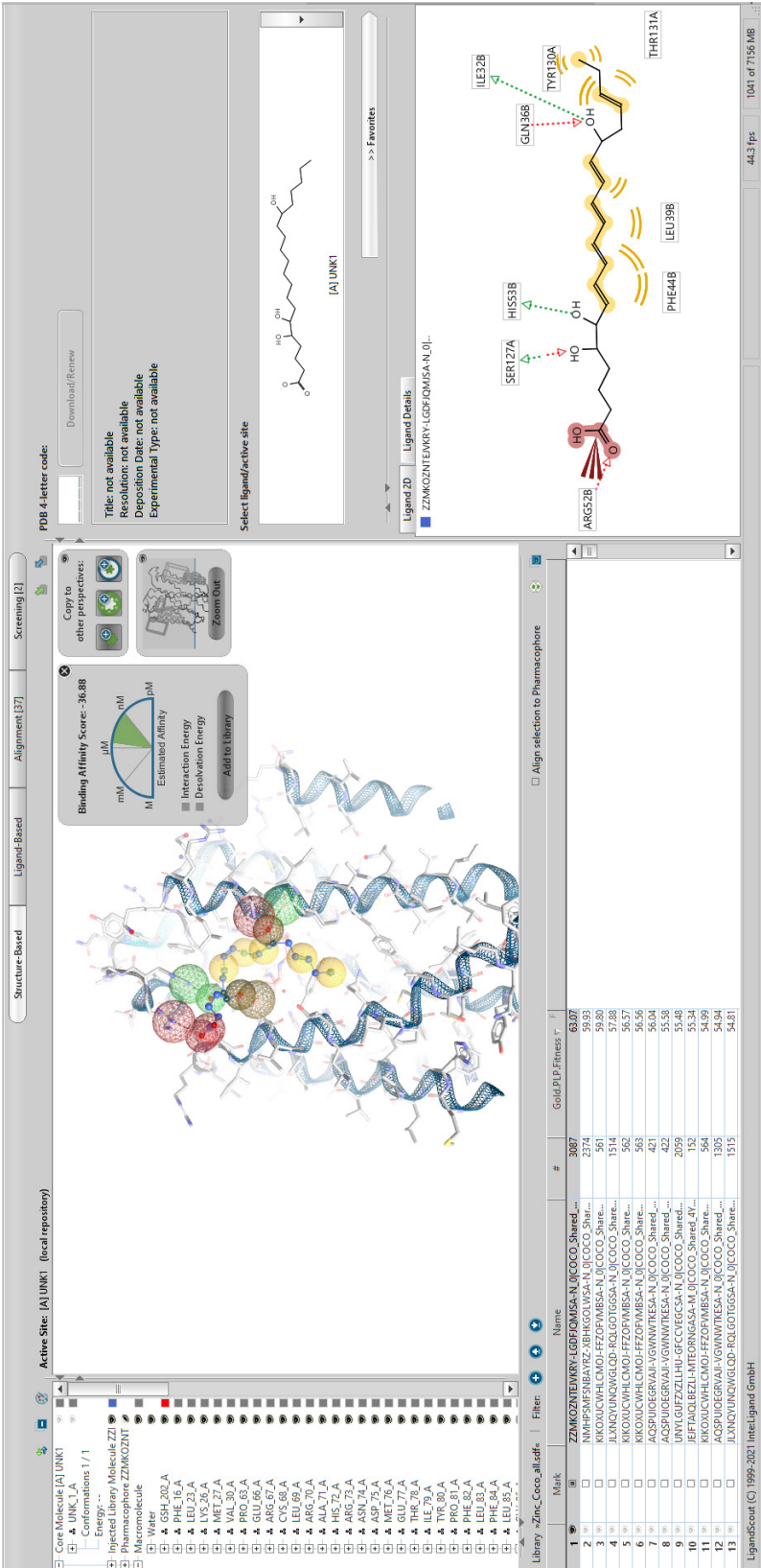


FIGURE 45 – exemple du travail de curage manuel à l'aide de LigandScout®

3.3.3.4 Résultats après les études d'amarrage moléculaire et du curage manuel

Lors du curage manuel, plusieurs composés ont été sélectionnés car ils remplissent les critères de filtration établis à partir des différents pharmacophores et des critères physico-chimiques, tout en présentant également un score élevé lors des études d'amarrage moléculaire. Ainsi ils sont susceptibles d'avoir une forte interaction avec mPGES-1. Ces composés sont présentés ci-après.

Deux composés ont été sélectionnés à partir de la base de données ZINC.

- Le composé 1 de dénomination IUPAC : 2-(6,7-dihydro-4H-thiéo[3,2-c]pyridin-5-yl)-1-[(2R,6R)-2-méthyl-6-phénylmorpholin-4-yl]éthanone

Ce composé présente un poids moléculaire calculé de 356,49 g/mol, un LogP calculé à l'aide de RDkit de 3,095. Il présente 4 groupements donneurs de liaison hydrogène, 1 groupement accepteur de liaison hydrogène et 3 liaisons permettant une rotation libre du ligand.

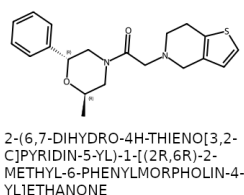


FIGURE 46 – Structure du composé 1. Réalisé avec RDkit.

Il présente une interaction avec mPGES-1, type donneur de liaison d'hydrogène avec le groupement amide et une autre de type accepteur de liaison hydrogène avec le groupement pyridine et avec la sérine 127A respectivement. Il existe également des interactions de Van der Waals avec la phénylalanine 44B et la leucine 39B pour le noyau phényle à une extrémité du ligand et avec la valine 128A, la leucine 135A et la thréonine 131A pour le noyau thiophène à l'autre extrémité. Le composé 1 est basique mais c'est la forme ammonium qui ressort car elle présente une meilleure score de docking (maximisant les interactions électrostatiques) cependant dans le site actif le microenvironnement peut favoriser la protonation de l'azote en dépit du fait que le pKa du composé est basique.

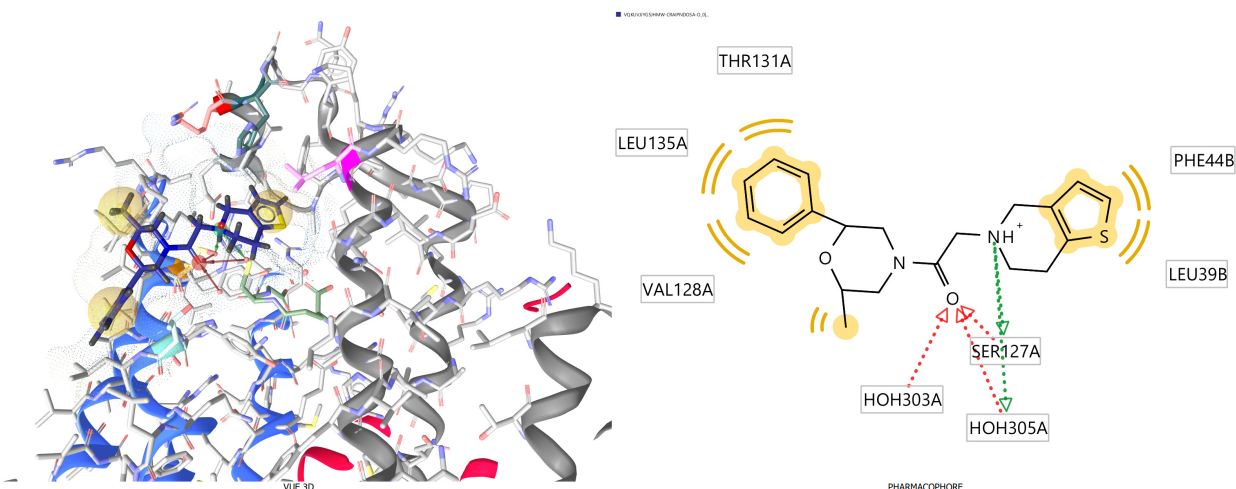


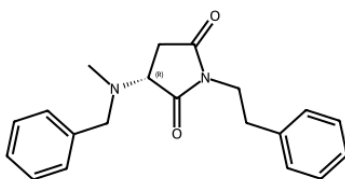
FIGURE 47 – À gauche, composé 1 après étude d'amarrage au sein de mPGES-1, avec le Gluthation en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose ; à droite, le pharmacophore en représentation 2D.

Ce composé 2 est le :

(3R)-3-[benzyl(méthyl)amino]-1-(2-phényléthyl)pyrrolidine-2,5-dione

Ce deuxième composé présente un poids moléculaire calculé de 322,408 g/mol, un LogP calculé de 2,489, un ensemble de 2 donneurs de liaison hydrogène, 1 accepteur de liaison hydrogène et 6 liaisons permettant une rotation libre du ligand.

Composé 2



(3R)-3-[benzyl(méthyl)amino]-1-(2-phényléthyl)pyrrolidine-2,5-dione

FIGURE 48 – Structure du composé 2. Réalisé avec RDkit.

Les interactions du composé 2 avec mPGES-1 sont : donneur de liaison hydrogène avec la sérine 127A et accepteurs de liaison hydrogène avec la thréonine 131A, ainsi que des interactions de Van der Waals avec la leucine 39B, la phénylalanine 44B pour un cycle de la benzylamine et avec, la leucine 135A et la valine 128A pour le cycle phényle.

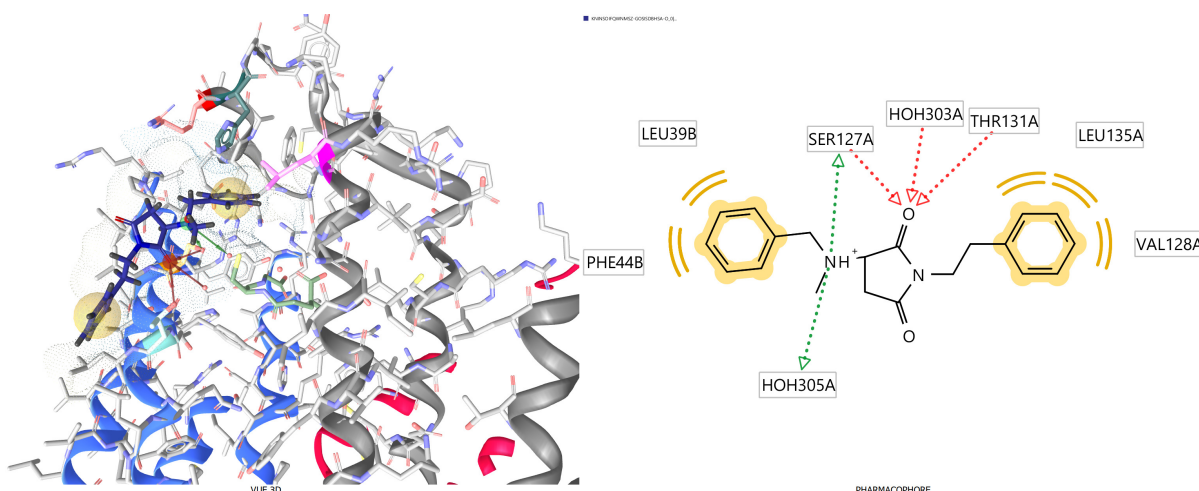


FIGURE 49 – À gauche, composé 2 issu de la base de données ZINC après étude d'amarrage moléculaire au sein de mPGES-1, avec le Gluthation en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose ; à droite, le pharmacophore en représentation 2D, le tout extrait de LigandScout®

Au sein de la base de données COCONUT, 2 composés ont également été sélectionnés :

— Ce composé 3 est la :N,N'-dibenzylpropane-1,3-diamine

Il a un poids moléculaire calculé de 254,37 g/mol, un LogP calculé de 2,6, un ensemble de 2 donneurs et accepteurs de liaison hydrogène ainsi que 8 liaisons permettant une rotation libre du ligand.

L'interaction entre le composé et mPGES-1 se compose de donneurs de liaison hydrogène avec la tyrosine 131A et la sérine 127A chacune avec un des deux azotes du composé 3 et de liaisons de Van der Waals avec l'isoleucine 32B avec un groupement phényle à une extrémité, la leucine 39B et la phénylalanine 44B avec le groupement phényle à l'autre extrémité.

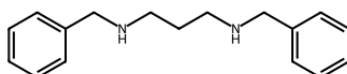
Dans la littérature ce composé est décrit présentant une activité antitumorale [224]

ou antivirale [225]

] notamment

pour le virus de l'hépatite C. Également, plusieurs brevets sont déposés pour son activité antipaludique.

Composé 3



N,N'-dibenzylpropane-1,3-diamine

FIGURE 50 – Représentation du composé 3 en deux dimensions. Réalisé avec RDkit.

Représentation du composé 3 issu de la base de données COCONUT

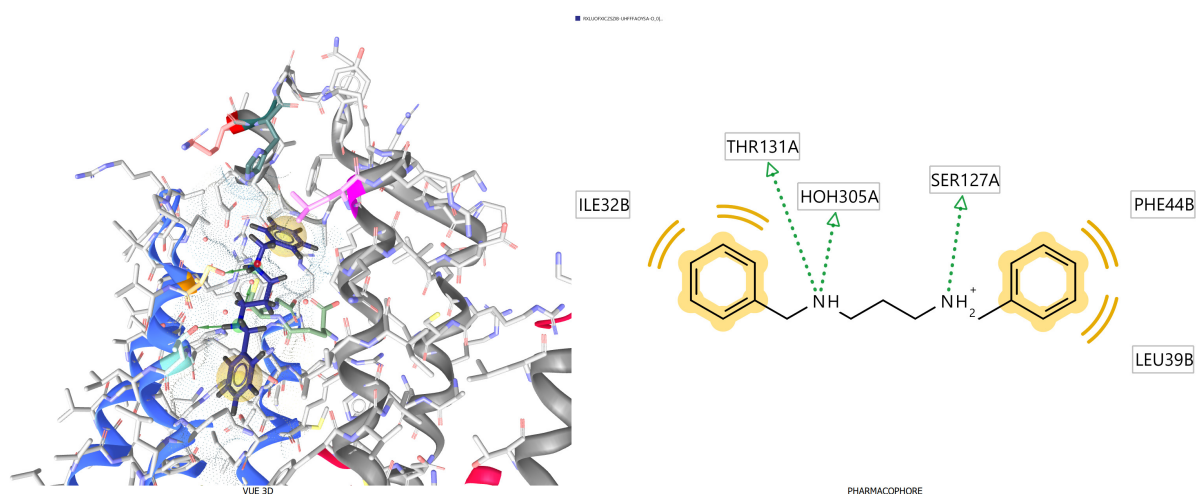


FIGURE 51 – À gauche, composé 3 issu de la base de données COCONUT après étude d'amarage moléculaire au sein de mPGES-1, avec le Gluthation en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose ; à droite, le pharmacophore en représentation 2D, le tout extrait de LigandScout®

— Le composé 4 est le : N-[(2S,3S,4S)-3-hydroxy-4-[(2-hydroxyphényl)méthylamino]oxolan-2-yl]méthyl]thiophene-2-carboxamide

Son poids moléculaire calculé est de 348,4 g/mol, son LogP calculé est de 1,6, il compatibilise 4 donneurs de liaison hydrogène et 6 accepteurs de liaison hydrogène ainsi que 6 liaisons permettant une rotation libre du ligand.

Composé 4

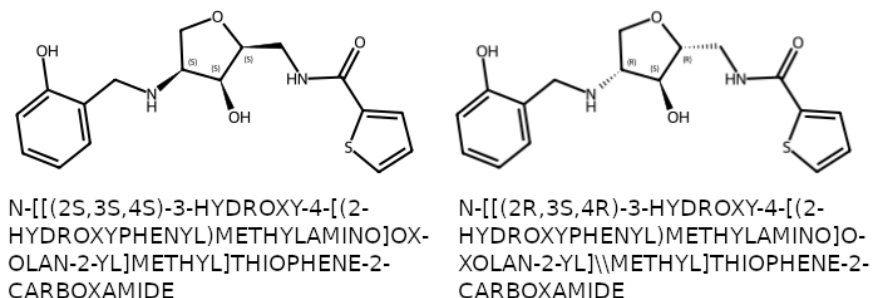


FIGURE 52 – Représentation du composé 4 en deux dimensions, avec son diastéréoisomère. Réalisé avec RDkit.

Un composé similaire mais présentant une stéréochimie différente (N-[[[(2R,3S,4R)-3-hydroxy-4-[(2-hydroxyphényl)méthylamino]oxolan-2-yl]méthyl]thiophene-2-carboxamide) a été testé contre la 12-lipoxygénase humaine mais s'est révélé inactif contre cette cible.

Le composé 4 présente au niveau du phénol des interactions de Van der Waals avec la leucine 39B et la phénylalanine 44B et une interaction de donneur de liaison hydrogène avec la tyrosine 53B. Il présente aussi des interactions de Van der Waals entre le groupement thiophène et l'alanine 31B, l'isoleucine 32B, la tyrosine 130A et la tyrosine 28B. Le composé 4 présente aussi des interactions de donneurs de liaison hydrogène avec la tyrosine 127A et la sérine 131A.

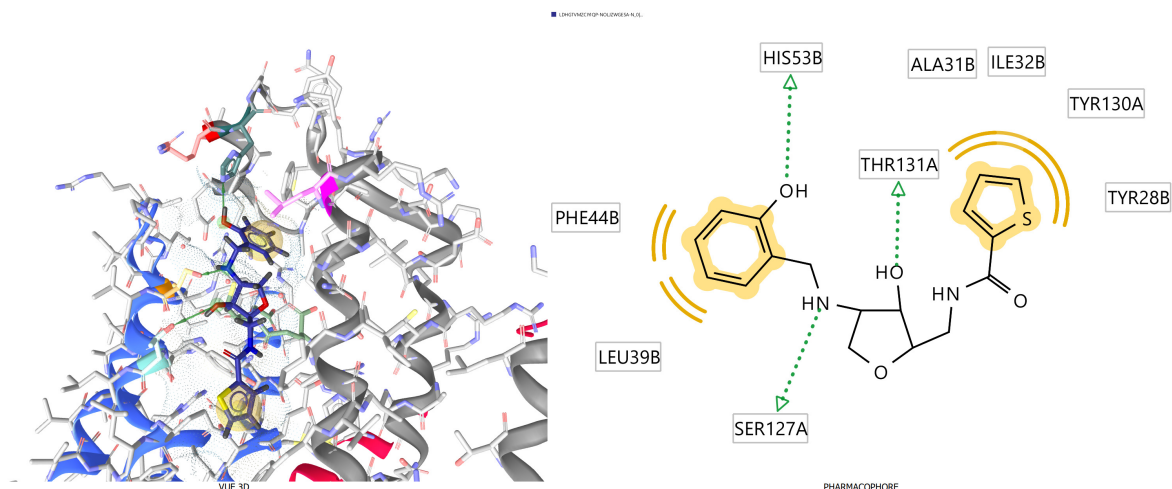


FIGURE 53 – À gauche, le composé 4 après étude d'amarrage moléculaire au sein de mPGES-1, avec le Gluthation en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose ; à droite, le pharmacophore en représentation 2D, extrait de LigandScout®

Parmi les composés de la base de données UNDP, 3 composés ont été sélectionnés :

- Le composé 5 est l'acide ((4S)4-((2S,3S,5S)5-éthyl-3-méthylloxolan-2-yl)-4-hydroxy-N-((2R,3S)1-hydroxy-3-méthylpentan-2-yl)pent-2-énimidique) (figure 55).

Composé 5

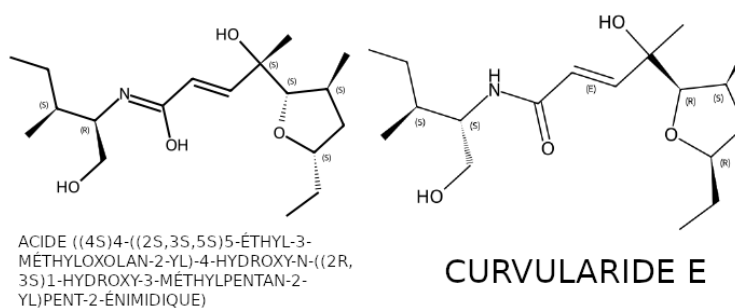


FIGURE 54 – Représentation du composé 5 en deux dimensions, avec sa comparaison. Réalisé avec RDkit.

Il présente une similitude structurale importante avec la curvularide E [226] issue de *Curvularia geniculata* mais avec une stéréochimie différente et également une

forme tautomère de la fonction amide en un groupement acide imidique.

Concernant les interactions supposées entre ce ligand et mPGES-1, il présente des interactions de donneurs de liaison hydrogène avec la sérine 127A, la thréonine 131A et l'histidine 53B et les différentes fonctions hydroxyle du composé 5. Également, une interaction d'accepteur de liaison hydrogène entre la thréonine 131A et l'oxygène du groupement oxolane est présente; de même entre celui-ci et des molécules d'eau. Enfin, le pharmacophore présente aussi de nombreuses interactions hydrophobes entre les chaînes aliphatiques et la phénylalanine 44B, la leucine 39B, l'alanine 123A, l'alanine 31B, l'alanine 138A, la leucine 135A, la tyrosine 130A et l'isoleucine 32B.

Représentation du composé 5 issu de la base de données UNDP

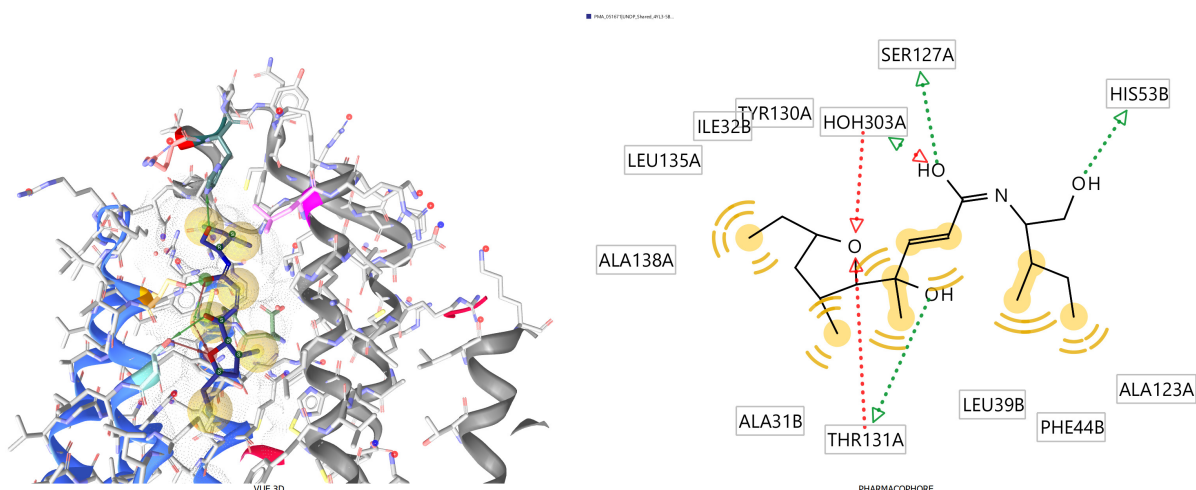
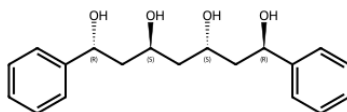


FIGURE 55 – À gauche, composé 5 issu de la base de données UNDP après étude d'amarrage moléculaire au sein de mPGES-1, avec le Gluthation en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose; à droite, le pharmacophore en représentation 2D, le tout extrait de LigandScout®

- Le composé 6 est la Diospongine C extrait de *Dioscorea spongiosa* L., sa dénomination IUPAC est (1R,3S,5S,7R)-1,7-diphénylheptane-1,3,5,7-tetrol.

Ce composé 6 a un poids moléculaire estimé de 316,4 g/mol, un logP estimé de 1,8, un ensemble de 4 donneurs de liaison hydrogène, un total de 4 accepteurs de liaison hydrogène, et 8 liaisons permettant une rotation libre du ligand. Ce composé 6 (57), présente des interactions de Van der Waals au niveau des cycles phényles, avec l'isoleucine 32B pour l'un et avec phénylalanine 44B et la Leucine 39B pour l'autre. Les différentes fonctions hydroxyle présentent des interactions donneurs de liaison hydrogène avec la sérine 127A, la thréonine 131A et l'histidine 53B et aussi accepteurs de liaison hydrogène avec la sérine 127A et la thréonine 131A.

Composé 6



(1R,3S,5S,7R)-1,7-diphenylheptane-1,3,5,7-tetrol

FIGURE 56 – Représentation du composé 6 en deux dimensions. Réalisé avec RDkit.

Représentation du composé 6 issu de la base de données UNDP

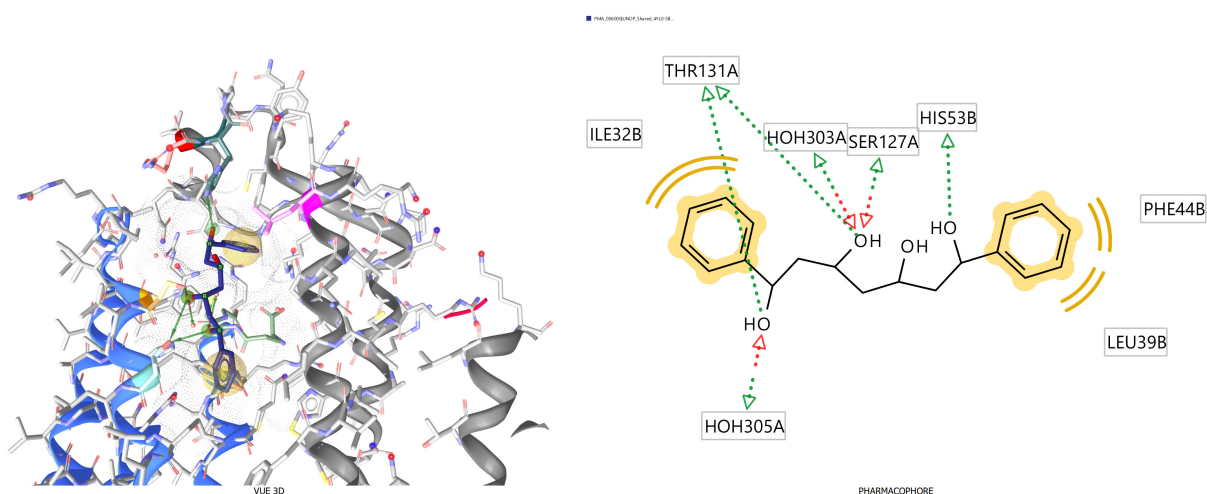
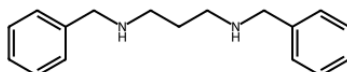


FIGURE 57 – À gauche, composé 6 issue de la base de donnée UNDP après étude d'amarrage moléculaire au sein de mPGES-1, avec le Gluthation en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose ; à droite, le pharmacophore en représentation 2D, le tout extrait de LigandScout®

- Ce composé 7, est le (1R,3S,5S)-1,7-diphenylheptane-1,3,5-triol [227] d'après la dénomination IUPAC. C'est un composé qui est synthétisé à la suite d'une hydrogénation par transfert asymétrique [228], qui présente un poids moléculaire calculé de 300,4 g/mol ; un LogP calculé de 2,9 ; un ensemble de 3 donneurs de liaison hydrogène et 3 accepteurs de liaison hydrogène ; et un ensemble de 8 liaisons permettant une rotation libre du ligand. D'après

Composé 7



N,N'-dibenzylpropane-1,3-diamine

FIGURE 58 – Représentation du composé 7 en deux dimensions. Réalisé avec RDkit.

J-global(service d'information scientifique de la "JST"(agence de financement de la recherche japonaise))[229]] ce composé ne fait actuellement l'objet d'aucun brevet ou projet de recherche. L'interaction potentielle de ce composé avec mPGES-1, se fait par des liaisons de Van der Waals au niveau des extrémités du composé, par les phényles en interaction avec la leucine 39B et la phénylalanine 44B d'une part et l'isoleucine 32B d'autre part. Il est aussi accepteur de liaison hydrogène avec la sérine 127A au niveau des hydroxyles et donneur de liaison hydrogène avec la même sérine et la thréonine 131A de mPGES-1 et aussi en interaction avec des molécules d'eau comme le montre la figure 59.

Ce composé, visible en figure 59, présente une grande similarité avec le composé précédent, le composé 6. En plus de cette similarité, les interactions retrouvées sont aussi similaires, hormis avec l'histidine 53B.

Représentation du composé 7 issu de la base de données UNDP

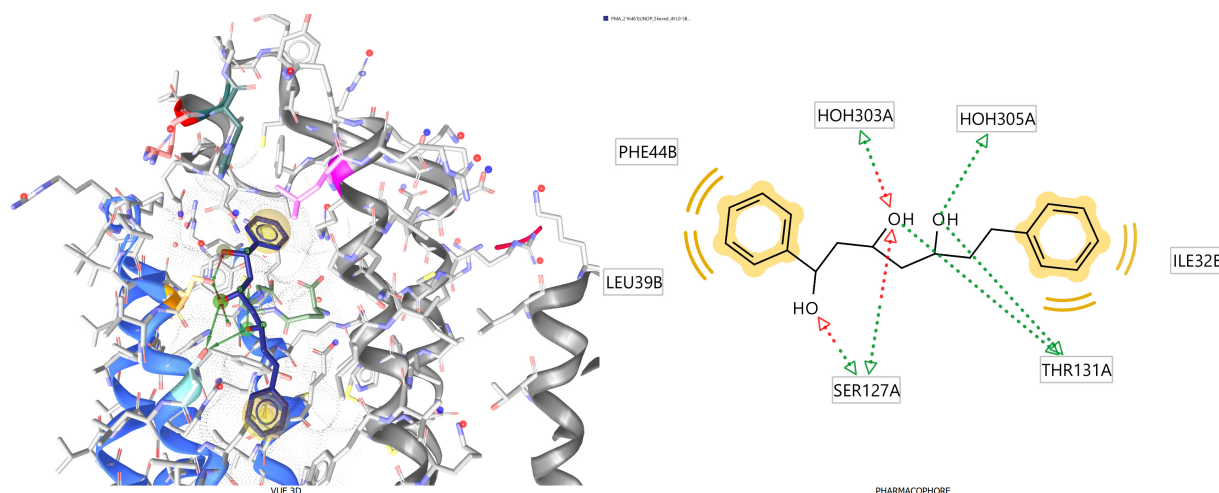
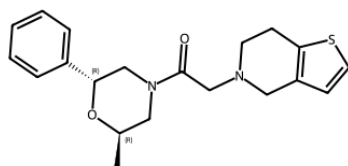


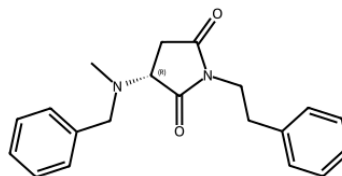
FIGURE 59 – À gauche, composé 7 issu de la base de données UNDP après étude d'amarrage moléculaire au sein de mPGES-1, avec le Glutathion en vert, la sérine A127 en orange, la thréonine A131 en cyan, l'histidine B53 en vert, l'arginine B52 en rouge et la leucine B39 en rose ; à droite, le pharmacophore en représentation 2D, le tout extrait de LigandScout®

Les composés sélectionnés parmi les différentes bases de données présentent une certaine diversité structurale, particulièrement au niveau des espaceurs des composés entre les 2 cycles à leurs extrémités hormis le composé 5. Ces espaceurs présentent une flexibilité importante car ils présentent peu de liaisons rigides en leur sein. Les différentes poses d'amarrage moléculaire montrent que les composés s'inscrivent dans les deux poches du site actif ; une pour les composés issus de la base de donnée ZINC et l'autre pour les bases de données COCONUT et UNDP.

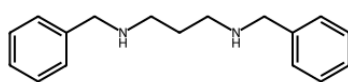
Composés vue d'ensemble



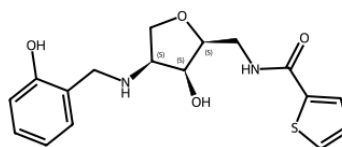
COMPOSE 1



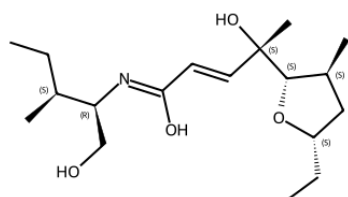
COMPOSE 2



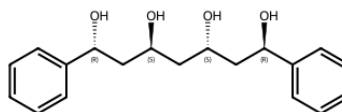
COMPOSE 3



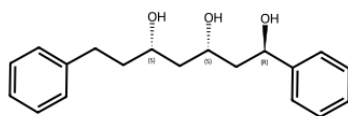
COMPOSE 4



COMPOSE 5



COMPOSE 6



COMPOSE 7

3.4 Méthodes

3.4.1 Obtention du jeu de données de validation - les inhibiteurs

Le jeu de données d'inhibiteurs actifs a été obtenu par interrogation de la base de données chEMBL (version chEMBL28), en ne sélectionnant que les composés possédant une CI_{50} inférieure à 2 μ M. Les données potentiellement dupliquées ou possiblement mal interprétées ont été également retirées du jeu de données. Le jeu de données a ensuite été traité à l'aide de RDkit (version 2020.09.1) pour établir un jeu de données de composés actifs directement importés dans LigandScout[®] (version 4.2) au format .ldb.

3.4.2 Obtention du jeu de données de validation - les decoys

Les decoys ont été obtenus à partir du jeu de données DUD-E. La construction du jeu de données decoys personnalisé s'est faite en passant l'ensemble du jeu de données d'inhibiteurs actifs au serveur pour la génération du jeu de données decoys. Brièvement, les decoys ont été estimés à partir du poids moléculaire, du coefficient de partition eau/octanol (miLogP), des accepteurs et donneurs de liaison hydrogène, des liaisons permettant une rotation libre de la molécule et de la charge nette. L'état de protonation des ligands a été généré à partir de Schrödinger Epik en passant les arguments '-ph 7.0 -pht 1.0 -tp 0.20'. Les propriétés moléculaires ont été calculées par Molinspiration's mib. À la suite, à partir de ZINC45 un protocole dynamique permet d'établir un ensemble de 3000 à 9000 decoys potentiels pour chaque état de référence d'un ligand. Puis, une empreinte ECFP4[230]

a été générée par la suite logicielle SciTegic. Plus particulièrement, une empreinte ECFP4 s'établit à partir des propriétés atomiques pour chaque atome de la molécule (nombre de voisins non hydrogène, l'ordre de liaison totale (excepté les liaisons aux hydrogènes), le numéro atomique, la charge partielle, la masse atomique, le nombre d'hydrogènes liés et si l'atome est inclus dans un cycle). Cet ensemble d'éléments est passé en entrée d'une fonction de hachage pour obtenir un identifiant unique pour chaque atome. Par itération, cette opération est répétée pour l'atome et ses voisins immédiats (le nombre d'itérations est marqué dans la signature par une valeur double du nombre d'itérations, 2 pour ECFP4, 3 pour ECFP6). Ensuite, les duplicats sont retirés. Puis cet ensemble d'identifiants est transformé en une suite de bits (généralement de 1024 bits de longueur). Les decoys générés ont ensuite été triés par leur score de Tanimoto selon la formule $T(A,B) = |A \cap B| / (|A| + |B| - |A \cap B|)$ à partir des empreintes (chaînes de bits) générées. Ainsi, le score de Tanimoto permet de classer les différents composés suivant la similarité de leurs empreintes par rapport au ligand. Il fluctue entre 0 et 1, 0 pour aucun bits commun entre les deux empreintes et 1 pour des empreintes identiques. Pour chaque ligand, le dernier quart le plus dissimilaire du ligand a été sélectionné. Les decoys ont ensuite été triés de nouveau en fonction de leur taux de duplication à travers les différents ligands puis assignés aux différents ligands à partir du ligand qui possède

le moins de decoys déjà assignés pour limiter les duplicats. Enfin, si disponible, un set de 50 decoys est tiré aléatoirement pour chaque ligand. Puis l'ensemble des fichiers présentant les decoys ont été assemblés en un fichier auquel les inhibiteurs issus du jeu de données de validation et présentant une CI_{50} supérieure à 3 μM ont été ajoutés. Le fichier a ensuite été converti au format ligand database(.ldb) à travers un script par LigandScout®.

Logigramme de construction du jeu de validation

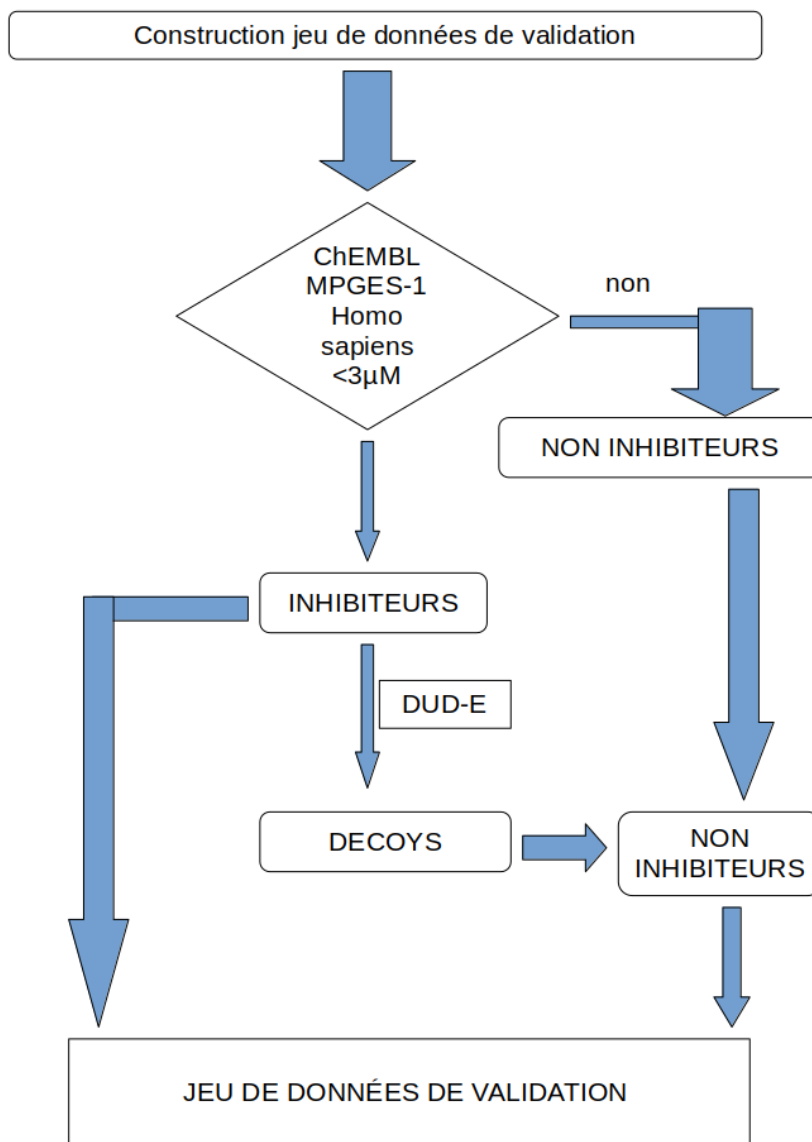


FIGURE 61 – Logigramme du jeu de données de validation à partir des inhibiteurs de mPGES-1 de ChEMBL étiquetés entre les inhibiteurs et les non inhibiteurs (non inhibiteurs connus et decoys).

3.4.3 Validation du pharmacophore

Prioritairement au criblage virtuel de grandes bases de données de composés chimiques, il est important de valider le pharmacophore établi par sa capacité à distinguer les inhibiteurs

actifs de mPGES-1 à d'autres molécules chimiquement proches, ne présentant pas cette activité. Par les activités de recherche, les bases de données comportent plus de données pour les composés portant une activités activatrice ou inhibitrice que de composés ne portant pas d'activité. Ces données sur les composés inactifs, parcellaires, ne permettent pas de maintenir un ratio inhibiteurs actifs/decoys important. La génération de composés "decoys" putatifs est nécessaire pour compléter les données expérimentales parcellaires. L'utilisation de decoys a commencé au début des années 2000 [195] à partir d'une sélection randomisée de composés. Les pratiques ont évolué pour aboutir à la sélection de molécules structurellement différentes mais présentant des propriétés chimiques proches des composés actifs. DUD-E [231] est un jeu de données couramment utilisé pour la génération de composés "decoys".

3.4.4 Exploration du jeu de données de validation

La visualisation et la réduction de dimension (UMAP) avant projection du jeu de données de validation comprenant les composés actifs et les "decoys", se sont faites à l'aide du paquet scikit et le calcul de descripteurs physico-chimiques (poids moléculaire, surface polaire topologique, nombre de liaisons permettant une rotation libre du ligand, nombre de donneurs de liaison hydrogène, nombre d'accepteurs de liaison hydrogène, LogP calculé) en utilisant rdkit, à l'aide d'un script.

3.4.5 Récupération des données pour construction du pharmacophore

La base de données Protein Data Bank (PDB) possède plusieurs complexes de mPGES-1 co-cristallisés avec un inhibiteur. Les fichiers PDB utilisés pour établir les modèles de pharmacophores sont d'une part 4YK5, 5TL9, 5T36, 5T37 et d'autre part 4YL0, 4YL1, 4YL3, 5BQH, 5BQI, 5BGQ.

3.4.6 Création des modèles de pharmacophores

La création et l'affinage des pharmacophores sont faites à l'aide du logiciel LigandScout® version 4.2. Le complexe biologique, inhibiteur-protéine est traité dans le logiciel afin de créer le pharmacophore dans un rayon de 7 Angströms autour de l'inhibiteur. Les volumes d'exclusion sont ajoutés afin de prendre en compte l'encombrement stérique, les volumes d'exclusion affectés aux molécules d'eau sont curés manuellement. Les pharmacophores sont ensuite interpolés deux à deux et les caractéristiques communes aux deux pharmacophores sont conservées. Pour cela des points d'ancrages communs aux deux pharmacophores sont utilisés : LEU39, ARG52, PHE44, VAL128, THR131, LEU132, LEU135.

3.4.7 Bibliothèques de composés

Plusieurs bibliothèques de composés ont été utilisées pour ces travaux. Une bibliothèque de composés issus de la base de données ZINC [232]. Une bibliothèque de composés issus de la base de données COCONUT [201]. une base de composés naturels et la base de données interne de l'unité SONAS. Chaque base de données, a fait l'objet de vérifications, vérification de la stéréochimie, modification de la 2D en 3D, minimisation énergétique. Par exemple pour la base de données COCONUT c'est 407300 composés qui ont été retenus.

3.4.8 Filtration physico-chimique

Les composés issus du criblage virtuel par pharmacophore sont filtrés par le nombre de donneurs de liaison hydrogène, le nombre d'accepteurs de liaison hydrogène, le LogP et du poids moléculaire (Tableau 2) à l'aide d'un script utilisant rdkit pour calculer ces descripteurs.

3.4.9 Minimisation énergétique

La minimisation énergétique des composés est importante pour la qualité des résultats des études d'amarrage moléculaire. Pour cela, la documentation du logiciel Gold recommande l'utilisation de "CSD Conformer Generation" version 2020.2. C'est ce qui a été utilisé pour minimiser l'ensemble des dérivés issus du criblage virtuel et après filtration.

3.4.10 Amarrage moléculaire

L'amarrage moléculaire s'effectue à l'aide du logiciel GOLD ("Genetic Optimisation for Ligand Docking") version GOLD © 2022 CCDC Software Ltd. La configuration suivante est appliquée :

FLAGS

internal_ligand_h_bonds = 1

flip_free_corners = 1

match_ring_templates = 0

flip_amide_bonds = 0

flip_planar_n = 1 flip_ring_NRR flip_ring_NHR

flip_pyramidal_n = 1

rotate_carboxylic_oh = flip

use_tordist = 1

postprocess_bonds = 1

rotatable_bond_override_file = DEFAULT

solvate_all = 1

TERMINATION

```
early_termination = 1
n_top_solutions = 3
rms_tolerance = 1.5
```

CONSTRAINTS

```
force_constraints = 0
```

COVALENT BONDING

```
covalent = 0
```

FITNESS FUNCTION SETTINGS

```
initial_virtual_pt_match_max = 3
relative_ligand_energy = 1
gold_fitfunc_path = plp
score_param_file = DEFAULT
```

WATER DATA

```
water 1 toggle spin 0 ID_HOH301
water 1 toggle spin 0 ID_HOH302
water 1 toggle spin 0 ID_HOH304
water 1 toggle spin 0 ID_HOH306
water 1 toggle spin 0 ID_HOH307
water 1 toggle spin 0 ID_HOH308
water 1 toggle spin 0 ID_HOH309
water 1 toggle spin 0 ID_HOH311
water 1 toggle spin 0 ID_HOH312
```

CONSTRAINTS

```
constraint protein_h_bond 10.0000 0.005000 1943
constraint protein_h_bond 10.0000 0.005000 3092
constraint protein_h_bond 10.0000 0.005000 2006
constraint protein_h_bond 15.0000 0.005000 3082 3080
constraint protein_h_bond 10.0000 0.005000 2802
interaction_restraint_weight = 50
constraint interaction_restraint 3 38 B ARG CD 1
constraint interaction_restraint 3 44 B PHE CD1 1
constraint interaction_restraint 3 39 B LEU CD2 1
```

```

constraint interaction_restraint 3 130 A TYR CE2 1
constraint interaction_restraint 3 32 B ILE CG1 1
constraint interaction_restraint 3 28 B TYR CD1 1
constraint interaction_restraint 3 128 A VAL CB 1
constraint interaction_restraint 3 125 A ILE CB 1

```

3.4.11 Nomenclature IUPAC

La nomenclature IUPAC des différents composés a été obtenue à partir du générateur "SMILES to IUPAC" sur (<https://app.syntelly.com/smiles2iupac>) [krasnovTransformerbasedArtificial] et vérifiée par la conversion inverse à l'aide d'OPSIN [loweChemicalNameStructure2011].

3.4.12 Programmation Python

La liste des paquets utilisés avec leur version ainsi que la version de Python utilisée sont disponibles en annexe.

4 Conclusion

Les médicaments anti-inflammatoires sont un besoin important de l'arsenal thérapeutique et sont très utilisés dans la prise en charge de l'inflammation aiguë et/ou chronique. Cependant les AINS et les AIS présentent chacun des effets indésirables (toxicité gastrique, toxicité rénale, hyperglycémie,...) notamment lors d'usages à long terme. C'est pourquoi le traitement de l'inflammation chronique est un enjeu important et la recherche et développement de nouveaux médicaments anti-inflammatoires reste un besoin. De nouvelles enzymes susceptibles de limiter l'inflammation sont ciblées. En ce sens, mPGES-1 est une cible intéressante. Elle intervient en aval des COX dans la cascade enzymatique des prostaglandines, susceptibles de présenter donc moins d'effets indésirables, ciblant spécifiquement la PGE2, principale prostaglandine pro-inflammatoire. C'est pourquoi elle fait l'objet depuis plusieurs années de recherches intensives aboutissant au développement de plusieurs classes d'inhibiteurs de mPGES-1, jusqu'au développement de plusieurs essais cliniques [235]. Cependant la toxicité des composés n'a pas permis d'aboutir à la mise sur le marché d'inhibiteurs de mPGES-1. La découverte de nouveaux inhibiteurs reste donc un besoin important. Pour cela, les composés d'origine naturelle présentent un intérêt car ils impliquent généralement moins d'effets indésirables que les composés synthétiques et sont à l'origine de la majorité des médicaments sur le marché. Pour cela, à partir des complexes co-cristallisés disponibles sur la PDB, plusieurs pharmacophores ont été développés et évalués. Ceux présentant les meilleures performances ont été utilisés pour filtrer plusieurs bases de données. Les composés présentant la plus grande compatibilité avec ces critères furent ensuite l'objet d'une étude d'amarrage moléculaire, puis le résultat de cette étude a été analysé manuellement. Le résultat est la sélection de plusieurs composés présentant une variabilité structurale homogène et susceptibles d'être de nouveaux inhibiteurs de mPGES-1. Ces différents composés s'établissent dans les deux sous-poches du site actif de mPGES-1. La majorité présente un cycle à chaque extrémité séparé par un espaceur qui présente plus de variétés. Cependant, cette étude reste prospective et n'a fait l'objet d'aucune confirmation entre les résultats obtenus à travers les modélisations utilisées et une corrélation biologique. Elle présente des résultats intéressants qui pourraient être enrichis d'une évaluation de la toxicité des composés trouvés car c'est ce qui a auparavant déjà limité le développement des inhibiteurs de mPGES-1. Également, une étude plus ambitieuse pourrait compléter ces résultats par des travaux de dynamique moléculaire permettant d'évaluer les interactions dans le temps entre mPGES-1 et un ou des ligands, pour pouvoir pondérer l'importance des différentes caractéristiques au sein des pharmacophores développés. Au delà de la poursuite de cette étude, ces résultats avec plusieurs nouveaux composés découverts pourraient aussi servir de base après fragmentation pour respecter la règle des "3", enrichir une librairie de fragments pour une étude de découvertes de médicaments basés sur les

fragments [236].

].

5 Bibliographie

.1 Annexe 1 : Représentation de composés chimiques pour l'informatique

Plusieurs formats sont disponibles pour la représentation de molécules chimiques pour la chimoinformatique. Parmi eux on trouve notamment les SMILES, les InChI, et le format MDL molfile. Ces différents formats coexistent car ils présentent chacun des avantages et inconvénients :

- InChI : Pour "INternational CHemical Identifier" consiste en une description linéaire et unique. Difficilement lisible pour l'homme, ce format sépare les liaisons et les atomes ne permettant pas une lecture aisée de la molécule. Il ne comporte pas les coordonnées des atomes. Comme c'est un format linéaire, la taille du fichier est petite. Exemple :
InChI=1S/C17H21NO4/c1-18-12-8-9-13(18)15(17(20)21-2)14(10-12)22-16(19)11-6-4-3-5-7-11/h3-7,12-15H,8-10H2,1-2H3/t12-,13+,14-,15+/m0/s1
- Smiles : Pour "Simplified Molecular Input Line Entry Specification" consiste aussi en une description linéaire, cependant pas toujours unique à moins qu'une forme canonique soit utilisée. Ce format décrit la molécule sous forme de graphe où les ramifications sont exprimées entre parenthèse et les liaisons simples et les hydrogènes sont implicites. C'est un format plus facilement lisible pour l'homme comparé au format InChI. Comme c'est aussi un format linéaire, la taille du fichier est petite. Il ne comporte pas non plus les coordonnées atomiques.

Exemple : C(C1=CC=CC=C1)(=O)O[C@@H]1[C@@H]([C@H]2CC[C@@H](C1)N2C)C(=O)OC

- Molfile : Ce format est non linéaire. C'est à dire qu'il est composé de plusieurs blocs. Un premier bloc d'entêtes avec le nom de la molécule, le programme qui a écrit le fichier et une ligne de commentaires, puis le bloc principal appelé table de connexion, divisé en deux parties. La première partie, le bloc "atom", avec un atome par ligne et ses coordonnées. La deuxième partie est le bloc de liaisons qui contient les connexions entre atomes. Enfin, le dernier bloc est un bloc de terminaison marquant la fin du fichier. C'est le format le plus volumineux parmi les 3 présentés.

Un exemple de fichier :

Représentation du format MDL Molfile

```

RDKit      20
22 24 0 0 0 0 0 0 0 0 0999 V2000
-4.7734 1.7567 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-5.9869 2.6383 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-5.8301 4.1301 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-7.0436 5.0118 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-8.4140 4.4017 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-8.5708 2.9099 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-7.3572 2.0282 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-4.9302 0.2649 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-3.4031 2.3668 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.1895 1.4851 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.8192 2.0952 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.3943 1.2135 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.2760 0.0000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.3943 -1.2135 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.0323 -0.7500 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.3463 -0.0067 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.0323 0.7500 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.7714 1.0486 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.6624 3.5870 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.8760 4.4687 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.7079 4.1971 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.8647 5.6889 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 2 1 0
2 3 2 0
3 4 1 0
4 5 2 0
5 6 1 0
6 7 2 0
1 8 2 0
1 9 1 0
10 9 1 6
10 11 1 0
11 12 1 0
12 13 1 1
13 14 1 0
15 14 1 1
15 16 1 0
15 17 1 0
17 18 1 0
11 19 1 6
19 20 2 0
19 21 1 0
21 22 1 0
7 2 1 0
16 10 1 0
17 12 1 0
M END

```

FIGURE 62 – Représentation de la même molécule que précédemment au format INCHI et SMILES, au format Molfile. Découpé en 3 blocs (en-tête, tables de connections, bloc de fin)

.2 Annexe 2 : Exemples de pharmacophores non sélectionnés

Représentation de pharmacophores non sélectionnés pour le criblage virtuel avec leurs statistiques

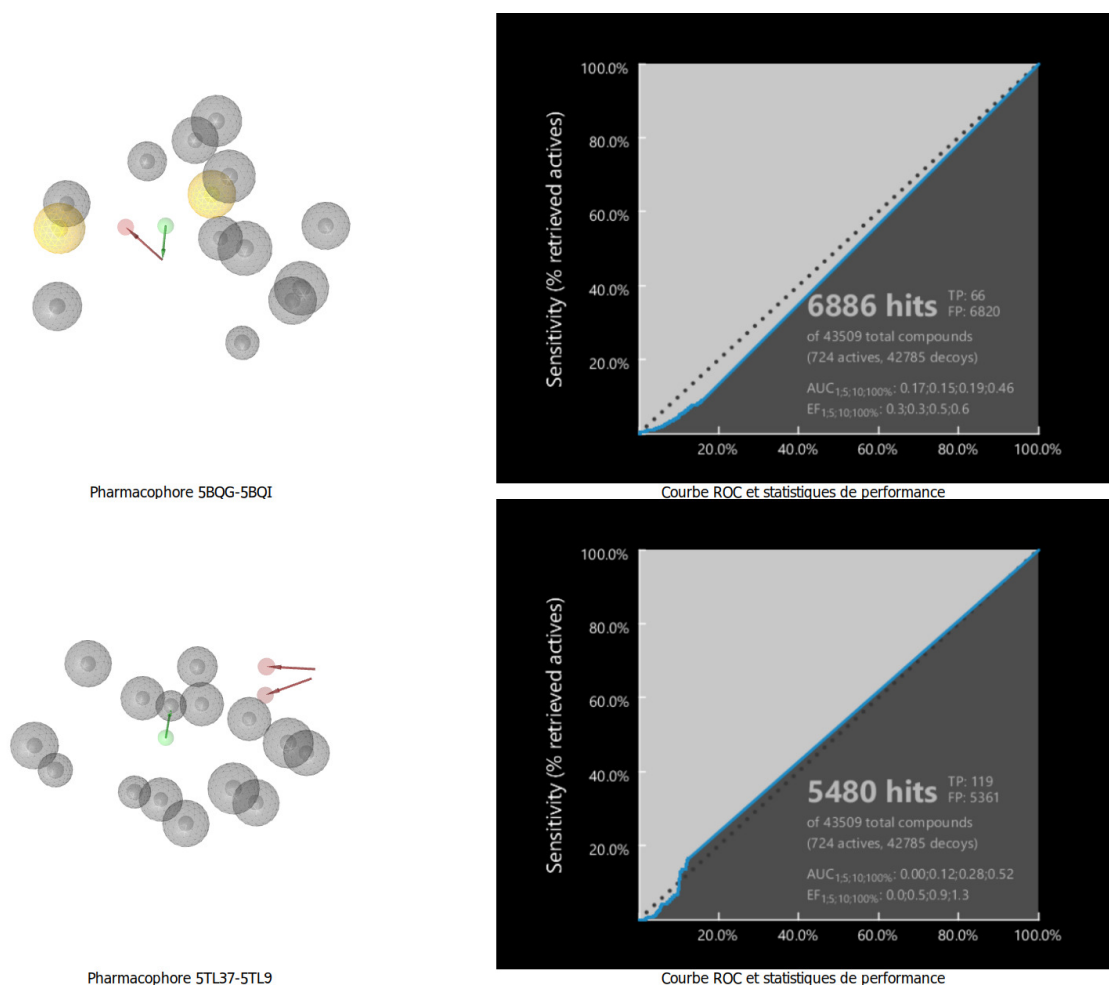


FIGURE 63 – Représentation de pharmacophores non sélectionnés pour le criblage virtuel avec leurs statistiques. À gauche le pharmacophore et à droite la courbe ROC avec les statistiques de performances.

.3 Annexe 3 : Exemples de poses d'amarrage moléculaire non retenues

Représentation de poses d'amarrage moléculaire non sélectionnées

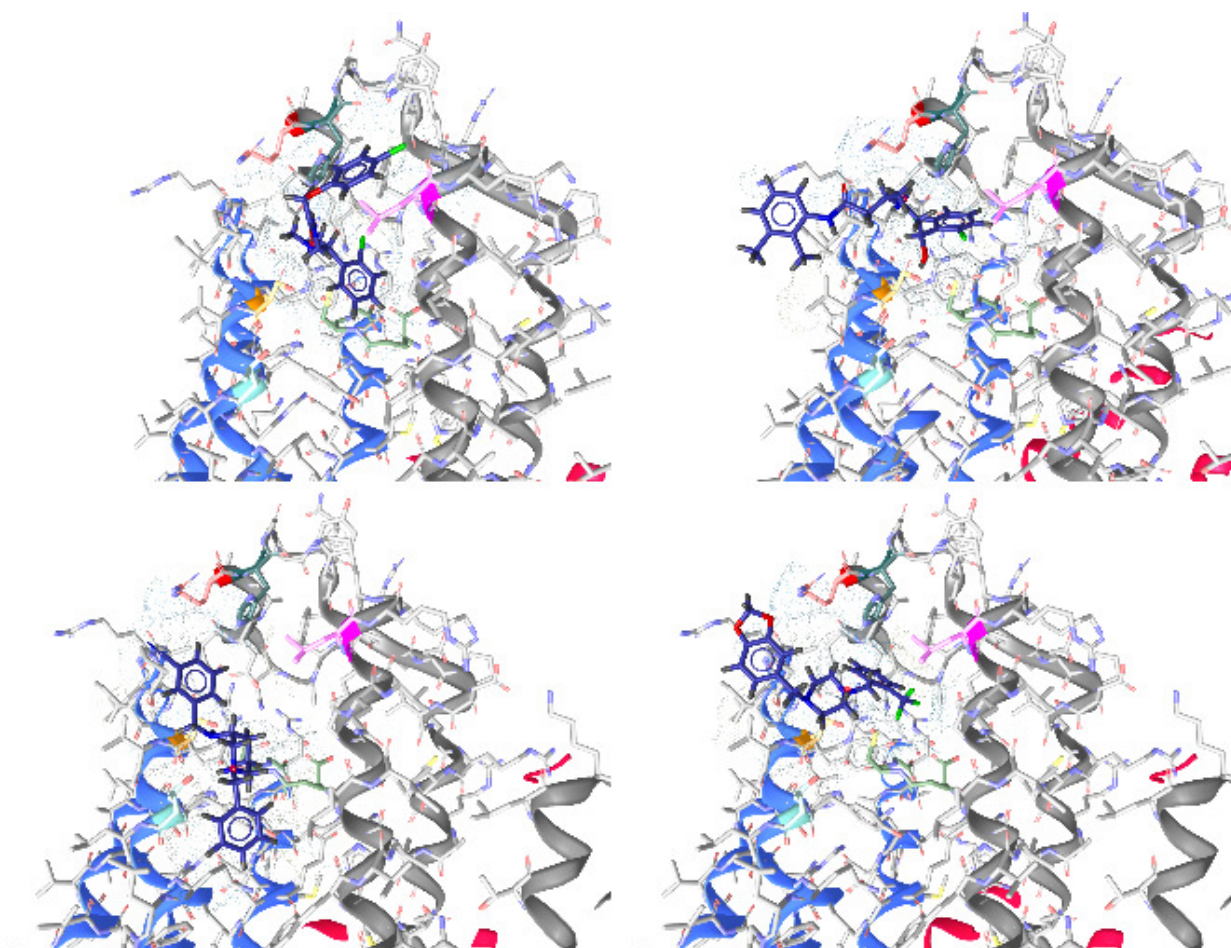


FIGURE 64 – *Représentation de poses d'amarrage moléculaire non sélectionnées et retirées lors du curage manuel*

Annexe 4 : Environnement Python utilisé lors de ces travaux

Tableau 3: Environnement Python, liste des paquets

Name	Version	Build	Channel
_libgcc_mutex	0.1	conda_forge	conda-forge
_openmp_mutex	4.5	1_gnu	conda-forge
bleach	3.3.0	pypi_0	pypi
bzip2	1.0.8	h7f98852_4	conda-forge
ca-certificates	2021.5.30	ha878542_0	conda-forge
cairo	1.16.0	h18b612c_1001	conda-forge
certifi	2021.5.30	py37h89c1867_0	conda-forge
cffi	1.14.5	pypi_0	pypi
chardet	4.0.0	pypi_0	pypi
chembl-webresource-client	0.10.3	pypi_0	pypi
colorama	0.4.4	pypi_0	pypi
cryptography	3.4.7	pypi_0	pypi
cycler	0.10.0	py_2	conda-forge
docutils	0.17.1	pypi_0	pypi
easydict	1.9	pypi_0	pypi
fontconfig	2.13.1	hba837de_1005	conda-forge

À suivre page suivante

Tableau 3: Environnement Python, liste des paquets (Suite)

Name	Version	Build	Channel
freetype	2.10.4	h0708190_1	conda-forge
gettext	0.19.8.1	h0b5b191_1005	conda-forge
glib	2.68.1	h9c3ff4c_0	conda-forge
glib-tools	2.68.1	h9c3ff4c_0	conda-forge
icu	58.2	hf484d3e_1000	conda-forge
idna	2.10	pypi_0	pypi
importlib-metadata	4.5.0	pypi_0	pypi
itsdangerous	1.1.0	pypi_0	pypi
jeepney	0.6.0	pypi_0	pypi
joblib	1.0.1	pyhd8ed1ab_0	conda-forge
jpeg	9d	h36c2ea0_0	conda-forge
keyring	23.0.1	pypi_0	pypi
kiwisolver	1.3.1	py37h2527ec5_1	conda-forge
lcms2	2.12	hddcbb42_0	conda-forge
ld_impl_linux-64	2.35.1	hea4e1c9_2	conda-forge
libblas	3.9.0	8_openblas	conda-forge
libboost	1.73.0	h3ff78a5_11	
libcblas	3.9.0	8_openblas	conda-forge
libffi	3.3	h58526e2_2	conda-forge
libgcc-ng	9.3.0	h2828fa1_19	conda-forge

À suivre page suivante

Tableau 3: Environnement Python, liste des paquets (Suite)

Name	Version	Build	Channel
libgfortran-ng	9.3.0	hff62375_19	conda-forge
libgfortran5	9.3.0	hff62375_19	conda-forge
libglib	2.68.1	h3e27bee_0	conda-forge
libgomp	9.3.0	h2828fa1_19	conda-forge
libiconv	1.16	h516909a_0	conda-forge
liblapack	3.9.0	8_openblas	conda-forge
libllvm10	10.0.1	he513fc3_3	conda-forge
libopenblas	0.3.12	pthread_h481-2303_1	conda-forge
libpng	1.6.37	h21135ba_2	conda-forge
libstdcxx-ng	9.3.0	h6de172a_19	conda-forge
libtiff	4.2.0	hdc55705_0	conda-forge
libuuid	2.32.1	h7f98852_1000	conda-forge
libwebp-base	1.2.0	h7f98852_2	conda-forge
libxcb	1.13	h7f98852_1003	conda-forge
libxml2	2.9.10	hb55368b_3	
llvmlite	0.36.0	py37h9d7f4d0_0	conda-forge
lz4-c	1.9.3	h9c3ff4c_0	conda-forge
matplotlib	0.1.9	pypi_0	pypi
matplotlib-base	3.4.2	py37hdd32ed1_0	conda-forge
natsort	7.1.1	pypi_0	pypi
ncurses	6.2	h58526e2_4	conda-forge

À suivre page suivante

Tableau 3: Environnement Python, liste des paquets (Suite)

Name	Version	Build	Channel
numba	0.53.1	py37hb11d6e1_1	conda-forge
numpy	1.20.2	py37h038b26d_0	conda-forge
olefile	0.46	pyh9f0ad1d_1	conda-forge
openjpeg	2.4.0	hf7af979_0	conda-forge
openssl	1.1.1k	h7f98852_0	conda-forge
packaging	20.9	pypi_0	pypi
pandas	1.2.4	py37h219a48f_0	conda-forge
patsy	0.5.1	py_0	conda-forge
pcre	8.44	he1b5a44_0	conda-forge
pillow	8.1.2	py37h4600e1f_1	conda-forge
pip	21.1	pyhd8ed1ab_0	conda-forge
pixman	0.38.0	h516909a_1003	conda-forge
pkginfo	1.7.0	pypi_0	pypi
pthread-stubs	0.4	h36c2ea0_1001	conda-forge
py-boost	1.73.0	py37ha9443f7_11	
pycparser	2.20	pypi_0	pypi
pygments	2.9.0	pypi_0	pypi
pyloco	0.0.139	pypi_0	pypi
pynndescent	0.5.2	pyh44b312d_0	conda-forge
pyparsing	2.4.7	pyh9f0ad1d_0	conda-forge
python	3.7.10	hffdb5ce_100- _cpython	conda-forge

À suivre page suivante

Tableau 3: Environnement Python, liste des paquets (Suite)

Name	Version	Build	Channel
python-dateutil	2.8.1	py_0	conda-forge
python_abi	3.7	1_cp37m	conda-forge
pytz	2021.1	pyhd8ed1ab_0	conda-forge
pyyaml	5.4.1	pypi_0	pypi
rdkit	2020.09.1.0	py37hd50e099_1	rdkit
readline	8.1	h46c0cb4_0	conda-forge
readme-renderer	29.0	pypi_0	pypi
requests	2.25.1	pypi_0	pypi
requests-cache	0.6.3	pypi_0	pypi
requests-toolbelt	0.9.1	pypi_0	pypi
rfc3986	1.5.0	pypi_0	pypi
scikit-learn	0.24.2	py37h18a542f_0	conda-forge
scipy	1.6.3	py37h29e03ee_0	conda-forge
seaborn	0.11.1	hd8ed1ab_1	conda-forge
seaborn-base	0.11.1	pyhd8ed1ab_1	conda-forge
secretstorage	3.3.1	pypi_0	pypi
setuptools	49.6.0	py37h89c1867_3	conda-forge
simplewebsoc- ketserver	0.1.1	pypi_0	pypi
six	1.15.0	pyh9f0ad1d_0	conda-forge
sqlite	3.35.5	h74cdb3f_0	conda-forge
statsmodels	0.12.2	py37h902c9e0_0	conda-forge

À suivre page suivante

Tableau 3: Environnement Python, liste des paquets (Suite)

Name	Version	Build	Channel
tbb	2020.2	h4bd325d_4	conda-forge
threadpoolctl	2.1.0	pyh5ca1d4c_0	conda-forge
tk	8.6.10	h21135ba_1	conda-forge
tornado	6.1	py37h5e8e339_1	conda-forge
tqdm	4.61.0	pypi_0	pypi
twine	3.4.1	pypi_0	pypi
typing	3.7.4.3	pypi_0	pypi
typing-extensions	3.10.0.0	pypi_0	pypi
umap-learn	0.5.1	py37h89c1867_1	conda-forge
url-normalize	1.4.3	pypi_0	pypi
urllib3	1.26.4	pypi_0	pypi
ushlex	0.99.1	pypi_0	pypi
webencodings	0.5.1	pypi_0	pypi
websocket-client	1.0.1	pypi_0	pypi
wheel	0.36.2	pyhd3deb0d_0	conda-forge
xorg-kbproto	1.0.7	h7f98852_1002	conda-forge
xorg-libice	1.0.10	h7f98852_0	conda-forge
xorg-libsm	1.2.3	hd9c2040_1000	conda-forge
xorg-libx11	1.7.0	h7f98852_0	conda-forge
xorg-libxau	1.0.9	h7f98852_0	conda-forge
xorg-libxdmcp	1.1.3	h7f98852_0	conda-forge

À suivre page suivante

Tableau 3: Environnement Python, liste des paquets (Suite)

Name	Version	Build	Channel
xorg-libxext	1.3.4	h7f98852_1	conda-forge
xorg-libxrender	0.9.10	h7f98852_1003	conda-forge
xorg-renderproto	0.11.1	h7f98852_1002	conda-forge
xorg-xextproto	7.3.0	h7f98852_1002	conda-forge
xorg-xproto	7.0.31	h7f98852_1007	conda-forge
xz	5.2.5	h516909a_1	conda-forge
zip	3.4.1	pypi_0	pypi
zlib	1.2.11	h516909a_1010	conda-forge
zstd	1.4.5	h9ceee32_0	

Développement de modèles pharmacophoriques et criblage virtuel de bibliothèques chimiques pour l'identification d'inhibiteurs de la prostaglandine E2 synthase 1 microsomale (mPGES-1)

RÉSUMÉ

L'inflammation aiguë est une réponse normale et nécessaire du système immunitaire de l'organisme. Parfois, ce processus perdure et/ou dysfonctionne. Il se transforme en un état inflammatoire chronique, associé au développement de nombreuses pathologies. Les traitements anti-inflammatoires actuels inhibent la synthèse des médiateurs inflammatoires lipidiques, notamment les prostaglandines. De plus, ces traitements anti-inflammatoires stéroïdiens et non-stéroïdiens sont associés à divers effets secondaires particulièrement lors de traitement à long terme. La mPGES-1 est une prostaglandine E2 synthase, dont l'expression est augmentée en cas d'inflammation. Elle fait l'objet d'études qui ont conduit à l'identification de plusieurs familles d'inhibiteurs sans aboutir à des essais cliniques satisfaisants permettant une mise sur le marché. Les complexes cristallographiques de mPGE-1 sont utilisés comme base à la construction de plusieurs modèles de pharmacophores. Ces modèles sont ensuite évalués à partir d'un jeu de données de validation. Les pharmacophores validés sont utilisés pour cribler plusieurs bases de données, sélectionnant plusieurs composés. Ces composés sont ensuite filtrés par des critères physicochimiques et ont fait l'objet d'études d'amarrage moléculaire. Ces études aboutissent à la sélection de 7 composés, potentiels inhibiteurs de mPGES-1.

Mots-clés : mPGES-1, pharmacophore, criblage virtuel, amarrage moléculaire, jeu de données

Development of pharmacophore models and virtual screening of chemical libraries for identification of microsomal prostaglandin E2 synthase 1 (mPGES-1) inhibitors

ABSTRACT

Acute inflammation is a normal response from the immune system for a return to normality but inflammation could be non-resolving leading to chronic inflammation associated with several diseases. Anti-inflammatory drugs, inhibit the synthesis of inflammatory lipid mediator including prostaglandins. Long-term treatment with those drugs is associated with several complications. mPGES-1 is a prostaglandin E2 synthase overexpressed in case of inflammation. Research on mPGES-1 has led to the discovery of several families of inhibitors without leading to clinical trials success. From a structure-based pharmacophore approach followed by a molecular docking study, several databases are screened for identification of new potential inhibitors of mPGES-1. Crystallographic complexes of mPGES-1 with inhibitors are used as basis for the construction of pharmacophore models. These pharmacophores are then evaluated using a validation dataset. Validated pharmacophores screen several databases, selecting many compounds then filtered with physico-chemicals criteria. The filtered results are passed by a molecular docking studies. Leading to the identification of seven compounds as potential inhibitors of mPGES-1.

Keywords : mPGES-1, pharmacophore, virtual screening, docking, dataset