



الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي و البحث العلمي

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE ET SCIENTIFIQUE

جامعة الاخوة منتوري قسنطينة1

Faculté des Sciences de la Nature et de la Vie كلية علوم الطبيعة و الحياة

قسم الكيمياء الحيوية و البيولوجيا الخلوية و الجزيئية

Département de Biochimie et Biologie Cellulaire et Moléculaire

Mémoire présenté en vue de l'obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Biochimie Appliquée

Intitulé :

ETUDE *IN SILICO* DE L'INHIBITION DE LA 3-CHYMOTRYPSIN
LIKE PROTEASE EN VUE DE COMBATTRE LE SYNDROME
RESPIRATOIRE AIGUE SEVERE (SRAS)

Présenté et soutenu par : - OULMI MANEL

SOUTENU LE : 14/07 / 2021

- ZOUBIRI MIRA

Jury d'évaluation :

Président du jury : Mr BENSEGUENI. A Pr. Université des Frères Mentouri Constantine1

Rapporteur : Mr CHIKHI. A Pr. Université des Frères Mentouri Constantine1

Examineur : Mr. MOKRANI E. H. Maitre de conférences B. UFM Constantine-1

Année universitaire

2020 – 2021

Résumé

ETUDE *IN SILICO* DE L'INHIBITION DE LA 3-CHYMOTRYPSIN LIKE PROTEASE EN VUE DE COMBATTRE LE SYNDROME RESPIRATOIRE AIGUE SEVERE (SRAS)

La pandémie actuelle de **COVID-19** a déclenché des efforts mondiaux pour l'identification rapide des vaccins et des traitements antiviraux spécifiques. **3CLpro** joue un rôle clé dans le cycle de vie de la réplication et de la maturation du coronavirus.

Le **Docking moléculaire** est une méthode souvent utilisée dans les phases initiales du processus de développement de nouveaux médicaments. Nous l'avons utilisé dans le but de concevoir de nouveaux composés à meilleur effet inhibiteur que « **Ebselen** » qui agit sur la cible **3CL pro**.

Le programme de docking moléculaire **Surflex** a été développé pour aider à la mise au point des molécules à activité thérapeutique. Avec un taux de valeurs de **RMSD** inférieures ou égales à **2 Å** de **61%** et un **coefficient de corrélation** de **0,601**, les performances du programme **Surflex** sont avérées. Ce test a été suivi d'une analyse visuelle de la superposition des ligands de référence et ceux obtenus par docking moléculaire.

Dans un deuxième temps, nous avons complété nos essais par une modélisation structurale en effectuant **63 mono substitutions** et **42 bi-substitutions**, parmi lesquelles quelques composés substitués ont eu des affinités améliorées. Le meilleur composé **n°47** parmi les mono substitués a atteint une affinité de **5.32 M⁻¹** et les meilleurs composés bi substitués **n°26** et **n°24** avec des affinités respectives de **6.21 M⁻¹** et **6.02M⁻¹**.

Enfin, l'étude des propriétés physicochimiques, pharmacocinétiques et de toxicité potentielle montrent que le composé **B24** a présenté un meilleur profil **ADMET**, par son métabolisme, et faible toxicité aigüe, contrairement au composé de référence et au composé qui a une plus grande valeur d'affinité **B26**. Nous le proposons donc comme nouvel inhibiteur éventuel de la **3CL pro**.

Mots clés : COVID-19, 3CLpro, Docking moléculaire, Inhibiteur, Surflex, RMSD, ADMET

Abstract

In silico study of the inhibition of «3-Chymotrypsine like protease» to combat severe acute respiratory syndrome (SRAS)

The current COVID-19 pandemic has sparked global efforts for the rapid identification of specific vaccines and antiviral treatments. **3CL pro** plays a key role in the coronavirus replication and maturation lifecycle.

Molecular docking is a method often used in the early stages of the development process of new drugs. We use it in order to design new compounds with better inhibitory effects than the compound «**ebsele**n» which acts on the target **3CL pro**. The **Surflex** molecular docking program was created to help the development of therapeutic molecules. With a rate of **RMSD** values less than or equal to **2 Å** of **61%** and a **correlation coefficient** of **0.601**. The performances of the **Surflex** program are proven. This test was followed by a visual analysis of the superposition of the reference ligands and those obtained by molecular docking.

In second step, we completed our tests with structural modeling by performing **63 mono substitutions** and **42 bi-substitutions**, among which some substituted compounds had improved affinities. The best compound **n° 47** among the mono substituted ones reached an affinity of

5.32 M⁻¹ and the best bi-substituted compounds **n° 26** and **n° 24** with respective affinities of **6.21 M⁻¹** and **6.02 M⁻¹**

Finally, the study of the physicochemical, pharmacokinetic and potential toxicity properties show that the compound **B24** presented a better **ADMET** profile, by its metabolism, and low acute toxicity, unlike the reference compound and the compound which has a greater d-value. affinity **B26**. We are therefore proposing it as a possible new inhibitor of **3CL pro**.

Keywords: COVID-19, 3CL pro, Molecular docking, Inhibitor, Surflex, RMSD, ADMET

المخلص

تطبيق الالتحام الجزيئي بواسطة برنامج Surflex لتطوير مثبطات جديدة لـ 3CLpro

اثارت جائحة COVID-19 الحالية جهودا عالمية للتعريف السريع بلقاحات معينة و علاجات مضادة للفيروسات. يلعب دورا رئيسيا في دورة حياة تكرر الفيروس التاجي و نضجه.

الالتحام الجزيئي هو طريقة تستخدم عادة في المراحل الاولى من عملية تطوير الادوية الجديدة. بهدف اكتشاف مركبات جديدة ذات تأثير مثبط افضل من مركب Ebselen الذي يؤثر على 3CL pro .

تم تطوير برنامج الالتحام الجزيئي Surflex للمساعدة في تطوير الجزيئات ذات النشاط العلاجي. حيث تحصلنا على معدل قيم RMSD اقل من او يساوي 2 Å بنسبة 61% و معامل الارتباط 0.601 حيث تم اثبات اداء برنامج Surflex.

تبع هذا الاختبار تحليل بصري لآلية التهام المركبات المرجعية و تلك التي تم الحصول عليها عن طريق الارساء الجزيئي.

ثانياً أكملنا اختباراتنا بالنمذجة الهيكلية عن طريق إجراء 63 استبدالاً أحادياً و 42 استبدالاً ثنائياً من بينها بعض المركبات المتبدلة التي حسنت التقارب. وصل أفضل مركب رقم 47 بين المركبات ذات الاستبدال الأحادي الى تقارب $5.32 M^{-1}$ و افضل المركبات ثنائية الاستبدال رقم 26 و رقم 24 مع تقاربات كل منهما $6.21 M^{-1}$ و $6.02 M^{-1}$ على الترتيب.

الكلمات الدالة : COVID-19 , 3CL pro , الالتحام الجزيئي المثبط , Surflex , RMSD , ADMET



Remerciement



Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.

Nous tenons tout d'abord, à remercier :

- notre encadreur Pr. CHIKHI. A, qui nous a accompagnés durant toute la période de cette étude avec beaucoup de patience et de savoir-faire. Sa gentillesse, sa disponibilité et ses qualités humaines nous ont été d'une aide précieuse.*

- Pr. BENSEGUENI A, pour avoir accepté la charge de présider le jury de ce mémoire.*

- Mr. MOKRANI E. H, pour avoir accepté de juger ce travail. Nous sommes honorés de la compter parmi les membres de ce jury.*

- À mes amis (es) de la promotion biochimie appliquée 2020*

- Nous tenons aussi à remercier tous les enseignants qui nous ont transmis notre savoir actuel tout au long de notre cursus.*

Dédicace



Je tiens à remercier en premier lieu ma reine celle qui a cru en moi dès le premier jour et m'a soutenu dans chaque décision que j'ai prise et m'a accordé sa confiance qui m'a poussé tout le long de mon chemin maman je te donne ma vie et c'est peu pour toi j'espère dieu te garde Pour nous j'espère te voir heureuse car c'est mon rêve ultime

Merci à mes sœurs Samia, Farah et Randa et leur petits : Timou Djana, Yanis Wassim, Mohamed et Adem merci à mes petites chouhou Nouza, Wissal, mima et Amina .et tout ma famille qui ont toujours cru en moi et en mes capacités à rêver encore plus fort.

Merci à mes proches et à mes amies Khawla et Yassmin, Je souhaite beaucoup de succès de réussite et de prospérité

Merci mon à mon amie et mon binôme : O.manel Pour son soutien moral, sa patience et sa compréhension tout long de ce projet.

Dédicace



A ma chère mère : leila

Tu m'inspires la force par un simple sourire, entre tes mains mes malheurs partent accompagnant un soupir, et en ta présence le bon Dieu ne cesse me bénir

A mon cher père : mourad

Comme un phare tu illumines et corriges mon cap ; qu'importent mes tombées, tu m'attrapes. A toi l'ami, le frère, le maître ; t'honorer est ma raison d'être.

A mes sœurs et mes frères : bouchra, anfel , alaa eddine et mouhamed amine

Avec qui je partage la joie et la peine, vos mot d'encouragement se comptent par centaines, grâce à vous en moi j'y crois et à présent sereine.

A mon Binôme : Z.mira

Je ne changerai pour rien au monde notre collaboration, merci pour ta sagesse, ton soutien, ta confiance ; j'ai trouvé en toi le pilier sur lequel je m'appuyais ; l'épaule sur lequel je pleurais. Ma confidente je te remercie pour ta sincérité, ton respect et ton amour

Manel

TABLES DES MATIERS

Résumé	
Liste des figures	
Liste des tableaux	
Liste des abréviations	
Introduction générale	1

Partie bibliographique

Chapitre 1 : généralités sur le COVID-19

Introduction	3
1. Mode de transmission	4
1.1. Contact et transmission de gouttelettes	4
1.2. Transmission aéroportée.....	5
1.3. Transmission fomite	5
2. Les tests et diagnostic	6
2.1. Les symptômes.....	7
2.2. La détection du matériel génétique du virus.....	7
2.3. La détection des protéines du virus.....	9
2.4. La détection de la réponse immunitaire contre le virus.....	11
3. Epidémiologie et statistique du COVID-19	13
3.1. l'échelle mondiale.....	13
3.2. l'échelle nationale.....	13

Chapitre 2 : la protéase principale (3CLpro)

1. Introduction	14
2. Les protéases virales	14
3. Les principales cibles corona virales	16
4. Main protéase (3CLpro, Mpro)	16
5. Structure du Mpro et le site actif	16
6. Les inhibiteurs de 3CLprotéase	17

Chapitre 3 : Le Docking Moléculaire

1. principe théorique	21
2. les types de Docking	22
3. Reconnaissance moléculaire	23
3.1. Interaction de Van Der Waals	23
3.2. Interactions ionique	24
3.3. Les liaisons hydrogène	25

3.4. L'effet hydrophobe	25
4. Outils du docking moléculaire	26
4.1. Le récepteur	26
4.2. Ligand 3D	26
5. Les principaux logiciels de docking	27
5.1. Surflex	28
6. Algorithmes de docking	28
6.1. Ajustement de fragment	28
6.2. Simulation par trajectoire	29
7. Les étapes du docking moléculaire	29
8. Applications	29

Chapitre 4 : Matériels et Méthodes

I. Matériel	30
1. Micro-ordinateur.....	30
2. Les banques de données	30
3. Les programmes.....	32
4. les sites internet	36
II. Méthodes	37
1. Docking moléculaire (D.M)	37
2. Tests d'évaluation de la fiabilité du docking.....	37
2.1 Coefficient de corrélation linéaire (r)	38
2.2 Le RMSD ou l'écart quadratique moyen.....	38
2.3. L'analyse visuelle	38
3. Le choix du complexe 4fkl.....	38
4. Préparation des molécules au docking moléculaire.....	40
4.1 Préparation de la cible	40
4.2 Construction des ligands	40
4.3 Les étapes du docking	40
5. Modification par substitution	42
6. Prédiction ADMET	42

Chapitre 5 : Résultats et discussion

1. Test d'évaluation de la fiabilité des programmes de docking	47
1.1.Le test RMSD	47
1.1.1. Le test RMSD par le programme Surflex.....	48
1.1.2. L'analyse visuelle des résultats du RMSD.....	48
1.2. Le test du coefficient de corrélation linéaire.....	50
1.2.1. L'analyse de corrélation des affinités obtenues par Surflex.....	51
2. Etude et recherche de nouveaux inhibiteurs de la 3CLpro	52
2.1. Etude des interactions enzyme-inhibiteur.....	53
2.1.1. Etude des interactions ebselen /3CLpro.....	53
2.2 Essais de modélisation moléculaire	54
2.2.1. Mono-substitution.....	54
2.2.2 Bi-substitution	59
3. Prédiction des propriétés ADMET	66
3.1. Propriétés Physico-chimique.....	67
3.2. Propriétés pharmacocinétiques	68
3.3. Les tests de toxicité.....	69
 CONCLUSION	 71
 LES REFERENCES BIBLIOGRAPHIQUES	
 ANNEXES	

Liste des figures

Figure 1 Transmission du SARS-Cov-2 via de petites particules en suspension dans l'air... 4	4
Figure 2 Transmission via des aérosols.....5	5
Figure 3 Transmission avec des surfaces contaminées.....6	6
Figure 4 la technique d'amplification de l'ADN par PCR.....8	8
Figure 5 Les différents résultats des bandes colorées possible.....10	10
Figure 6 Les étapes de développement des anticorps au stade de l'infection..... 10	10
Figure 7 les différents résultats de tests antigéniques rapides.....11	11
Figure 8 Répartition spatiale du COVID-19 par région.13	13
Figure 9 Représentation schématique de la production de particules virales par une cellule.....15	15
Figure 10 La structure générale de 3CL pro et le site actif.....17	17
Figure 11 Structure de l'inhibiteur Ebselen.....18	18
Figure 12 Structure de l'inhibiteur Bocéprivir.....18	18
Figure 13 Structure de l'inhibiteur GC-376.....19	19
Figure 14 Structure de l'inhibiteur α -kétoamides.....19	19
Figure 15 Structure de l'inhibiteur Tideglusib.....20	20
Figure 16 Structure de l'inhibiteur Carmofur.....20	20
Figure 17 Docking protéine-ligand.....22	22
Figure 18 Modèle clé-serrure.....22	22
Figure 19 Interaction de Van Der Waals.....23	23
Figure 20 La liaison ionique.....24	24
Figure 21 Liaison hydrogène.....24	24
Figure 22 Auto-organisation de molécules amphiphiles dans un solvant aqueux les zones .orange représentent les parties hydrophiles et les queues bleues les chaînes hydrophobes...25	25
Figure 23 Utilisation des différents logiciels de docking en 2010-2011.....27	27
Figure 24 Interface d'accueil de PDB.....31	31
Figure 25 Interface d'accueil de Pubchem.....32	32
Figure 26 Interface de logiciel Surfex.....32	32
Figure 27 Interface de logiciel Open Babel.....33	33
Figure 28 Interface de logiciel ArgusLab.....33	33
Figure 29 Interface de logiciel ViewerLite.....34	34
Figure 30 Interface de logiciel Biovia discovery studio visualizer.....35	35

Figure 31 Interface de l'excel.....	35
Figure 32 La page d'accueil du serveur ADMETlab.....	36
Figure 33 La page d'accueil du serveur PreADMET.....	37
Figure 34 Structure 3D du complexe 6Y2F issu de la PDB.....	40
Figure 35 Schéma explicatif de test d'Ames.....	45
Figure 36 Résultats en pourcentage des valeurs de RMSD obtenus par « surflex ».....	48
Figure 37 Courbe de corrélation entre LogIC50 et final-0 obtenus par Surflex.....	51
Figure 38 Structure de l'inhibiteur ebselen.....	53
Figure 39 Diagramme 2D des interactions de composé ebselen/3CLpro avec final0=3.51M-1.....	53
Figure 40 Représentation des radicaux R1.R2.R3 de la structure de base d'inhibiteur ebselen.....	54
Figure 41 le positionnement du composé M47 dans le site actif de 3CLpro.....	57
Figure 42 Diagramme 2D des interactions de composé M47/3CLpro.....	59
Figure 43 Représentation des radicaux R1.R2 du composé m47 issu de la mono-substitution.....	60
Figure 44 le positionnement du composé B24 dans le site actif de 3CLpro.....	62
Figure 45 Diagramme 2D des interactions de composé B24/3CLPro.....	64
Figure 46 le positionnement du composé B25 dans le site actif de 3CLpro.....	64
Figure 47 Diagramme 2D des interactions de composé B25/3CLpro.....	67

Liste des tableaux

Tableau 1 Présentation de différent résultats obtenir par teste antigénique et leur interprétation.....	11
Tableau 2 Les micro-ordinateurs utilisés et leurs caractéristiques.....	30
Tableau 3 : Code PDB des complexes 3clpro /inhibiteur avec leur résolution.....	
Tableau 4 Superposition des trois ligands des complexes (1cgh, 4j0y et 2who) et ceux calculés par docking moléculaire avec Surfex.....	49
Tableau 5 Données de l'analyse par la corrélation linéaire entre l'affinité obtenue par docking moléculaire (Surflex) et les logIC50.....	51
Tableau 6 les structures et les affinités des composés issus de la mon-substitution.....	55
Tableau 7 : les liaisons établies par le composé M47 avec le site actif.....	58
Tableau 8 les structures et les finals O des composés issus de la bi-substitution.....	60
Tableau 9 les liaisons hydrogènes établies par le composé B24 avec le site actif.....	63
Tableau 10 les liaisons hydrogènes établies par le composé B25 avec le site actif.....	65
Tableau 11 propriétés physico-chimique des composés obtenus à partir de serveur AdmetLab.....	67
Tableau 12 propriétés pharmacocinétique des composés obtenus à partir de serveur AdmetLab.....	68
Tableau 13 : le profil de toxicité relative des composés candidats obtenus à partir de serveur PreADMET.....	69

Liste des abréviations

COVID-19 : Corona Virus Disease 2019

SRAS-Cov-2 : Syndrome Respiratoire Aigu Sévère du Coronavirus 2

MERS : Syndrome Respiratoire du Moyen-Orient

3CLpro : 3C-Like protéase

Mpro : Main protéase

PL-pro : Papaine-Like-protéase

OMS : Organisation mondiale de la santé

ICIV : Comité International de Taxonomie des Virus

PCR: chain reaction de polymerisation

RT-PCR : la réaction de polymérisation en chaîne par transcription inverse

POC : Test Point-Of-Care

TROD : Tests rapides d'orientation de diagnostic

IgM : immunoglobulines M

IgA: immunoglobulines A

IgG : immunoglobulines G

RBD : Receptor Binding Domain

IC : Confiance intervalle

CPA : concentré plaquettaire d'Aphérèse

FDA : Food and Drug administration

VHC : Virus de l'hépatite C

VIH : Virus de l'immunodéficience Humaine

ADN : Acide DésoxyriboNucléique

ADNc : Acide DésoxyriboNucléique Complémentaire

ARN : Acide RiboNucléique

RMN : Spectroscopie Résonance Magnétique Nucléaire

PDB : Protein Data Bank (banque de donnée protéique)

3D : Trois dimensions

2D : Deux dimensions

Å : Angstrom

RMSD : Root Mean Square Deviation

IC₅₀ : Inhibitory concentration 50

r: Coefficient de corrélation

ADMET : Absorption, Distribution, Métabolisme, Elimination et Toxicité

BBB : Barrière Hémato-encéphalique

Da : Dalton

LogP : Coefficient de partage Eau/Octanol

hERG : human Ether-à-go-go-Related Gene

μM: Micro molaire

I - Partie
bibliographique

Introduction

La conception d'un nouveau médicament est un processus excessivement laborieux et dispendieux mais, qui a un rôle très important dans les sciences médicales actuelles pour la prévention et le traitement des maladies ainsi que la protection de la santé publique.

Les différentes approches pour la conception d'un médicament peuvent différer au niveau de la méthodologie utilisée pour parvenir à la conception d'un produit acceptable et commercialisable. Aujourd'hui avec le développement de l'outil informatique, l'identification et la mise au point d'une molécule possédant l'activité biologique recherchée, se base, de plus en plus sur des approches permettant la prédiction de l'affinité d'un ligand à l'égard d'une cible thérapeutique d'intérêt. Les résultats les plus prometteurs seront vérifiés *in vitro* puis *in vivo*, dans des cellules voire des organismes modèles [1] [2].

Dans le présent travail, nous nous sommes focalisées sur la recherche, par docking moléculaire de nouveaux inhibiteurs de SARS coronavirus 3-Chymotrypsine like protéase (3CLpro), cible thérapeutique pour le traitement de l'infection par SARS-COV2 [3].

Dans le premier temps, la performance du programme de docking moléculaire surflex a été évaluée via trois tests consistant en la détermination du RMSD, à une analyse visuelle de la superposition des ligands générés par le logiciel et ceux de référence et au calcul du coefficient de corrélation.

La molécule montrant le meilleur pouvoir inhibiteur nous permettra dans un premier temps d'étudier les différentes interactions formées avec la cible d'intérêt, et dans un deuxième temps de point de départ pour identifier et développer d'autres inhibiteurs meilleurs en procédant à la substitution de certains groupements.

Enfin, plusieurs critères physicochimiques, pharmacocinétiques et de toxicité potentielle des meilleurs inhibiteurs obtenus dans ce travail ont été simulés à l'aide des serveurs swissADME et PreADMET

Notre manuscrit est divisé en deux grandes parties :

1. Une première dite : partie bibliographique comprend trois chapitres dont ;

- Le premier est consacré aux généralités sur le COVID-19, le mode de transmission, les tests, le diagnostic, et quelques statistiques...
- Le deuxième : donne un aperçu sur notre cible « 3CLpro » et ses inhibiteurs.
- Le troisième chapitre : décrit le docking, ses algorithmes et ses logiciels

2. une seconde dite : partie pratique, comprend deux chapitres dont ;

- Le premier présente le matériel et les méthodes utilisés dans cette étude.
- Le second : expose les résultats et leurs interprétations suivis d'une conclusion générale

En fin quelques perspectives qui clôturent ce travail.

Chapitre 1

Récemment, à la fin de 2019, Wuhan, un centre commercial émergent de Chine, a connu une épidémie d'un nouveau coronavirus qui a tué plus de dix-huit cents et infecté plus de soixante-dix mille personnes au cours de cinquante premiers jours de l'épidémie. Le nouveau virus a été nommé nouveau coronavirus 2019 (2019-nCov) par les chercheurs chinois. Le comité international de taxonomie des virus (ICTV) a nommé le virus comme le SARS-Cov-2 et la maladie comme le COVID-19. [4, 5, 6]

Le 12 janvier 2020, la Commission nationale de la santé de Chine a publié des détails supplémentaires sur l'épidémie, suggérant une pneumonie virale. [7] Le 30 janvier, l'OMS a annoncé une urgence de santé publique de portée internationale (USPPI) pour l'épidémie de 2019-nCoV. Plus tard, la transmission interhumaine a été confirmée. Au 31 janvier, 51 séquences du génome entier de 2019-nCoV provenant de différents laboratoires et régions avaient été soumises à la base de données GISAID [8]. Le 11 mars, l'OMS a officiellement reconnu le COVID-19 comme une pandémie.

Le SRAS-Cov-2 est le troisième coronavirus humain le plus virulent, suivi par le CoV-SRAS et MERS-CoV exprimant le taux de mortalité le plus élevé. [7]

Les coronavirus sont de grands virus reconnus comme des virus hautement évolutifs, avec une fréquence élevée de recombinaison génomique et de mutation. [9]

Il existe six espèces connues de coronavirus humains (HCoV-NL63, HCoV-229E, HCoV-OC34, HCoV-HKU1, SARS-CoV et MERS-CoV) associées à plusieurs maladies des voies respiratoires. [10] Le nouveau SARS-COV-2 est la septième souche de coronavirus humain qui est classiquement placée dans genre Beta coronavirus présentant une similitude de séquence nucléotidique de 89,1% et 60% avec les coronavirus SRAS et MERS, respectivement. [11]

Le génome des coronavirus est composé d'un simple brin d'acide ribonucléique (ARN) portant l'information pour produire plusieurs protéines qui permettent au virus de se multiplier dans la cellule infectée et de créer de nouvelles particules virales infectieuses. [12]

L'infection par SARS-Cov-2 provoque principalement des maladies respiratoires allant d'une maladie bénigne à une maladie grave et à la mort, et certaines personnes infectées par le virus ne développent jamais de symptômes. [13] Qu'elles présentent ou non des symptômes, les personnes infectées peuvent être contagieuses et le virus peut se propager d'elles à d'autres personnes. [14]

I. Mode de transmission du coronavirus

Pour chaque épidémie, la connaissance des voies de transmission de l'agent infectieux est très importante. [15]

En ce qui concerne le COVID -19, l'hypothèse standard est que les voies de transmission suivantes sont dominantes : par contact direct, indirecte ou étroite (à moins de 1 mètre d'intervalle) avec des personnes infectées par des sécrétions infectées telles que les petites particules liquides lorsqu'elle tousse, éternue, parle, chante ou respire fortement. Ces particules liquides sont de tailles différentes, allant des plus grosses gouttelettes respiratoires aux plus petits aérosols et la contamination des surfaces et les objets. Une autre voie de transmission qui retient également l'attention de la communauté scientifique est une voie fécale-orale. [15]

I.1. Contact et transmission de gouttelettes :

Ont un diamètre $> 5 - 10 \mu\text{m}$, la transmission respiratoire de gouttelettes peut se produire lorsqu'une personne infectée émet des gouttelettes contenant des particules virales en toussant, en éternuant et en parlant. Ces gouttelettes atterrissent sur la muqueuse respiratoire ou la conjonctive d'une autre personne, généralement à une distance d'environ deux mètres, mais peut-être plus loin [16].

Les gouttelettes peuvent également se déposer sur des objets fixes ou mobiles et peuvent être transférées à une autre personne lorsqu'elle être en contact avec ces objets.

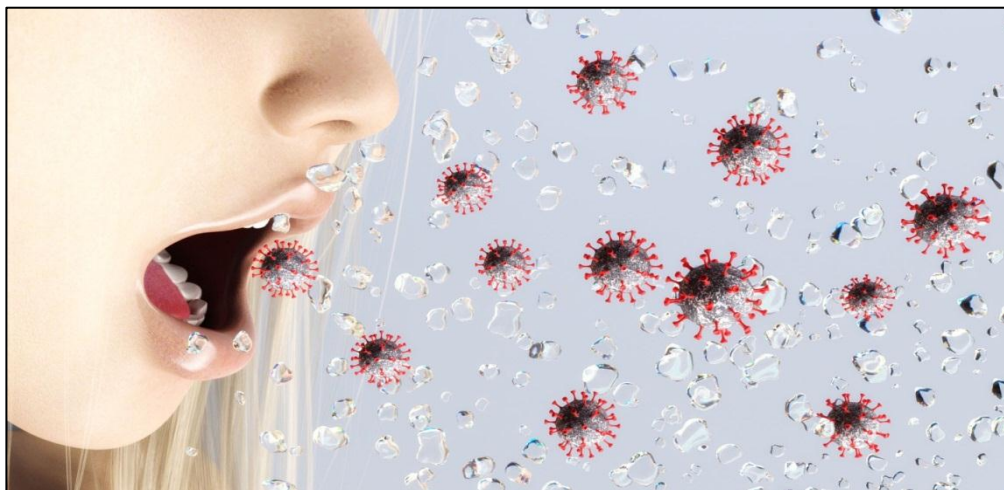


Figure1 Transmission du SARS-Cov-2 via de petites particules en suspension dans l'air. [17]

I.2. Transmission aéroportée :

La transmission par aérosols, des particules plus petites que 5 μm , définie comme la propagation d'un agent infectieux causée par la dissémination de noyaux de gouttelettes (aérosols) qui restent infectieux lorsqu'ils sont en suspension, dans l'air sur de longues distances et dans le temps [13]. Peut se produire dans des environnements spécifiques, en particulier dans des espaces intérieurs, bondés et insuffisamment ventilés, où les personnes infectées passent de longues périodes de temps avec d'autres [14].

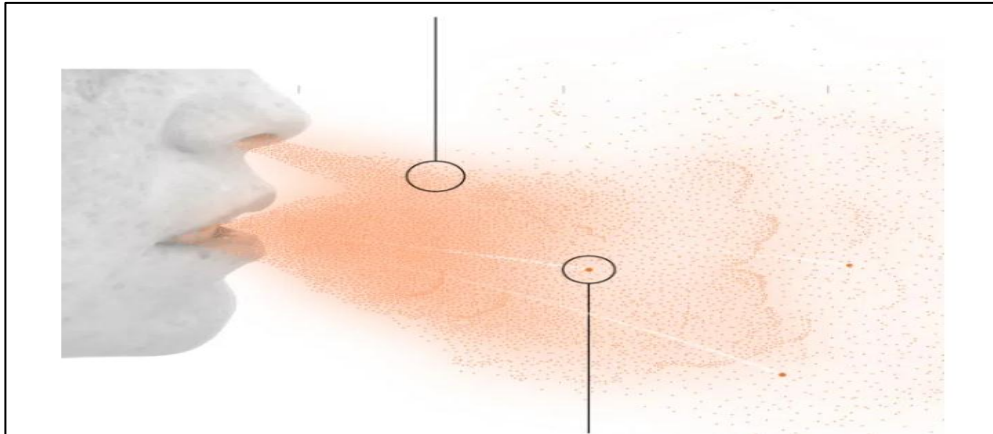


Figure 2 Transmission via des aérosols. [18]

I.3. Transmission fomite (contact indirecte) :

Les sécrétions respiratoires ou les gouttelettes expulsées par les personnes infectées peuvent contaminer les surfaces et les objets, créant les fomites. Peuvent être trouvées le virus ces surfaces pendant des périodes allant de quelques heures à plusieurs jours, selon l'environnement ambiant (y compris la température et l'humidité) et le type de surface, en particulier à forte concentration dans les établissements de soins de santé où les patients COVID-19 étaient traités [16].

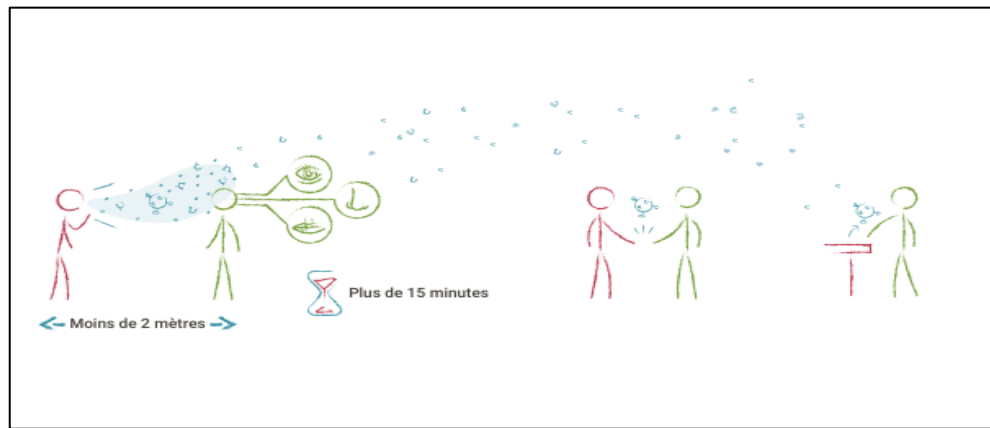


Figure 3 Transmission avec des surfaces contaminées. [19]

La transmission peut également se produire indirectement en touchant ou des objets contaminés par les virus d'une personne infectée puis en touchant la bouche, le nez ou les yeux. Cependant, la transmission fomite est considérée comme un mode probable de transmission du SARS-Cov-2 [16].

Le risque de transmission du SRAS-CoV-2 est par ailleurs augmenté dans des espaces restreints, ventilés de façon inadéquate, à forte densité d'occupants et lorsque la durée d'exposition est prolongée. Dans ces conditions spécifiques, la distance maximale à laquelle on peut être infecté demeure imprécise, mais il est peu probable que ce soit au-delà de quelques mètres. [19]

II. les tests et diagnostic

Poser rapidement le diagnostic de COVID-19 – récemment renommé COVID – avec le plus d'exactitude possible est la pierre angulaire du contrôle de la pandémie. Mais cette évidence partagée par tout un chacun se heurte au caractère véritablement multiforme de l'infection par le SARS-CoV2, appelée COVID-19 : à côté des formes restant totalement asymptomatiques, on observe en effet des formes légères ou paucisymptomatiques, des formes modérées à sévères, ces dernières nécessitant des soins hospitaliers, et enfin des formes très graves nécessitant une admission en soins intensifs et une ventilation assistée. L'ensemble de ces manifestations de l'infection virale est susceptible de contribuer à la transmission du virus dans les collectivités.

Parmi les tests diagnostiques qui nous permettent de confirmer le COVID, la réaction de transcription inverse suivie d'une réaction de polymérisation en chaîne quantitative en temps réel (RT-qPCR), et le test de diagnostic rapide basé sur la détection de l'antigène spécifique du SARS-CoV-2 sont deux méthodes utilisées dans la phase précoce des

manifestations infectieuses. Les tests de détection des anticorps sériques (ELISA et test de flux latéral) sont utilisés dans la phase ultérieure et après la guérison. En l'absence de « test de référence » parfait, les sensibilités et spécificités respectives citées dans les différentes études doivent être considérées avec prudence et ne seront donc que brièvement commentées. [20]

II.1. Les Symptômes

La covid-19 affecte les individus de différentes manières. La plupart des personnes infectées développent une forme légère à modérée de la maladie et guérissent sans hospitalisation.

Symptômes les plus fréquents :

- fièvre
- toux sèche
- fatigue

Symptômes moins fréquents :

- courbatures
- maux de gorge
- diarrhée
- conjonctivite
- maux de tête
- perte de l'odorat ou de goût
- éruption cutanée, ou décoloration des doigts ou des orteils

Symptômes graves :

- difficultés à respirer ou essoufflement
- sensation d'oppression ou douleur au niveau de la poitrine
- perte d'élocution ou de motricité. [21]

II.2. La détection du matériel génétique du virus

La PCR, polymérase chaîne réaction ou réaction de polymérisation en chaîne, est une technique d'amplification enzymatique permettant d'obtenir un grand nombre de copies identiques d'un fragment d'ADN à partir de moins de 1 pictogramme d'une gène, soit une amplification de l'ordre du milliard. [22]

❖ Fonctionnement de la PCR

L'ARN présent dans le prélèvement doit tout d'abord être purifié par ajout de différents solvants. Cette étape d'extraction, qui se termine par la resuspension de l'ARN

dans de l'eau, dure entre une et deux heures. La RT-PCR elle-même comprend deux étapes majeures. L'ARN doit tout d'abord être transformé en ADN par une enzyme transcriptase réverse (RT). Cette enzyme prend l'ARN comme modèle pour synthétiser une séquence d'ADN dit complémentaire (ADNc). L'ADNc du virus, si celui-ci est présent dans le prélèvement, est ensuite fortement amplifié par une réaction de polymérase en chaîne (PCR) quantitative. Cette réaction a lieu en trois phases. Une dénaturation de l'ADNc par chauffage à 95°C pour séparer les deux brins qui le composent, une hybridation des amorces aux extrémités de la séquence recherchée, puis une élongation grâce à l'action d'une enzyme ADN polymérase à 58°C. Les amorces sont des séquences d'ADN simple brin spécifique du virus. Ce sont elles qui garantissent la spécificité de la réaction d'amplification. Seuls les brins d'ADNc fixant ces amorces sont amplifiés. La durée d'un cycle de PCR est de l'ordre d'une minute. Il est répété 45 fois pour obtenir une multiplication exponentielle de la séquence d'ADN cible. C'est cette phase d'amplification qui confère au test RT-PCR une très haute sensibilité. [23]

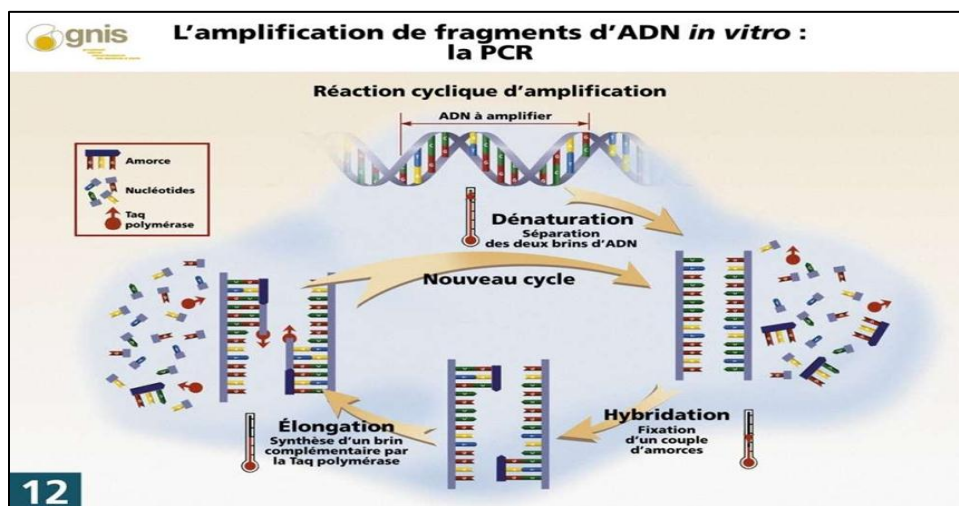


Figure 4 la technique d'amplification de l'ADN par PCR[22]

Sa fiabilité dépend de nombreux facteurs. La qualité du prélèvement est critique. Celui-ci doit être réalisé assez profondément dans les cavités nasales du patient à l'aide d'un grand coton-tige ce qui nécessite une bonne maîtrise. Il a aussi été observé que le virus pouvait être indétectable dans les voies respiratoires supérieures, mais présent dans les poumons. En conséquence, on estime que la fiabilité du test RT-PCR, malgré sa très haute spécificité ($\approx 100\%$) et sensibilité, n'est que de 60-80 % pour identifier un individu infecté. Cette fiabilité décroît avec le temps car le virus est éliminé par la réponse immunitaire. Elle n'est plus que de 40-50 % entre 15 et 39 jours post infection. Ce

pourcentage peut sembler faible, mais il est similaire à celui des tests de détection par RT-PCR du virus influenza. [23]

II.3. La détection des protéines du virus

Des tests rapides pouvant être réalisés sans passer par un laboratoire, directement sur le terrain (tests dits Point-of-Care, POC), ont également été développés pour détecter la présence du virus. [23]

Les tests dits « antigène rapide » permettent la détection des protéines du virus chez un individu en quelques minutes. Un prélèvement est réalisé dans les cavités nasales, comme pour le test RT-PCR. La présence des protéines virales est mise en évidence à l'aide d'anticorps spécifiques (deux types d'iso types d'anticorps : IgG et IgM) de ces protéines couplées à une enzyme permettant une réaction colorimétrique sur une languette, comme pour un test de grossesse disponible en pharmacie. [23]

À la différence de tests par RT-PCR, les tests antigène rapide ne comportent pas de phase d'amplification du signal et ne détectent le virus que lorsqu'il est présent à un titre élevé. Ils sont donc moins sensibles et fiables que le test RT-PCR pour identifier un individu infecté. Ces tests sont considérés comme des tests rapides d'orientation de diagnostic (TROD). En cas de résultat négatif, il est prudent de confirmer le test antigène rapide par un test RT-PCR. [23]

❖ Fonctionnement de la teste antigénique rapide

Il existe plusieurs types de tests mais le plus courant consiste à fixer des anticorps anti-IgG et IgM humaines sur la surface de la cassette et de coupler un antigène du virus avec des particules d'or colloïdal. Si l'échantillon du patient contient des anticorps anti-SARS-CoV-2 alors ces anticorps se fixeront à l'antigène présent dans la zone de conjugaison de la cassette et le complexe formé migrera jusqu'à anti-IgG et/ou IgM humaines fixés sur la membrane. On verra alors apparaître une bande colorée (voir les différents résultats possibles plus bas). [24]

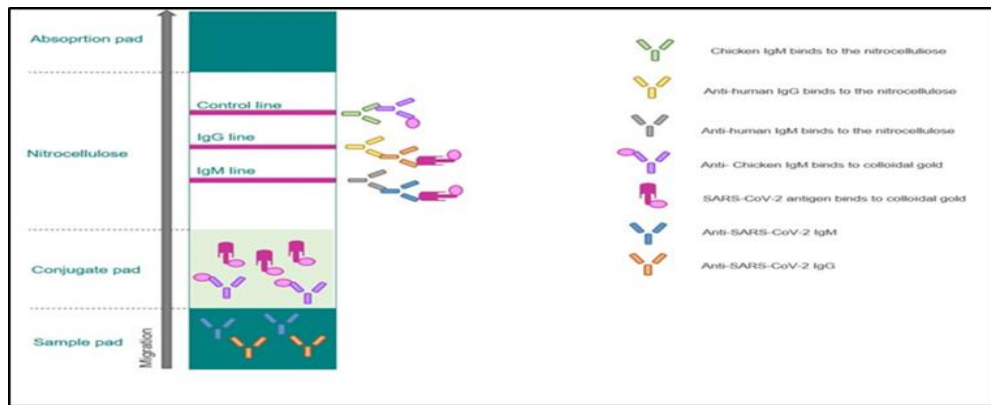


Figure 5 les différents résultats des bandes colorées possible. [24]

Le niveau d'anticorps IgM commence à augmenter environ une semaine après l'infection initiale, tandis que les IgG apparaissent plus tard que les IgM (généralement dans les 14 jours suivant l'infection) et peuvent durer 6 mois, voire plusieurs années, ce qui signifie que les IgG servent d'indicateur d'une infection antérieure. Les patients qui sont infectés par le SARS-CoV-2 peuvent être rapidement identifiés par la surveillance simultanée des IgM et des IgG. la détection des anticorps IgM / IgG a été utilisée comme l'une des méthodes de diagnostic recommandés. [24]

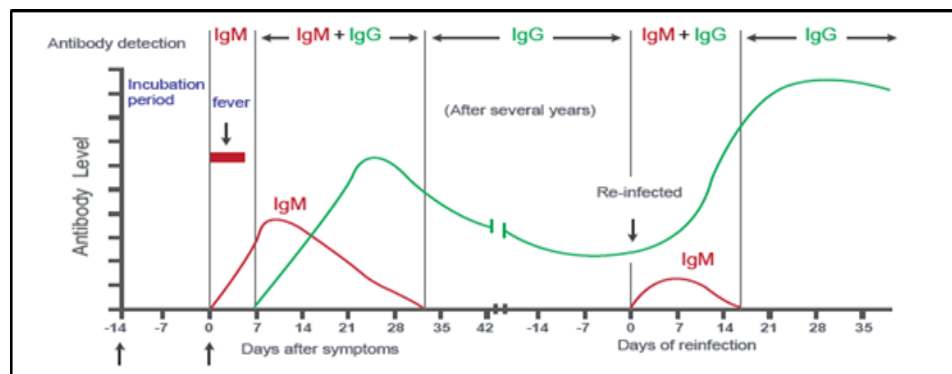


Figure 6 les étapes de développement des anticorps au stade

De l'infection [24]

Un échantillon peut être positif si des anticorps IgM et/ou IgG sont présents. [24clinique]

Il existe différentes cassettes pour les tests rapides. En général, pour la détection qualitative des IgG et des IgM en même temps, il y a 3 lignes différentes : une pour les IgG, une pour les IgM et une pour le contrôle. Et pour être validé, ce test doit présenter une ligne positive pour le contrôle (C) [24]

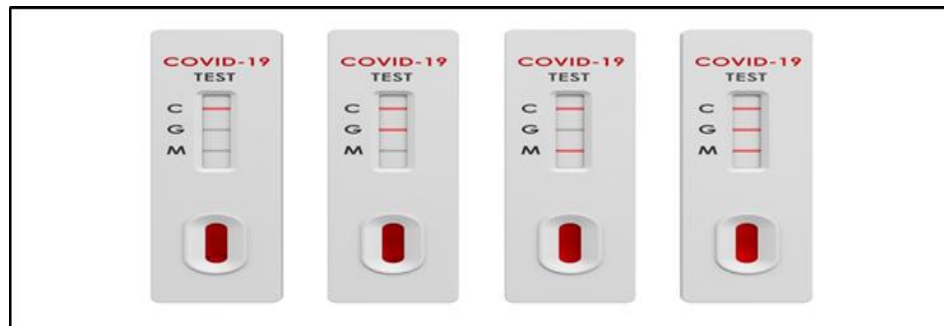


Figure 7 les différents résultats de tests antigéniques rapides [24]

Résultats	Interprétation
IgM+ /IgG+	Infection récente au SARS-cov-2
IgM+ /IgG-	Infection récente au SARS-cov-2
IgM- /IgG+	Infection antérieure au SARS-cov-2
IgM- /IgG-	Pas d'infection ou pas d'anticorps détectables pendant le début de l'infection

Tableau 1 présentation de différent résultats obtenir par teste antigénique et leur interprétation. [24]

II.4. La détection de la réponse immunitaire contre le virus

En réponse à l'infection, le système immunitaire de l'hôte produit des anticorps spécifiques contre les protéines du virus. Une partie de ces anticorps empêchent le virus de se fixer sur les cellules de l'hôte et sont dits neutralisants. [23]

Les anticorps contre le virus sont présents dans le sang des individus infectés. Ils peuvent être détectés en réalisant un test immuno-enzymatique. Des protéines recombinantes du virus, synthétisées in vitro par génie génétique, sont fixées sur un support et capturent les anticorps spécifiques présents dans le sérum du patient. La présence d'anticorps est ensuite révélée par une réaction enzymatique qui libère un composant coloré. [23]

❖ Test ELISA

C'est une technique immuno-enzymatique de détection qui se fait en laboratoire et qui permet de visualiser une réaction antigène-anticorps grâce à une réaction colorée produite par l'action sur un substrat d'une enzyme préalablement fixée à l'anticorps. L'utilisation d'anticorps monoclonaux rend la détection spécifique et la réalisation d'une gamme en parallèle (droite de référence réalisée en diluant de manière sériée avec un contrôle positif) permet de quantifier les anticorps du patient présents dans le sang.

L'ELISA nécessite la réalisation de différentes étapes successives : antigène spécifique du virus SARS-CoV-2 (la protéine N contenu dans la nucléocapside virale où le récepteur de liaison du virus dit RBD (Receptor Binding Domain) est fixé pendant une nuit dans le fond d'un puits d'une plaque 96 puits (« coating ») ; les anticorps présents dans l'échantillon de plasma du patient vont se fixer spécifiquement sur l'antigène. Un anticorps de détection va ensuite fixer les anticorps humains à doser. Ces anticorps de détection sont couplés à une enzyme qui en présence de son substrat le transforme en produit de réaction détectable et mesurable grâce à l'apparition d'une coloration. L'intensité de celle-ci est proportionnelle à la quantité d'enzymes présentent et donc à la concentration d'anticorps recherchés. [20]

Certaines de ces étapes prennent plusieurs heures. Le test ELISA ne peut être effectué sur une goutte de sang. Ce n'est donc pas un test rapide et il ne peut être réalisé au lit de la malade présence d'anticorps humaine anti-SARS-CoV-2 dans les échantillons de sérum. La validation sur le plasma et la goutte de sang est en cours d'évaluation. [20]

III. Epidémiologie et statistique du COVID-19

III.1. à l'échelle mondiale

À 10h50 CEST, le 16 juin 2021, **176 156 662** cas confirmés de COVID-19, dont **3815 486** décès, ont été signalés à l'OMS. Un total de **2310 082 345** doses de vaccin ont été administrées.

III.2. à l'échelle nationale

Données communiquées à l'OMS au cours des dernières 24 heures. Dernière mise à jour : 16 juin 2021. Nouveaux cas : **354**, cas confirmés : **1330742**, décès confirmés **3579**. [25]

Le 15 juin 2021, le nombre de nouveaux cas et de décès liés au COVID-19 a continué de diminuer, avec plus de 2,6 millions de nouveaux cas et 72 000 nouveaux décès signalés dans le monde. Alors que le nombre de cas signalés dans le monde dépasse désormais 175 millions, au cours de la semaine dernière, l'incidence hebdomadaire de cas la plus faible depuis février 2021 a été signalée. Des baisses du nombre de nouveaux cas hebdomadaires, par rapport à la semaine précédente, ont été signalées dans toutes les Régions, à l'exception de la Région africaine. Le nombre de nouveaux décès signalés au cours de la semaine dernière a diminué dans toutes les régions, à l'exception des régions de l'Afrique et de l'Asie du Sud-Est. [26]

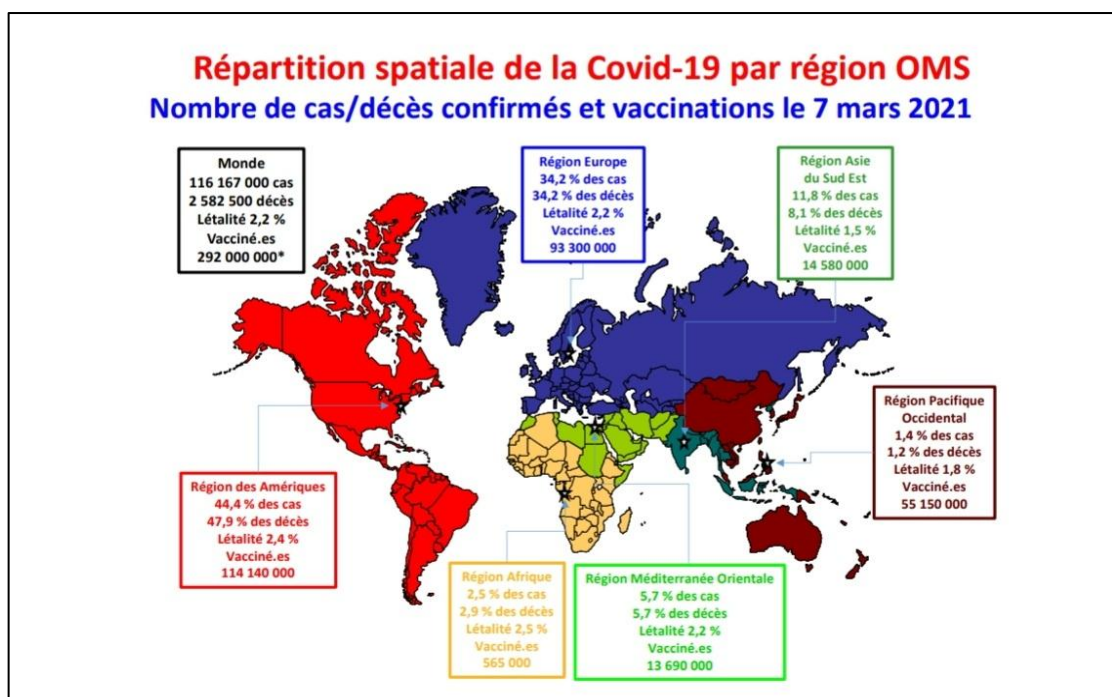


Figure 8 Répartition spatiale du COVID-19 par région

Chapitre 2

Introduction

Nous sommes de plus en plus souvent confrontés à des maladies infectieuses à l'origine d'épidémies, pour lesquelles il n'existe pas encore de traitement efficace. La plupart de ces épidémies sont dues à l'émergence de microorganismes pathogènes, notamment des virus, qui se transmettent naturellement d'animaux vertébrés aux humains. L'ampleur de l'épidémie va ensuite dépendre de la capacité de transmission du pathogène entre individus contaminés. [27]

Plusieurs options peuvent être envisagées pour contrôler ou prévenir les infections émergentes à 2019-nCoV, notamment les vaccins, les anticorps monoclonaux, les peptides, les thérapies par interféron et les médicaments à petites molécules. Cependant, de nouvelles interventions sont susceptibles de nécessiter des mois voire des années pour se développer. [28]

Les connaissances expérimentales et informatiques recueillies tout au long de ce processus sont essentielles pour identifier les points faibles et les principaux sites d'interaction dans la cible, et développer des procédures de criblage appropriées pour trouver des molécules interférant avec ou modulant sa fonction. [29]

L'une des protéines les plus importantes du virus est une enzyme, appelée **3CLpro** ou protéase principale, déjà identifiée comme une cible pharmacologique importante également dans les virus du SRAS et du virus du syndrome respiratoire du Moyen-Orient (MERS). Par conséquent, il est essentiel d'acquérir une compréhension plus approfondie de la structure et de la fonction de 3CLpro afin de cibler efficacement cette enzyme. [30]

I. Les protéases virales

Les protéases virales sont essentielles à la production de virions par la cellule infectée car elles participent à une étape de maturation en coupant les polys-protéines virales synthétisées à des sites bien précis. Elles sont spécifiques d'un virus donné et agissent sur les protéines virales, ainsi que sur quelques protéines de la cellule hôte, pour favoriser la réplication virale [31]. C'est pour cela qu'elles suscitent un grand intérêt pour stopper la propagation des épidémies virales.

Plusieurs étapes sont nécessaires pour la formation de nouvelles particules virales :

- ❖ Entrée du virus dans la cellule qu'il va infecter.

- ❖ Copier de l'ARN viral.
- ❖ Synthèse du polyprotéine et sa découpe en protéines fonctionnelles de tailles plus petites
- ❖ Enfin, reconstitution de nouvelles particules virales infectieuses et leur relargage à l'extérieur de la cellule infectée.

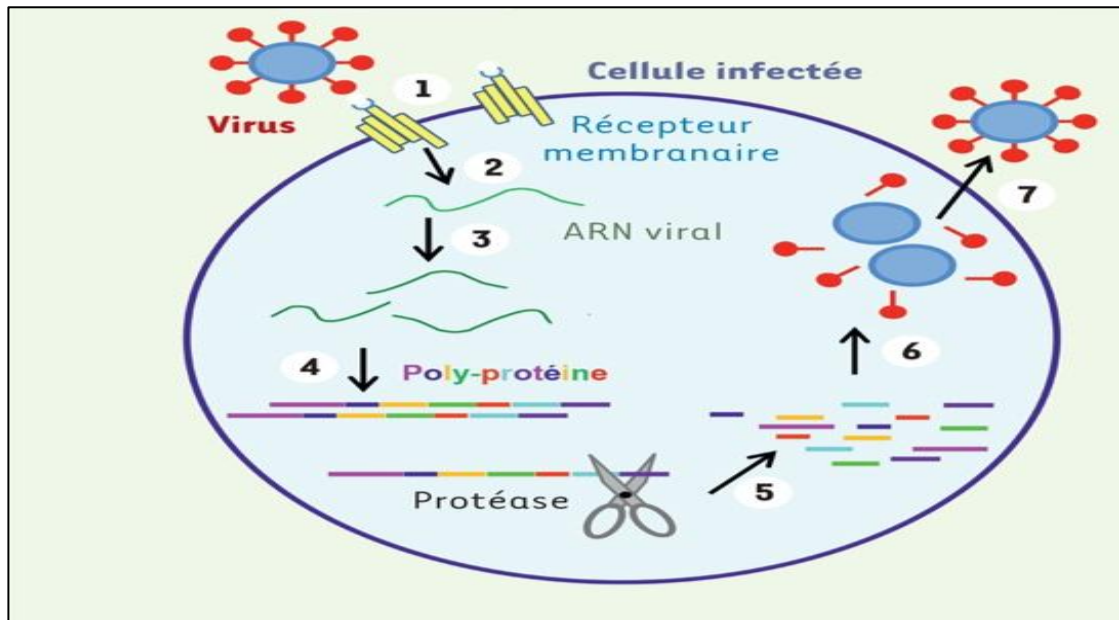


Figure 9 Représentation schématique de la production de particules virales par une cellule [27].

Pour comprendre le fonctionnement des protéases virales, il faut certes connaître leur séquence d'acides aminés, mais aussi leur structure tridimensionnelle. Cela permet alors de caractériser précisément le site catalytique de l'enzyme, qui sera la cible des molécules à usage thérapeutique capables de le bloquer. Les coronavirus possèdent deux protéases [32].

- ❖ 3CLpro joue un rôle clé dans le cycle de vie de la réplication et de la maturation du coronavirus.
- ❖ PLpro, a des propriétés différentes de celles de 3CLpro, notamment une activité déubiquitinase.

II. Les principales cibles corona virales

Parmi les cibles corona virale étudiées dans le passé, la principale protéase (Mpro) a fait l'objet d'une attention particulière, [33] notamment suite à la première flambée de SRAS-CoV au début des années 2000. D'autres cibles coronavirales comprennent la protéine de pointe (S), l'ARN-polymérase, hélicase et la protéase de type papaine (PL pro). [34]

III. La protéase principale (3CLpro, Mpro)

Est une cystéine protéase, sa séquence protéique est identique à 96 % à celle des protéases 3CLpro des autres coronavirus, qui ont été très étudiées [35]. La **Mpro** clive exclusivement les séquences polypeptidiques virales 11 fois au niveau de motifs Leu-Gln-(Ser/Ala/Gly). L'inhibition de l'activité de cette enzyme bloquerait la réplication virale. Elle est positionnante comme une cible médicamenteuse idéale car, aucune protéase de cellule hôte humaine n'est connue avec cette spécificité de substrat. [36]

IV. Structure du 3CL pro et le site actif

La structure 3D de 3CLpro du SRAS-Cov-2, comme d'autres coronavirus, se compose de trois domaines :

Le domaine I (8-101 résidus d'acides aminés) et II (102-184 résidus d'acides aminés) sont essentiellement des barils bêta et ressemblent à la chymotrypsine tandis que le domaine III (201 à 306 résidus d'acides aminés) était principalement composé d'hélices alpha. La région de liaison au substrat, située à la fente des Domaines I et II, se compose de His41 et Dyade catalytique Cys145, dans laquelle Cys agit comme nucléophile tout en Son agit comme un accepteur de protons. En plus du Centre catalytique, il existe deux autres sous-sites profondément enfouis (S1 & S2) et trois sous-sites peu profonds (S3-S5). Le sous-site S1 se compose de His163, Glu166, Cys145, Gly143, His172 et Phe140 Tandis que S2 est constitué de résidus d'acides aminés Cys145, His41 et Thr25 ; principalement impliqué dans les interactions hydrophobes et électrostatiques. Les sous-sites Shallow S3-S5 comprennent Met49, His41, Résidus d'acides aminés Met165, Glu166 et Gln189. Ces sous-sites peu profonds peuvent tolérer différentes fonctionnalités. [37] [38]

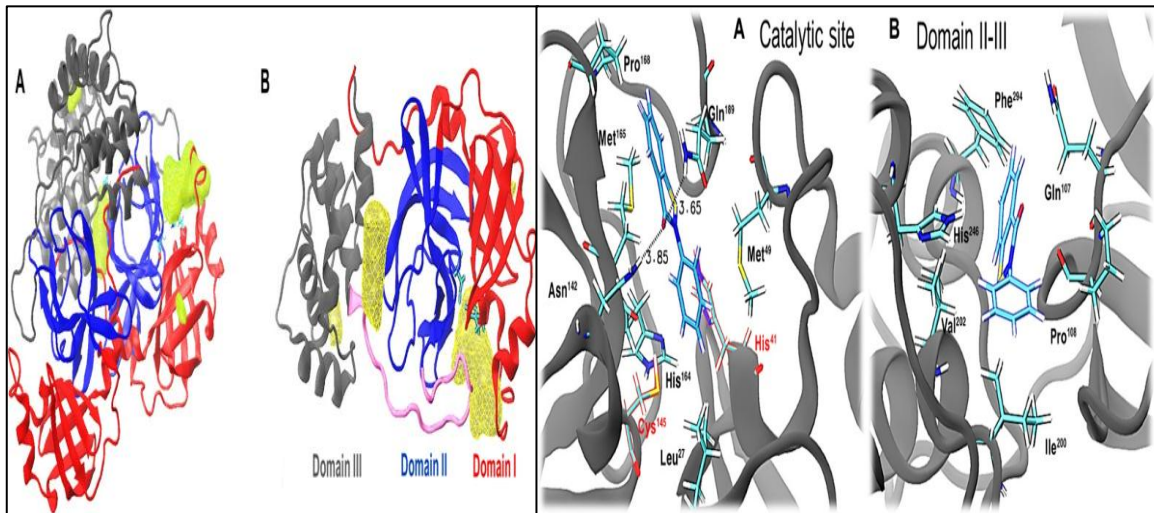


Figure 10 La structure générale de 3CL pro et le site actif. [39]

V. Les inhibiteurs de 3CLprotéase

Les agents antiviraux approuvés pour le traitement des patients atteints de COVID-19 restent indisponibles. La réutilisation des médicaments d'antiviraux approuvés contre d'autres virus tels que le VIH ou le virus Ebola est l'une des stratégies les plus pratiques pour développer des agents antiviraux efficaces contre le SRAS-Cov-2. Une combinaison de médicaments réutilisée peut améliorer l'efficacité du traitement, et une conception de médicaments basés sur la structure peut être utilisée pour cibler spécifiquement le SRAS-Cov-2. [40]

L'inhibition de la protéase principale ou de l'activité 3CLpro est une stratégie prometteuse pour contrôler l'infection émergente du SRAS-CoV-2.

1. Ebselen

Ebselen (2-phényl-1,2-benzosélenazol-3 (2H) -one) a été décrit pour la première fois comme une enzyme catalysant la réaction du glutathion peroxydase [41]. La principale protéase du coronavirus SARS-CoV-2 est une cible médicamenteuse potentielle. L' IC_{50} de l'ebselen est de 0,67 μ M [80].

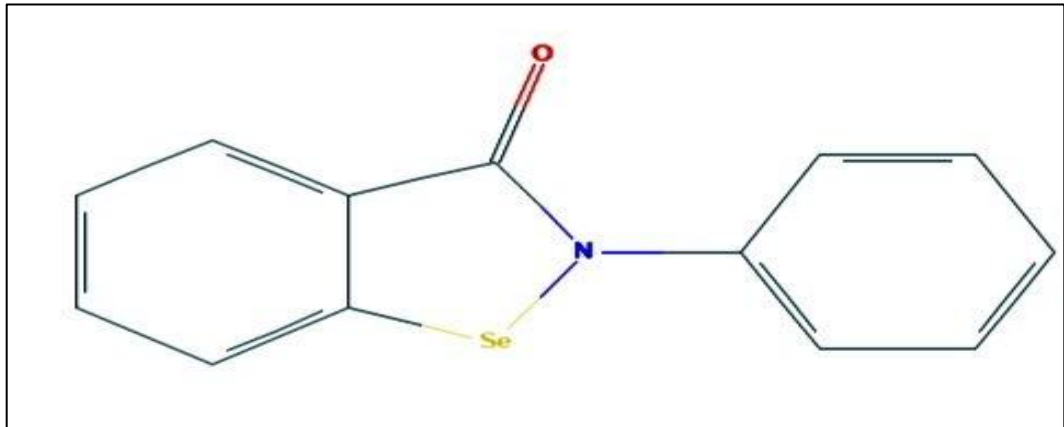


Figure 11 Structure de l'inhibiteur **Ebselen** [42]

2. Bocéprivir :

Un médicament anti-VHC approuvé par la FDA, inhibe l'activité enzymatique de M pro avec une IC_{50} de 4,13 μ M contre le SRAS-CoV-2 dans le test CPE viral cellulaire. [43]

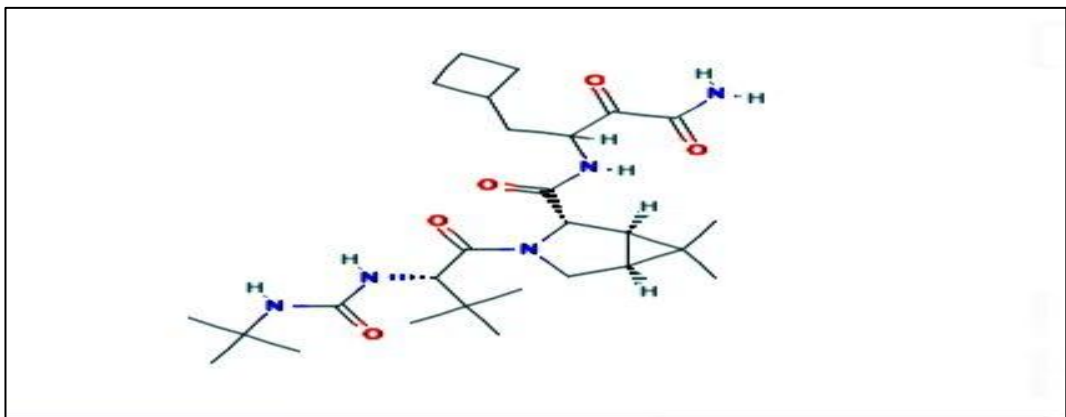


Figure 12 Structure de l'inhibiteur **Bocéprivir** [42]

3. GC-376

Un médicament vétérinaire expérimental actuellement en cours de développement pour traiter la péritonite infectieuse féline. Présente des propriétés pharmacocinétiques in vivo optimales et une efficacité antivirale in vivo dans le modèle de chat infecté par le virus

de la péritonite infectieuse féline. [44, 45] Est un inhibiteur de protéase du type 3CL pro qui arrête le clivage et l'activation des protéines virales fonctionnelles nécessaires à la réplication et à la transcription dans les cellules hôtes avec une valeur d'IC₅₀ de 0,62 μM.

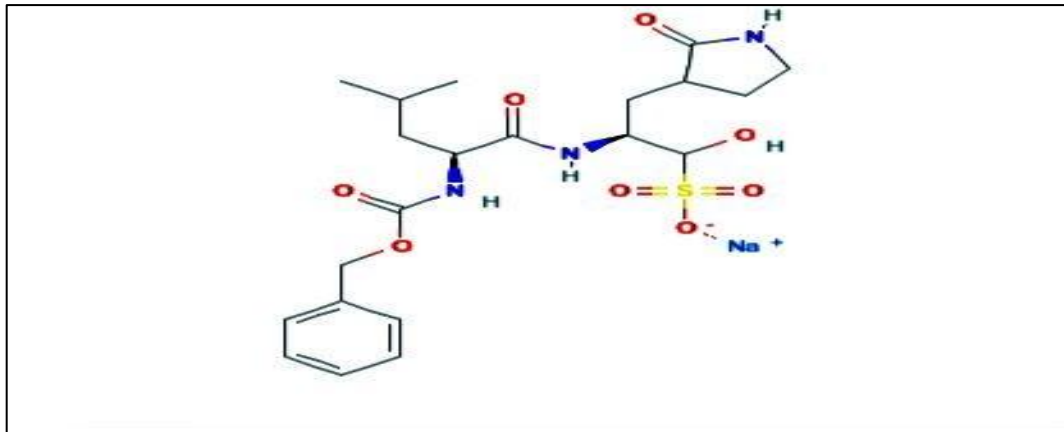


Figure 13 Structure de l'inhibiteur GC-376 [42]

4. α-kétoamides

A été retenue pour son efficacité d'insertion dans le site catalytique de la protéase, ce qui bloque son activité et permet de réduire considérablement la production virale dans des cellules du poumon humaines infectée avec le SRAS-Cov-2 in vitro, ainsi que dans des modèles d'études précliniques [35]. Elle est inhibe l'activité enzymatique de Mpro avec une valeur de IC₅₀ de 0.67 μM.

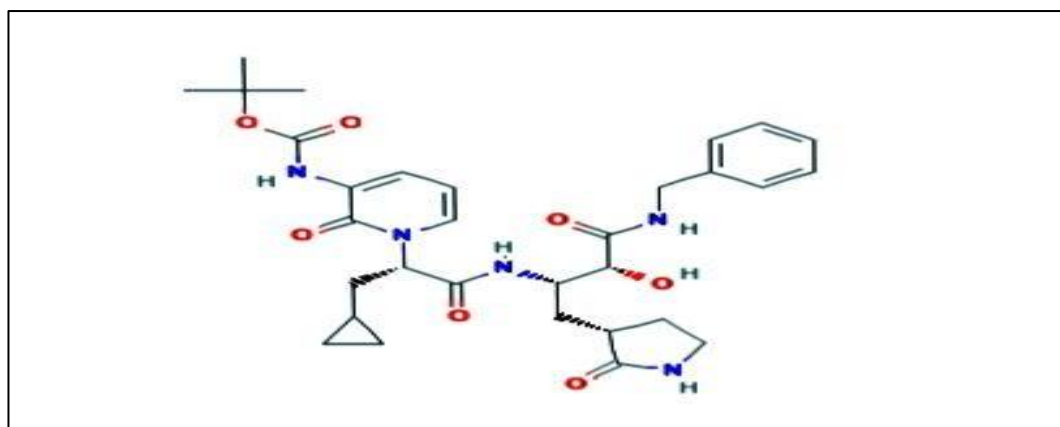


Figure 14 Structure de l'inhibiteur α-kétoamides [42]

5. Tideglusib

Est un médicament à petite molécule disponible par voie orale de la classe des thiazolidinones. [42] Il agit comme un inhibiteur irréversible du glycogène synthase kinase 3. Inhibe la principale protéase avec une IC_{50} de 1.55 μ M.

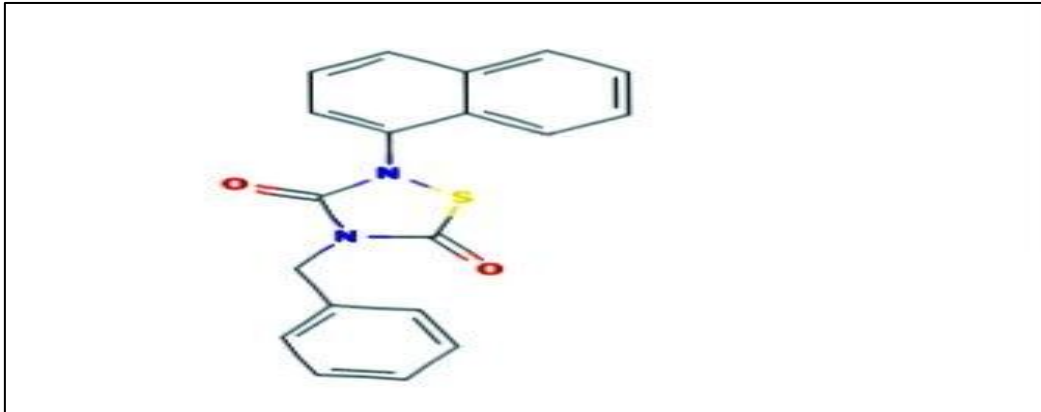


Figure 15 Structure de l'inhibiteur **Tideglusib** [42]

6. Carmofur

Est un dérivé antinéoplasique antimétabolite (analogue à la pyrimidine) du 5-fluorouracile, inhibe la céramidase acide humaine en modifiant de manière covalente le résidu catalytique C143. [46] le carmofur inhibait la réplication virale du SRAS-CoV-2 avec une IC_{50} de 1.82 μ M.

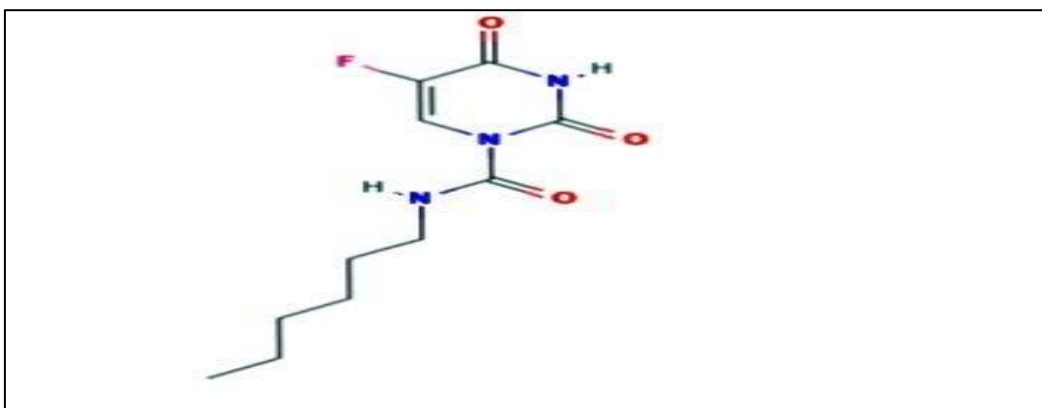


Figure Structure de l'inhibiteur **Carmofur** [42]

Chapitre 3

I. Principe théorique

La conception d'un nouveau médicament est extrêmement longue et onéreuse depuis la première phase de recherche jusqu'à sa mise sur le marché. En effet, le temps nécessaire à la découverte d'un nouveau médicament est estimé en moyenne entre 12 et 14 ans pour un cout total de 800 millions à 1 milliard de dollars [47]. C'est dans ce contexte que les techniques de docking in silico ont été développées.

Docking (ancrage ou amarrage, en français) est le nom donné aux simulations moléculaires dans lesquelles différentes approches sont combinées pour étudier les modes d'interaction entre deux molécules. Dans la plupart des cas, il s'agit d'un récepteur macromoléculaire (cible de docking) dont la structure tridimensionnelle est connue et d'une petite molécule (ligand) [48].

Une simulation de docking comprend essentiellement deux étapes : le docking proprement dit et le scoring.

❖ La première (le docking) est l'étape de sélection :

Consistant à placer le ligand dans le site actif de la protéine et à échantillonner les conformations, positions et orientations (poses) possibles, en ne retenant que celles qui représentent les modes d'interaction les plus favorables. Bien que pouvant être faite « manuellement », cette étape est la plus souvent exécutée de manière automatisée à l'aide d'algorithmes de docking, ce qui améliore la vitesse et la précision des simulations. [49]

❖ La deuxième (le scoring) est l'étape de classement :

Qui consiste à évaluer l'affinité entre le ligand et la protéine dans chacun des complexes sélectionnés lors de l'étape précédente. Cette tâche est accomplie par des outils informatiques nommées fonction de score. Celle-ci évaluent d'abord les différentes contributions énergétiques (parfois entropique) pour la stabilisation du complexe protéine-ligand, et ordonnent ensuite ces complexes, permettant ainsi l'identification des modes d'interaction les plus probable [49].

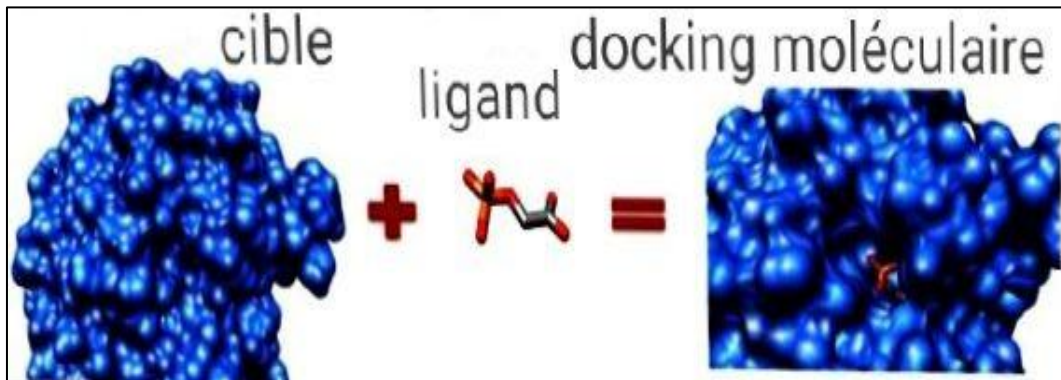


Figure 17 Docking protéine-ligand.

La résolution d'un problème de docking implique la combinaison entre un meilleur algorithme de recherche qui devrait créer un nombre optimum de configuration et une fonction de score qui sert à évaluer et filtrer ces configurations [50].

I. Les types de docking

Il existe trois types de docking moléculaire : le premier est le docking rigide qui est basé sur le modèle « clé-serrure », proposé en 1894 par Emile Fischer, ce docking considère la protéine et le ligand comme corps rigide. Cette procédure est la plus simple et la plus rapide à réaliser, car la taille de l'espace de recherche est beaucoup plus petite, mais si la conformation du ligand n'est pas correcte, la probabilité de trouver un ajustement complémentaire sera plus faible. Elle demeure toutefois encore souvent employée pour l'amarrage protéine-protéine.

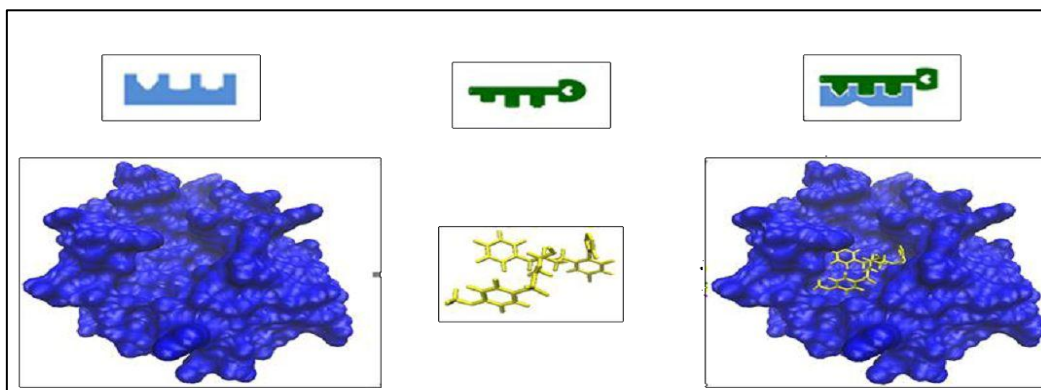


Figure 18 Modèle clé-serrure. [51]

Le deuxième type est le docking semi-flexible qui est souvent utilisé pour l'amarrage protéine-ligand. Le ligand était considéré comme flexible et la protéine gardée rigide. Cependant, certaines protéines présentent naturellement des régions de grande flexibilité, subissant des réarrangements considérables en présence d'un ligand. Dans ce cas, négliger la flexibilité de la protéine peut mettre en péril la fiabilité des résultats du docking, et rend nécessaire l'utilisation d'approches capables de tenir compte de la flexibilité du système entier.

En revanche, la troisième classe qui est le docking flexible traite la flexibilité des deux molécules, mais la flexibilité permise reste limitée. [52] [53] [54].

II. Reconnaissances moléculaire

Le docking moléculaire s'intéresse particulièrement à l'étude des interactions intervenant lors de la formation des complexes moléculaires telle que : les interactions hydrophobes, les interactions électrostatiques et les liaisons hydrogènes [55] [56].

III.1 Interaction de Van Der Waals

Les interactions de Van Der Waals, également connues sous le nom d'interactions de London ou de forces de dispersion, impliquent des dipôles instantanés. Elles sont de faible intensité (ordre de grandeur de Kcal.mol^{-1}), mais elles sont néanmoins très importantes en pratique à cause de leur très grand nombre dans les complexes protéine-ligand [51].

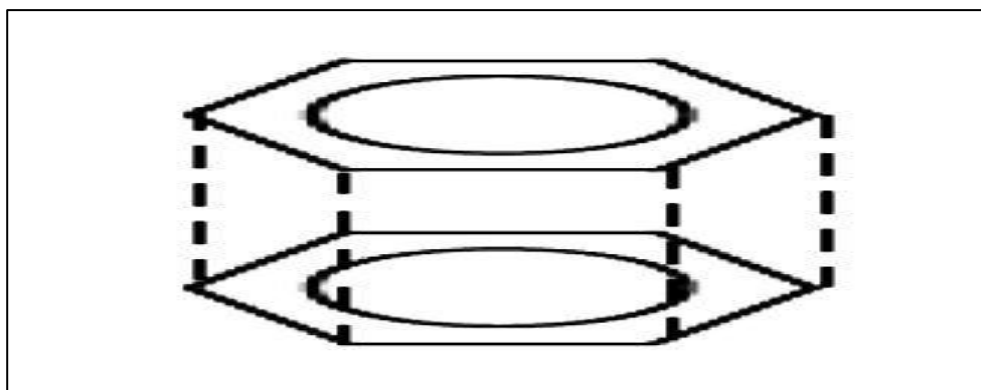


Figure 19 Interaction de Van Der Waals. [56]

III.2 Interactions ioniques

Les interactions ioniques ou ponts salins concernent les fortes interactions électrostatiques attractives entre deux ions de charges opposées [51]. La distance pour une liaison saline, par exemple, est de 3.7 à 4.5 Å, et sa force peut atteindre 8 Kcal/mol.

En présence du solvant, ces interactions sont favorisées du fait de la polarité et de la petite taille de la molécule d'eau [56].

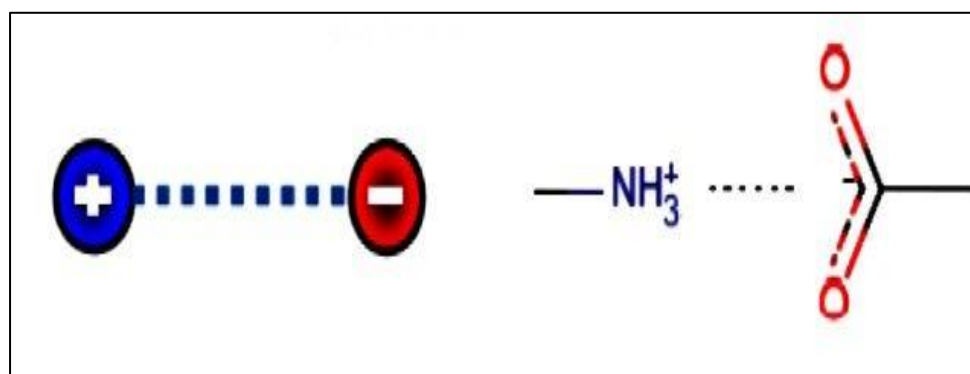


Figure 20 La liaison ionique. [51]

III.3 les liaisons hydrogène

La liaison hydrogène est une liaison physique concerne la force attractive qui s'opère entre deux groupes d'atomes accepteur et donneur impliquant un atome d'hydrogène [57]. Cette liaison avec une force de l'ordre de 2 à 4 Kcal/mol, agit dans les systèmes biologiques à très courte distance de 0.9 à 2.8 Å. elles sont peu nombreuses et s'adaptent très bien à la flexibilité où l'angle peut varier de 120° à 180° [55].

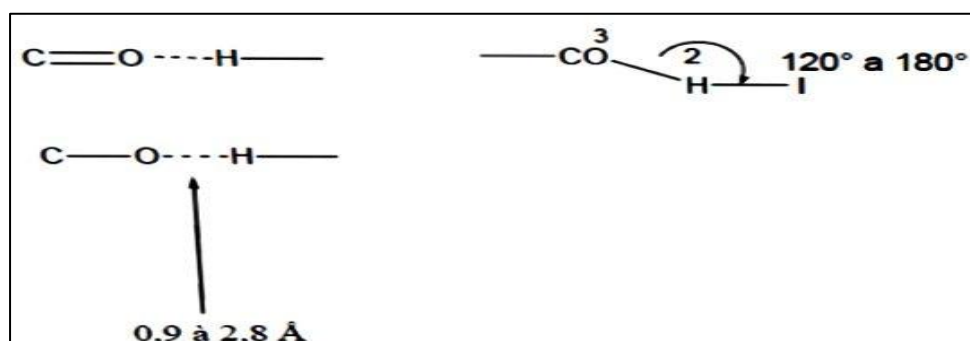


Figure 21 Liaison hydrogène. [56]

III.4. L'effet hydrophobe

Les molécules ou groupes d'atomes non polaires ne sont pas capables de former des liaisons hydrogènes et ne peuvent donc pas s'hydrater [56] [57]. L'effet hydrophobe est la tendance que possèdent ces groupes à se rassembler par coalescence de façon à minimiser les contacts avec l'eau. Les liaisons hydrophobes possèdent une énergie de l'ordre de 7 kcal/mol [51] [57].



Figure 22 Auto-organisation de molécules amphiphiles dans un solvant aqueux

Les zones orange représentent les parties hydrophiles et les queues bleues les chaînes hydrophobes [51].

III. Outils du docking moléculaires

Les éléments indispensables pour pratiquer un docking moléculaire se confinent en un récepteur macromoléculaire de nature protéique, une petite molécule appelée ligand et un programme de docking qui prédit la conformation la plus favorable du ligand au sein du récepteur choisi [58].

IV.1 Le récepteur

La Bank PDB est la plus grande archive de données structurales de macromolécules biologiques (protéine, ADN, ARN,...etc.) dont les protéines. Elle a été créée en 1971 au Brookhaven National Laboratory. Les structures 3D que propose cette banque sont accessibles gratuitement pour les utilisateurs <https://www.rcsb.org/> . Cette banque contient 177009 séquences de structures protéiques obtenues soit par cristallographie (rayons X), soit par RMN. Les coordonnées atomiques des structures se trouvent dans un format de référence, appelé PDB, lu par différents logiciels de visualisation, de modélisation, de criblage virtuel...etc. Donc il suffit de consulter la PDB et d'inscrire le

nom de la structure pour télécharger le fichier du récepteur au format **PDB** indispensable à l'opération de docking moléculaire [59].

Il est important de signaler que certaines protéines ne sont pas encore disponibles dans cette base de données, il devient alors possible de construire la structure 3D de la cible souhaitée, en faisant appel à la modélisation par homologie [60].

IV.2 Ligands 3D (Chimiothèques)

Le choix du ligand est une étape très importante pour le succès du docking moléculaire. Dans ce cadre, le ligand doit être sous forme 3D. Pour y obtenir, il existe deux moyens :

- Le premier moyen consiste à utiliser des ligands de la PDB ou de la littérature qu'on peut dessiner, optimiser et enregistrer sous différents formats (pdb, mol, mol2...etc.) grâce à des programmes de construction moléculaire 3D tels que Titan, ChemDraw, Arguslab, Marvin, Sybyl...etc.
- Le second moyen est souvent d'aspect commercial consistant à consulter des bases de données de structures chimiques appelées chimiothèques [58]. Parmi les chimiothèques virtuelles nous citons :

❖ **Chimiothèque nationale Française:** cette chimiothèque a été créée en 2003. Elle fédère les collections de produits de synthèse et d'extraits naturels existant dans les laboratoires publics français et en assure la valorisation scientifique et industrielle. En 2019, la chimiothèque nationale française regroupe plus de 74.000 molécules. Cette chimiothèque est consultée gratuitement via le lien : <http://chimiotheque-nationale.enscm.fr> [61].

❖ **Pubchem :** est une banque de données américaine de molécules chimiques gérée par le National Center for Biotechnology Information, branche de la Bibliothèque nationale de médecine des États-Unis, sous l'autorité de la Nationale Institutes of Health. Actuellement, la PubChem répertorie plus de 97.000.000 composés en mettant en ligne, gratuitement. Cette chimiothèque est consultée gratuitement via une interface utilisateur Web : <http://pubchem.ncbi.nlm.nih.gov> [42].

❖ **Zinc:** C'est une base de données contenant des composés disponibles à vendre appropriés pour le criblage virtuel. Gouvernée par le Laboratoire Shoichet dans le département de chimie pharmaceutique à l'Université de Californie à San Francisco,

cette base est accessible gratuitement pour l'ensemble des scientifiques. Elle compte actuellement environ 35 millions de composés qui peuvent être achetés tout simplement. Cette chimiothèque est consultée gratuitement via le lien : <http://zinc.docking.org>. [62]

IV. Principaux logiciels de docking

Dans le domaine de docking moléculaire, plusieurs logiciels ont été utilisés pour étudier les différentes interactions existant entre deux entités moléculaires (protéine-ligand). À l'heure actuelle, plus de 30 programmes de docking moléculaire sont disponibles que ce soit commerciaux ou libres [63]. Ces programmes présentent des algorithmes de recherche et des fonctions de score variées. Il est intéressant de remarquer que les trois logiciels AutoDock, Glide et GOLD font partie des logiciels de docking les plus populaires [64]. Les plus fréquemment cités sont respectivement : AutoDock, GOLD, Flex, DOCK et ICM [65] (Voir figure N°19).

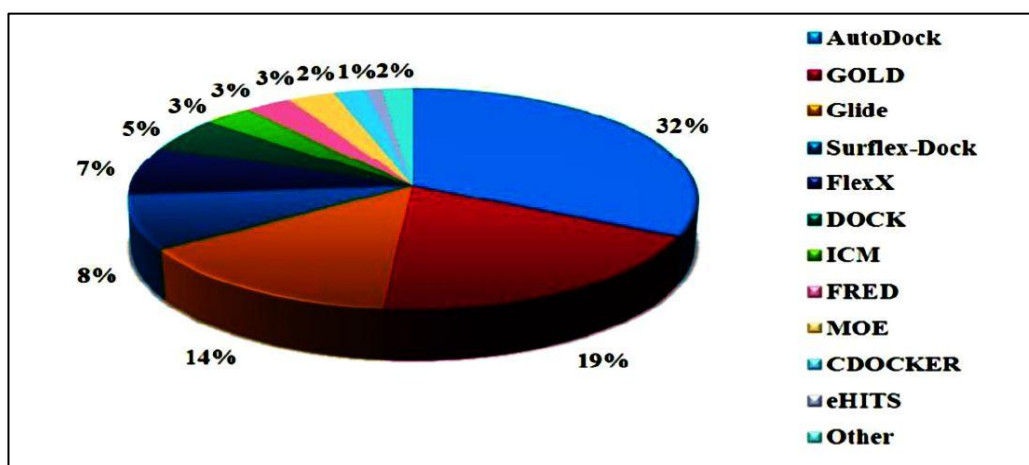


Figure 23 Utilisation des différents logiciels de docking en 2010-2011.

(Yuriev. E et al., 2013)

V.1 Surflex

Parmi les programmes de Docking moléculaire, nous avons utilisé le programme Surflex pour réaliser le travail de recherche. Ce programme repose sur l'ajustement de fragments en commençant par construire un modèle « en négatif » du site actif dit protomol.

Ce modèle est constitué par le volume accessible du site et les points d'interaction tel que les sites de liaison hydrogène, les charges ou les sites lipophiles. Son algorithme

est appelé également de fragmentation incrémentale car, dans un premier temps, la molécule est découpée en parties rigides et en liaisons flexibles. Par la suite, les parties rigides sont positionnées en premier et le ligand est reconstruit de proche en proche, explorant diverses conformations lors de la mise en place des liaisons flexibles ce qui permet un criblage rapide de vastes bibliothèques de composés. [65]

V. Algorithmes de docking

Les algorithmes de docking ont été conçus pour rechercher de façon objective, rapide et efficace les modes d'association protéine-ligand les plus favorables [56]. Ces algorithmes peuvent être divisés en deux catégories principales selon la méthode utilisée :

- ❖ L'ajustement de fragments.
- ❖ Simulation par trajectoire [58].

VI.1 L'ajustement de fragments (Reconstruction)

Dans un premier temps le ligand est découpé en parties rigides et flexibles. Entre les points ou des rotations sont possibles, une ou plusieurs ancres rigides sont définies. Ensuite, une première partie rigide est mise en interaction avec le récepteur, puis les parties flexibles sont ajoutées de manière successive avec une exploration des angles de torsion [66]. Des logiciels tels que DOCK ou SURFLEX utilisent cette approche qui permet notamment un criblage rapide.

VI.2 Simulation par trajectoire (Aléatoire)

Cette approche, est plus précise : à partir d'une position initiale aléatoire, à l'extérieur du site actif, le ligand considéré dans son entier explore le site étudié par la répétition successive de mouvements et d'évaluations de l'interaction ligand-récepteur [12] AutoDock fait partie de la seconde catégorie [67].

VI. Les étapes du docking moléculaire

Les étapes clés communes à tous les protocoles de docking, commencent d'abord par le choix des structures 3D de la macromolécule cible et le ligand. Puis, chaque structure doit être préparée conformément aux exigences de la méthode utilisée. Après le docking, les résultats doivent être analysés [64].

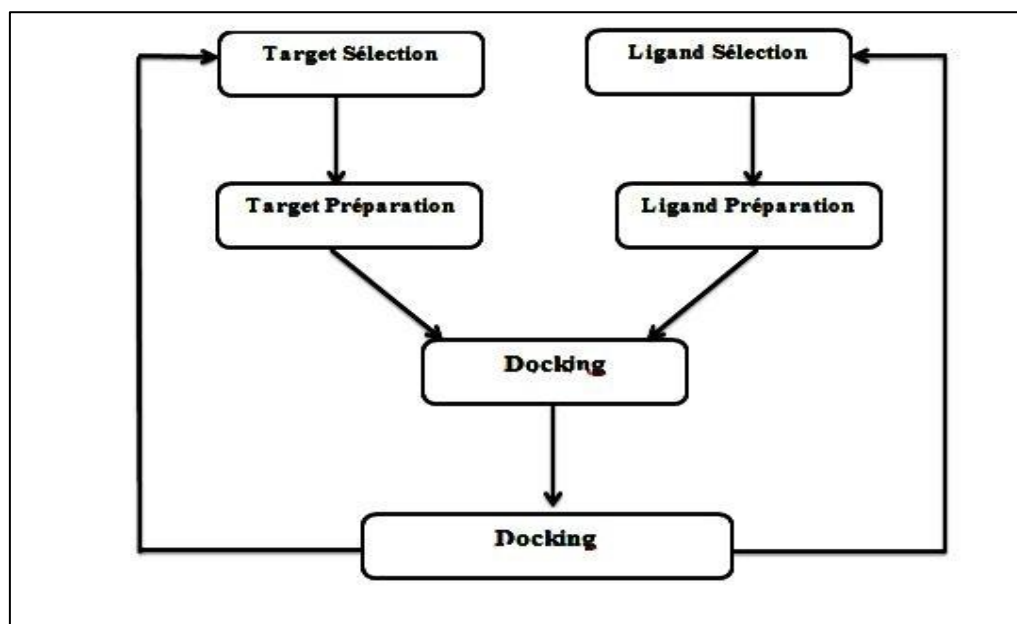


Figure 24 Les étapes clés du docking moléculaire.

VII. Application

Il y a deux principales applications au programme de docking :

- ❖ La plus ancienne est la prédiction du mode d'interaction.
- ❖ La seconde à optimiser des molécules ayant déjà une activité avec le récepteur.

Plusieurs méthodes sont utilisées pour évaluer les performances des différents programmes de docking pour chaque application. L'aptitude d'un algorithme à trouver l'emplacement correct du ligand par rapport à son récepteur est habituellement déterminée au moyen de la déviation quadratique moyenne ou RMSD du modèle conçu par le logiciel vis-à-vis de la structure du cristal [68]. La valeur admise est une différence maximale de 2 Å au-delà de laquelle la prédiction est considérée comme non adéquate [69].

Partie pratique

Chapitre 4

Matériels et

Méthodes

Dans le cadre de ce travail, nous nous sommes intéressées aux méthodes de docking moléculaire afin de proposer *in silico* de nouveaux inhibiteurs de l'enzyme 3CLpro. Pour y arriver, nous avons utilisé le matériel suivant :

I. Matériel

1. Micro-ordinateur

Notre travail a été effectué en utilisant deux micro-ordinateurs dont les caractéristiques sont rassemblées dans le tableau suivant :

Tableau 2 Les micro-ordinateurs utilisés et leurs caractéristiques

	Micro-ordinateur 1	Micro-ordinateur 2
Mémoire (RAM)	2.00 Go	2.00 Go
Type de Système	Système d'exploitation 32 bits	Système d'exploitation 64 bits
Processus	Intel(R) celeron(R) 2957u@1.40GHz 1.40GHz	Intel(R) celeron(R) Cpun3060@1.60GHz 1.60GHz
Système D'exploration	Windows 7 édition intégral version 2009	Windows 10 professionnel

La réalisation des calculs de docking moléculaire a nécessité l'utilisation d'un ensemble de programmes informatiques, la consultation de plusieurs sites internet et certaines bases et banques de données.

2. Banques de données

Le réseau internet a été très utile dans notre étude pour l'obtention des complexes protéiques (protéine-ligand), pour la simulation des propriétés physicochimiques et pharmacocinétiques de nos meilleurs inhibiteurs.

2.1. PDB (Protein Data Bank)

La structure 3D de l'enzyme 3CLpro faisant l'objet de ce travail provient de la PDB « Protein Data Bank ». Il s'agit de la plus grande archive de données structurales de macromolécules biologiques permettant de consulter et de télécharger gratuitement les structures 3D des protéines biologiques via son site web : <http://www.rcsb.org/pdb/>. [59]



Figure 24 Interface d'accueil de PDB

2.2. Pubchem

Cette chimiothèque a été utilisée pour la recherche et le téléchargement de l'inhibiteur de la 3CLpro. Les différentes conformations tridimensionnelles des modèles proposés par la PubChem sont téléchargeables sous l'extension *.sdf*. On peut consulter cette chimiothèque via le lien : <https://pubchem.ncbi.nlm.nih.gov/>. [42]



Figure 25 Interface d'accueil de Pubchem

3. Les programmes

3.1. Surflex version 1.3

Le docking moléculaire a été réalisé à l'aide du programme **Surflex version 1.3 (2005)**. Nous a permis de réaliser un docking moléculaire rapide afin d'arrimer les ligands dans un environnement constitué d'acides aminés avec une bonne précision. [70]

```

C:\M19>surflex-dock rms 6y2f-1.mol2 final-0.mol2
Surflex License: C:\sfdocklicense
The Surflex software is the property of Biopharmics LLC.
Surflex licensed to SFDock0036 (Gerhard Vergoten, Univ. Lille) for internal use
only.
Rotatable bonds: 1 1
Actual rmsd: 7.968
Heavy atom rmsd: 4.965
Similarity: 0.355
-0.49] [ret_conf -0.42] [ret_conf -0.28] [ret_conf -0.52] [ret_conf -0.55] [ret
_conf -0.42] [ret_conf -0.37] [ret_conf 0.43] [ret_conf 0.43] [ret_conf -0.28] [
ret_conf -0.42] [ret_conf -0.48] [ret_conf 0.27] [ret_conf -0.42] [ret_conf 0.43
] [ret_conf -0.55] [ret_conf -0.55] [ret_conf -0.37] [ret_conf -0.37] [ret_conf
0.46] [ret_conf 0.28] [ret_conf 0.43] [ret_conf 0.27] [ret_conf 0.43] [ret_conf
0.27] [ret_conf 0.46] [ret_conf -0.18] final-0: 0.51 (crash -0.13 (self -0.00))
final-1: 0.51 (crash -0.13 (self -0.00))
final-2: 0.50 (crash -0.17 (self -0.00))
final-3: 0.50 (crash -0.17 (self -0.00))
final-4: 0.45 (crash -0.19 (self -0.00))
final-5: 0.45 (crash -0.19 (self -0.00))
final-6: 0.45 (crash -0.19 (self -0.00))
final-7: 0.37 (crash -0.32 (self -0.00))
final-8: 0.27 (crash -0.23 (self -0.00))
final-9: 0.27 (crash -0.23 (self -0.00))

```

Figure 26 Interface de logiciel Surflex

3.2. Open Babel

Le programme « Open Babel » version (2.0.2) est un programme libre qui nous a facilité la conversion des données biologiques d'un format à un autre en tenant compte

d'une protonation adéquate au pH physiologique. [71] Les différents formats que « Open Babel » prend en charge comprennent : pdb, mol, mol2, sdf, smi...etc.

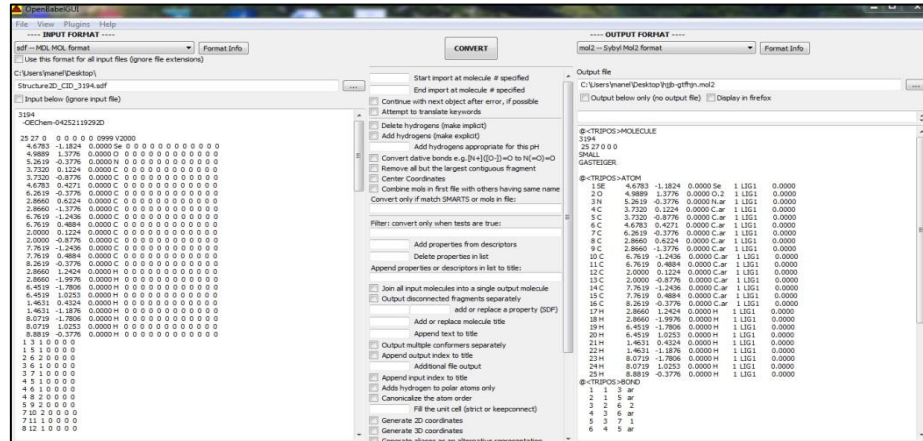


Figure 27 Interface de logiciel Open Babel

3.3. ArgusLab

Est un programme de docking moléculaire et de dessin de divers ligands, il permet également de séparer les complexes (protéine-ligand) téléchargés de la PDB. [72] Il utilise le format *.pdb* et n'a pas besoin d'une préparation préalable. [70]

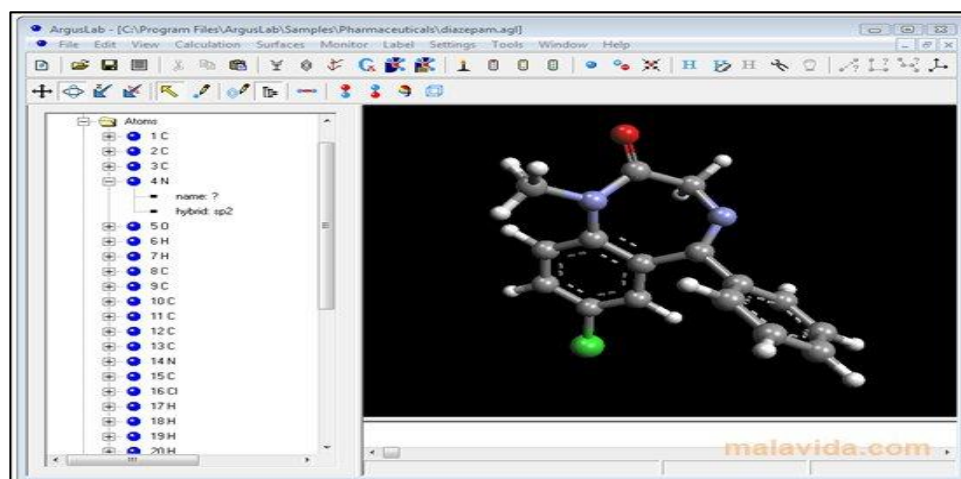


Figure 28 Interface de logiciel ArgusLab

3.4. ViewerLite4.2

Il est utilisé, dans sa version la plus récente 4.2.0, pour la visualisation des interactions formées entre les différents inhibiteurs et le site actif de l'enzyme étudiée.

[73]

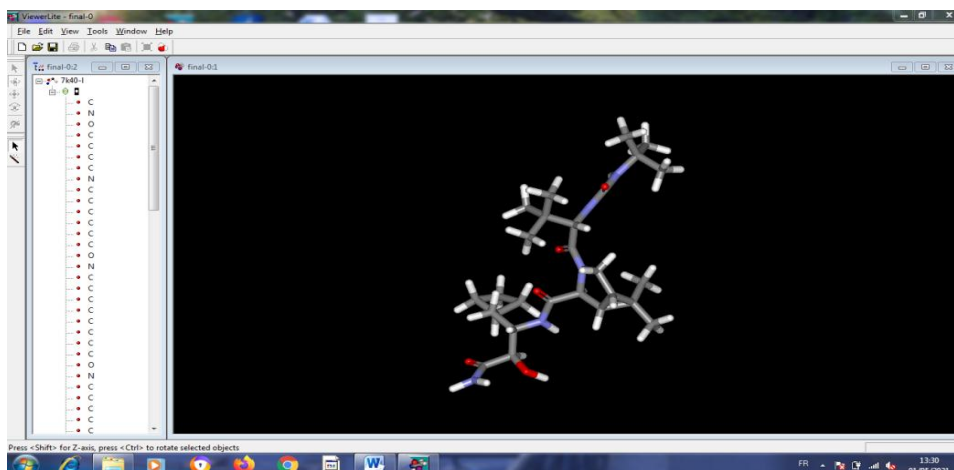


Figure 29 Interface de logiciel ViewerLite

3.5. Biovia Discovery studio visualizer

Est une suite de programmes fournissant des fonctionnalités pour visualiser et analyser des données biologiques et chimiques. [74]

Il a été utilisé pour la visualisation des résultats obtenus, la génération des diagrammes 2D des interactions ligand-cible et la construction des ligands.

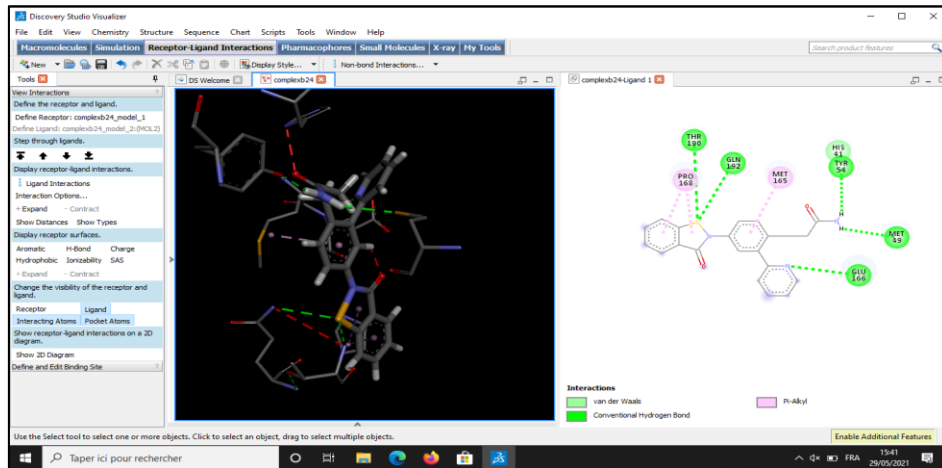


Figure 30 Interface de logiciel Biovia discovery studio visualizer

3.6. Excel Microsoft 2010

Excel est un logiciel de la suite bureautique Office de Microsoft. Il permet de créer facilement des tableaux, des graphes...etc. Ce logiciel a été utilisé pour générer la courbe de corrélation linéaire.

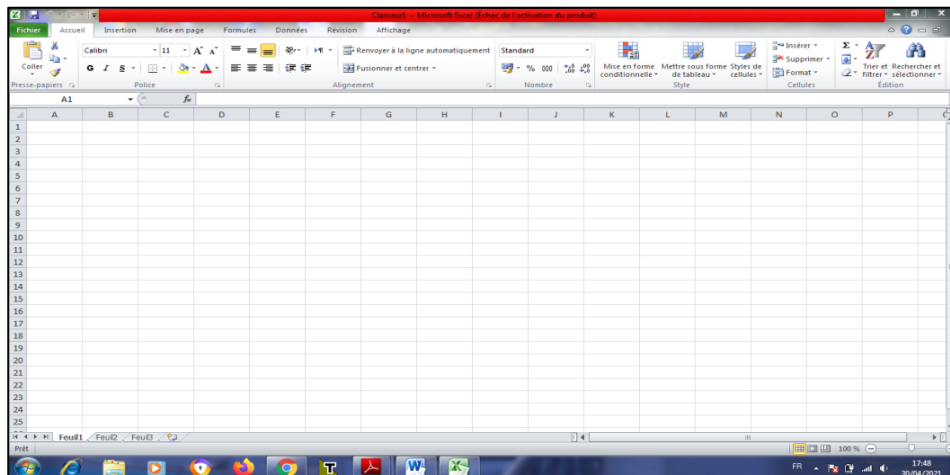


Figure 31 Interface de l'excel

4. Outils Web

4.1. Serveur ADME/T lab

Est une plate-forme en ligne permettant aux utilisateurs d'effectuer une analyse de druglikeness, la prédiction des points finaux ADMET, l'évaluation systématique et la recherche similarité, ce qui donne une compréhension globale d'un lot de composés et peut réaliser le processus de criblage rapide. On peut consulter ce serveur via le lien : <http://admet.scbdd.com/>. [75]

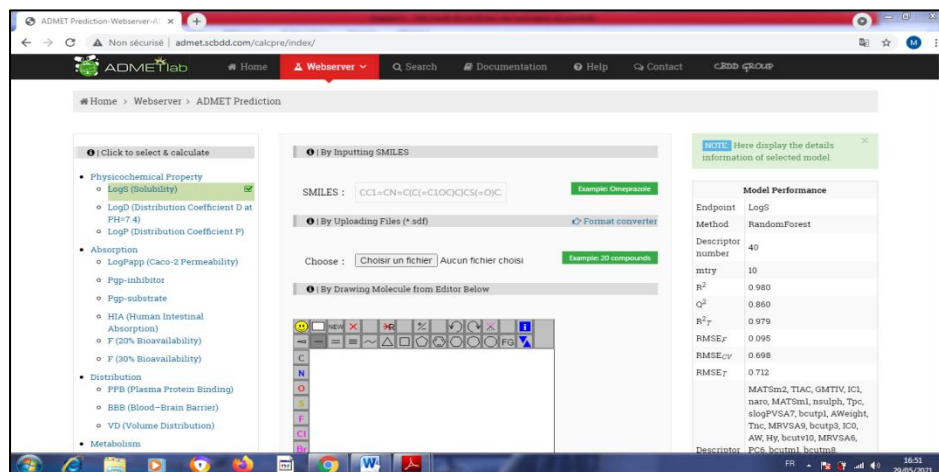


Figure 32 La page d'accueil du serveur ADMETlab

4.2. PreADMET

Est une application web qui permet une prédiction rapide des propriétés physico-chimiques et des propriétés médicamenteuses. Elle fournit des informations numériques utiles concernant l'absorption, la distribution, le métabolisme, l'excrétion et la toxicité (ADMET) d'un composé chimique. [76] Dans le présent travail, ce serveur nous a été utile pour la prédiction de la toxicité potentielle de nos inhibiteurs d'intérêt.

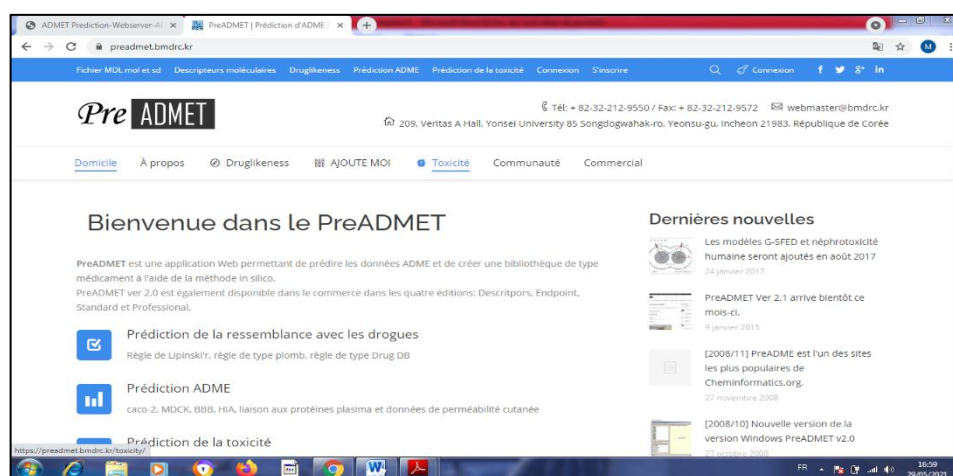


Figure 33 La page d'accueil du serveur PreADMET

II. Méthodes

1. Docking moléculaire

Le docking moléculaire vise à prédire la structure d'un complexe moléculaire à partir des molécules extraites, dans lesquelles différentes approches sont combinées pour étudier les modes d'interaction entre deux molécules. Les logiciels de docking sont donc des outils très utiles en biologie, pharmacie et médecine, car la plupart des principes actifs sont de petites molécules (ligand) qui interagissent avec une cible biologique d'intérêt thérapeutique. Le récepteur macromoléculaire étant le plus souvent une protéine, le terme docking seul est couramment employé pour désigner un « docking protéine-ligand ». [56]

Tests d'évaluation de la fiabilité du docking

Il est indispensable d'étudier l'efficacité des logiciels d'arrimage moléculaire afin de valoriser et crédibiliser les résultats fournis le long de notre travail. Pour cela, nous avons tenté d'évaluer la performance de logiciel, Surflex. L'évaluation de la fiabilité comportera trois tests de fiabilité :

- ❖ le premier concerne le **coefficient de corrélation (r)**.
- ❖ le deuxième, le **RMSD**.
- ❖ Le troisième l'**analyse visuelle**.

1.1. Coefficient de corrélation linéaire (r)

La corrélation est une quantification de la relation linéaire entre des variables continues. Le coefficient de corrélation c'est une mesure de l'association ou de la relation entre deux variables. Cette standardisation permet d'obtenir une valeur qui variera toujours entre -1 et +1, peu importe l'échelle de mesure des variables mises en relation. Le test de corrélation (contrairement à la régression simple) ne propose pas d'identifier une variable dépendante et une variable indépendante. On ne cherche qu'à déterminer l'absence ou la présence d'une relation linéaire significative entre les variables. [77]

Il existe diverses façons de mesurer la relation ou la liaison, selon qu'il s'agisse de variables quantitatives. [78]

Notre étude de corrélation a concerné le score obtenu à partir du docking moléculaire par le logiciel : Surflex et l'activité biologique traduite par IC_{50} , et elle s'est étalée sur 30 complexes tirés de la PDB.

2.2. Le RMSD ou l'écart quadratique moyen

L'aptitude d'un algorithme à trouver l'emplacement correct du ligand par rapport à son récepteur est habituellement déterminée au moyen de la Déviation Quadratique Moyenne ou RMSD entre le modèle du ligand simulé par docking moléculaire vis-à-vis de celui de référence (déterminé expérimentalement). La valeur seuil admise est une différence maximale de 2 Å [68] [69].

Dans notre cas, le test de fiabilité du programme Surflex par RMSD a été réalisé sur 100 complexes protéine-ligand pris de manière arbitraire via la PDB.

- ❖ Les calculs de RMSD des 100 complexes étudiés sont présentés dans l'annexe.

2.3. L'analyse visuelle

L'analyse visuelle faisant suite au test par RMSD est une étape complémentaire pour vérifier la fiabilité des simulations. Cette étape nous permet de vérifier la superposition du modèle de chaque ligand simulé vis-à-vis de sa pose de référence.

2. Le choix de la cible

La 3CLpro est la cible thérapeutique que nous avons étudiée dans notre travail, sa structure tridimensionnelle est disponible sous format pdb via la banque de données PDB. Cette dernière nous a proposé plusieurs complexes de structures cristallographiques de protéase principale (3CLpro) dont 9 meilleurs sont représentés dans le tableau.

Code pdb	Code ligand	Résolution
1Q9U	MRD	2.37 A°
1UK4	ATO	2.50A°
2V6N	MES	1.98A°
6Y2G	O6K	2.20 A°
6Y2F	O6K	1.95 A°
6LU7	PRD	2.16A°
6WTT	B15	2.15A°
7BFB	9JT	2.05A°
7NBR	USG	2.40 A°

Tableau 4 : code PDB des complexes 3clpro /inhibiteur avec leur résolution.

La résolution de la protéine est une des données reflétant la qualité des structures ayant permis de construire le modèle cristallographique. Une résolution proche de 1A° permet de distinguer tous les atomes y compris les hydrogènes. [79]

C'est en se basant sur ce critère qu'on a choisi de mener notre étude sur le complexe 6Y2F de la protéase principale du SARS-cov2 avec le tert-butyl-3.4-dioxo-1-2-oxo-1,2 carbamate (alpha-acétamide 13b) (O6K) ayant la plus petite résolution (1.95A°) et en plus c'est la cible du notre inhibiteur de référence à partir de l'article sur lequel nous avons choisi de travailler [80] , la protéine a été ensuite téléchargée sous format *pdb* et préparée grâce au programme Arguslab.

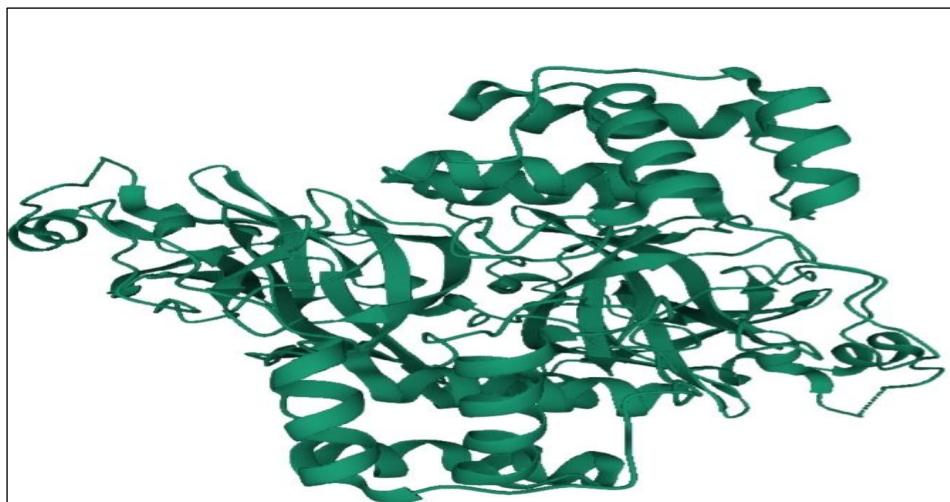


Figure 34 Structure 3D du complexe 6Y2F issu de la PDB

3. Préparation des molécules au docking moléculaire

3.1. Préparation de la cible

Un fichier de structure PDB typique n'est pas adapté pour une utilisation immédiate. Une préparation précoce de la structure protéique à l'aide d'Arguslab est indispensable avant de passer au docking moléculaire. La préparation de la cible consiste à éliminer les molécules d'eau, le ligand de la cristallisation. La cible est enregistrée dans un fichier au format *pdb* et ensuite convertie en *mol2* à l'aide d'OpenBabel

3.2. Construction des ligands

Les ligands utilisés dans ce travail ont été construits grâce au programme Biovia Discovery Studio Visualizer. Ce dernier nous a permis de faire la construction de 63 inhibiteurs de main protéase issus de la littérature chaque inhibiteur a subi par la suite une optimisation de sa géométrie grâce à la fonction de minimisation d'énergie avant de l'enregistrer sous format *mol2*.

3.3. Les étapes du docking par Surflex

Nous avons réalisé les étapes du docking moléculaire par le logiciel Surflex selon le protocole suivant [81] :

1ère étape : **génération du protomol**

L'identification du site actif par la construction d'une pseudo-molécule appelée « protomol » est une étape réalisée grâce à la commande suivante :

```
Surflex-dock proto ligand.mol2 protéine.mol2 pl
```

❖ 2ème étape : **le docking moléculaire**

Cette étape nécessite la disponibilité de la protéine format .mol2, le protomol généré dans l'étape précédente et les ligands sous format .mol2. Le docking moléculaire par «Surflex» permet d'avoir les dix (10) meilleures positions (conformations) de chaque ligand au sein du site actif étudié, classées par leur score (Final-0, Final-1Final-9). Cette étape a été réalisée en utilisant la commande suivante :

```
Surflex-dock dock ligand.mol2 pl-protomol.mol2 protéine.mol2
```

❖ 3ème étape : **calcul du RMSD**

Le RMSD considéré est celui des atomes lourds ; Cette étape a été réalisée en utilisant la commande suivante :

```
Surflex-dock rms ligand.mol2 final-0.mol2
```

❖ 4ème étape: **optimisation du RMSD**

Cette étape est supplémentaire pour optimiser quelques valeurs du RMSD. Cette étape a été réalisée en utilisant la commande suivante :

```
Surflex-dock oprms ligand.mol2 final-0.mol2
```

4. Modification par substitution

La modélisation moléculaire sert à construire de nouvelles molécules ou le remplacement de quelques groupements chimiques par d'autres. Ces substitutions ont été réalisées dans le but de découvrir de nouveaux inhibiteurs plus efficaces. L'étude de la modélisation moléculaire par substitution permet d'améliorer l'affinité de manière significative par un choix judicieux des fragments à substituer et de proposer de nouveaux inhibiteurs meilleurs que notre ligand de référence.

5. Prédiction ADMET

Outre une affinité qui lui confère une activité biologique, un candidat médicament doit être facile à administrer et doit atteindre sa cible en concentration suffisante pour être efficace. C'est pourquoi la prédiction des propriétés ADMET joue un rôle crucial dans le processus de la conception des médicaments. [82]

ADME est une abréviation pour « Absorption, Distribution, Métabolisme, Excrétion », elle décrit la disposition d'un composé pharmaceutique au sein d'un organisme. [83] Alors que le « T » représente la toxicité potentielle de ce composé.

Dans ce travail, nous avons calculé et simulé les propriétés physico-chimiques, pharmacocinétiques et la toxicité des composés issus de la substitution en utilisant les serveurs ADMETlab et PreADMET.

5.1. Propriétés physico-chimiques

5.1.1. Règle de Lipinski

Christopher A. Lipinski a défini un ensemble de critères permettant de déterminer si un composé chimique possède des propriétés physico-chimiques qui en feraient de lui un médicament oralement actif chez l'homme. Il observe que les molécules ayant le plus de chances d'être absorbées par voie orale doivent satisfaire au moins trois des caractéristiques suivantes [84] :

- 1) Le poids moléculaires doit être inférieur à 500 Da.
- 2) Le coefficient de partage octanol-eau (logP) doit être inférieur à 5.
- 3) Le nombre d'accepteurs de liaison hydrogène doit être inférieur ou égal à 10.

- 4) Le nombre de donneurs de liaison hydrogène doit être inférieur ou égal 5.
- 5) Le nombre de liaisons flexibles doit être inférieur à 15.

5.1.2. Règle de Veber

Avec un objectif similaires, Veber et al. Etendent la règle de Lipinski en ajoutant deux contraintes sur la surface polaire qui doit être inférieur à 140 \AA^2 , et le nombre de liaisons de rotation qui doit être inférieur ou égal à 10. [85]

5.1.3. LogP

Le log P est égal au logarithme du rapport des concentrations de la substance étudiée dans l'octanol et dans l'eau. $\text{Log P} = \log (\text{Coct}/\text{Ceau})$. Cette valeur permet d'appréhender le caractère hydrophile ou hydrophobe (lipophile) d'une molécule. Diverses études statistiques ont mis en évidence les valeurs optimales de log P pour qu'un composé (typiquement un médicament) soit absorbé par le corps humain :

- ❖ Pénétration au cerveau: 2,0
- ❖ Absorption orale: 1,8
- ❖ Absorption sublinguale: 5
- ❖ Absorption percutanée: 2,6 [84]

5.2. Propriétés pharmacocinétiques

5.2.1. BBB perméabilité

La capacité de nos meilleurs inhibiteurs à traverser la barrière hémato-encéphalique a été simulé à l'aide au serveur ADMETlab. Ce critère est important tant pour les médicaments psychotropes, qui doivent exercer leur action dans le système nerveux central, que pour les médicaments destinés aux organes périphériques. [86]

5.2.2. Inhibition des cytochromes P450

Les cytochromes P450 (CYP) sont des enzymes hépatiques intervenant dans le métabolisme des substrats endogènes ou exogènes, notamment médicamenteux. Les principaux CYP impliquées dans le métabolisme des médicaments sont : CYP1A2,

CYP2C9, CYP2C19, CYP2D6, CYP3A4. Toutefois, en fonction de leur degré d'affinité pour ces enzymes, certains médicaments peuvent inhiber le métabolisme. C'est pourquoi il vaut mieux choisir ceux qui ne sont pas inhibiteurs des cytochromes P450. [87]

5.3. Toxicité potentielle

Pendant le développement d'un candidat-médicament, la sécurité est toujours le problème le plus important, y compris une variété de toxicités et d'effets indésirables. [88]

L'évaluation de la toxicité d'un composé s'appuie sur des études qualitatives ou quantitatives adéquates. Dans cette étude, le serveur PreADMET a été utilisé pour simuler certains critères de toxicité selon les tests suivants :

5.3.1. Test d'Ames

Le test d'Ames permet de déterminer le potentiel mutagène d'un composé chimique. [89] Les souches de *Salmonella typhimurium* utilisées dans ce test sont des souches porteuses d'une mutation dans un des gènes gouvernant la synthèse de l'acide aminé histidine, ce qui rend ces souches incapables de pousser. Avec une fréquence très faible, ces mutations His-reversent spontanément vers His⁺ et donc les cellules retrouvent leur capacité à pousser sur un milieu dépourvu d'histidine. Cette fréquence de réversion peut augmenter en exposant les bactéries His⁻ à des agents mutagènes. Ainsi, le test d'Ames permet de quantifier l'induction de ces mutations réverses His⁺. [90]

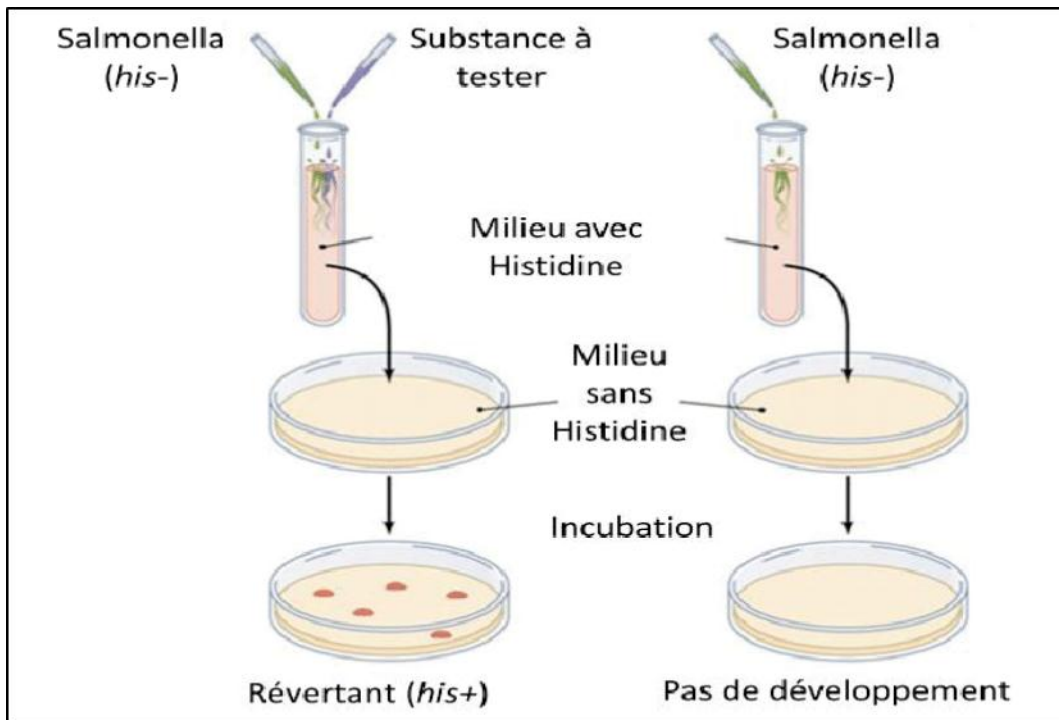


Figure 35 Schéma explicatif de test d'Ames [91].

5.3.2. Inhibition de hERG

Le gène hERG (human Ether-à-go-go-Related Gene) code pour un canal potassique générateur de tension qui est un composant clé dans la formation du potentiel d'action cardiaque. Le blocage pharmacologique du canal hERG provoque une mort subite [92]. La simulation des effets des composés sur le canal hERG offre des informations précieuses pour vérifier les effets secondaires proarythmiques potentiels [93].

5.3.3. la toxicité aigüe de *Daphnia*

Daphnia magna est un micro crustacé d'eau douce. L'une des raisons pour laquelle ce modèle biologique est couramment utilisé en toxicologie c'est qu'il est facilement cultivé en laboratoire grâce à son cycle de vie court et sa fécondité élevée [94] [95].

5.3.4. la toxicité aigüe sur les algues

Ce test permet d'évaluer la toxicité aigüe liée à l'utilisation du composé d'intérêt. Il consiste à déterminer le taux d'inhibition de la croissance des différentes espèces d'algues à savoir : (*Chlorella vulgaris*, *Selenastrum caprinortum* et *Scenedesmus subspicatus*) [96].

5.3.5. la toxicité aigüe sur le Medaka

Le médaka japonais, (*Oryziaslatipes*), est largement utilisé dans les tests de toxicité aigüe [97]. C'est un petit poisson d'aquarium approprié pour évaluer l'oestrogénicité des composés perturbateurs du système endocrinien [98].

5.3.6. la toxicité aigüe sur Minnow

Pimephales promelas est une espèce de poisson largement utilisée pour une grande variété d'applications de recherche portant sur des sujets tels que l'évaluation de toxicité aigüe. Les poissons sont petits et faciles à élever en laboratoire. Ils ont un cycle de vie relativement court, engendrant des centaines d'oeufs avec des taux de fertilisation et d'éclosion élevés [99].

Chapitre 5

Résultats et

Discussion

1. Test d'évaluation de la fiabilité des programmes de docking

Le docking est le moyen qui calcule l'ensemble des interactions suite à une orientation spatiale optimale du ligand ; assurant la stabilité du complexe et déclenchant un signal responsable de l'effet thérapeutique. Le calcul concerne l'intensité de ce signal qu'il traduit en un score basé sur la puissance de l'affinité ou l'énergie attribuée par l'ensemble des interactions.

Cette méthode nous a été utile dans notre quête de nouveaux inhibiteurs théoriques de 3CL protéase et vu son importance plusieurs logiciels sont développés parmi lesquels nous avons choisi : Surflex (V) dont la fiabilité a été évaluée par trois tests : le test RMSD et le test de corrélation et l'analyse visuelle pour vérifier leur reproductibilité théorique par rapport aux résultats expérimentaux.

1.1. Le test RMSD

Ce test permet d'estimer la possibilité du programme à reproduire l'orientation spatiale expérimentale du ligand à l'intérieur du site actif, donc ce test se base sur une comparaison après superposition de la conformation expérimentale du ligand de référence et celle prédite par le logiciel de docking après calcul des interactions, et qui convient à la conformation du ligand ayant la plus haute affinité (final-0) par Surflex.

La prédiction est approuvable seulement si la valeur du RMSD est inférieure ou égale à 2 Å.

L'évaluation de la fiabilité par le RMSD a été effectuée sur 100 complexes (protéine-ligand) que nous avons téléchargés de la PDB et regroupés dans le tableau de l'annexe.

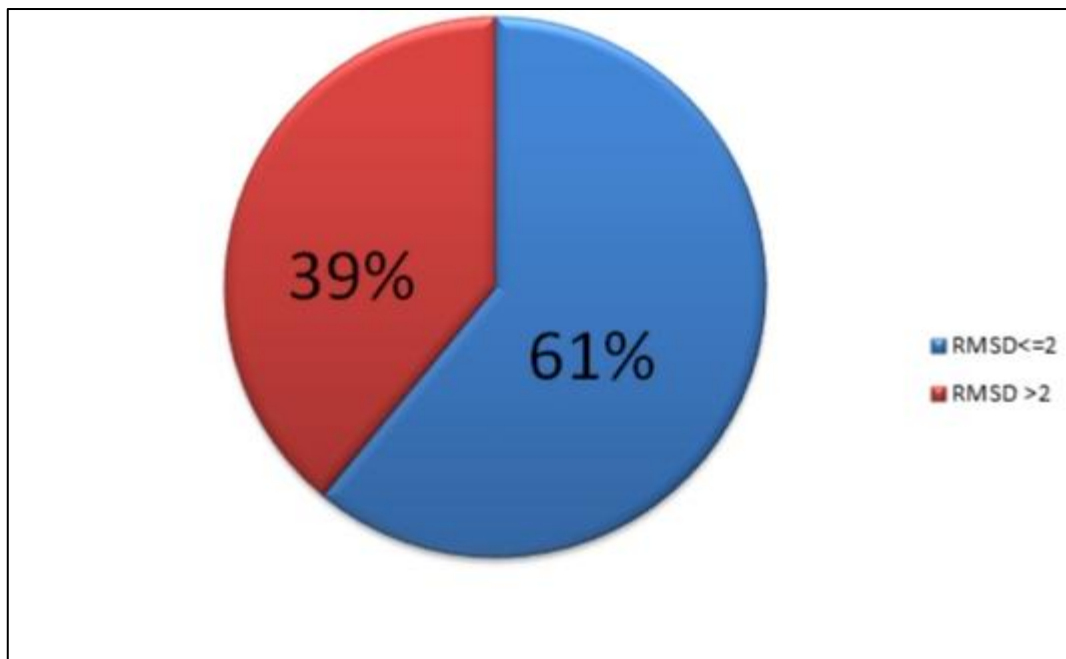


Figure 36 Résultats en pourcentage des valeurs de RMSD obtenus par « surflex »

1.1.1. Le test RMSD par le programme Surfex

Sur un total de 100 complexes, 61 molécules ont fourni un résultat positif avec un RMSD $\leq 2\text{Å}$, ce qui se traduit par une valeur de 61% des résultats obtenus.

Il est à noter que dans des études comparables des pourcentages meilleurs ont été trouvés : 63% à partir de 300 complexes (Chikhi. A. et Bensegueni. A., 2008) et 85,8% sur 134 complexes (Teniou. S., 2012). Le programme Surfex est un outil de modélisation moléculaire de haute performance.

1.1.2. L'analyse visuelle des résultats du RMSD

L'évaluation de la fiabilité par le test RMSD ne s'accomplit qu'après vérification que les ligands (expérimental et prédit) sont superposables atome par atome.

À propos des prédictions par le logiciel de docking : **Surflex**, nous avons choisi trois résultats de docking de différents complexes dont le premier est à RMSD inférieur à 2Å le deuxième est aux alentours de 2Å ; quant au troisième, il est nettement supérieur à 2Å .

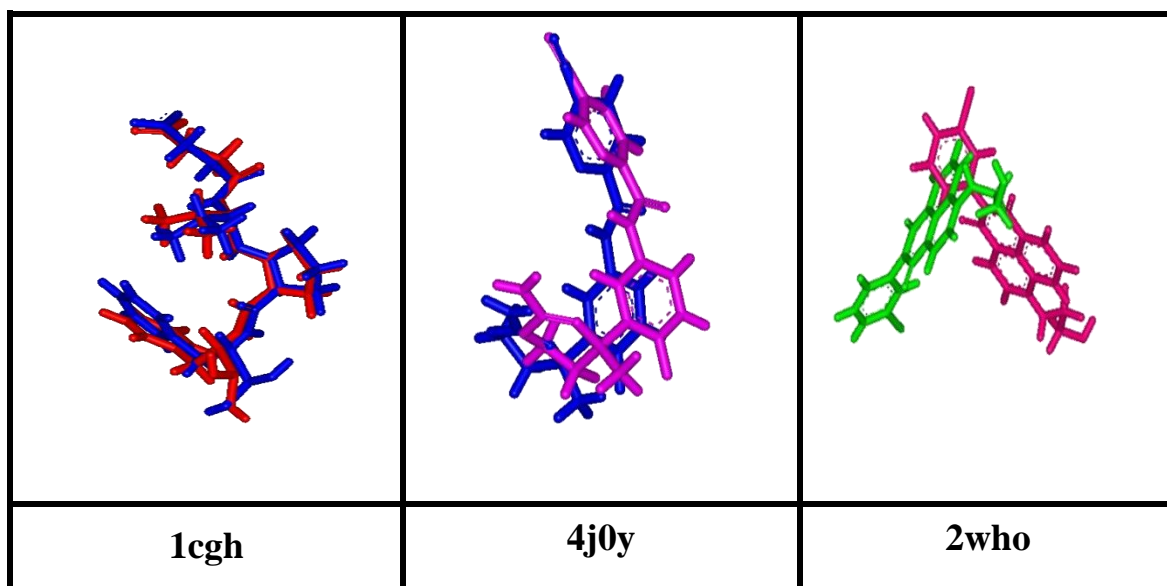


Tableau 4 Superposition des trois ligands des complexes (1cgh, 4j0y et 2who) et ceux calculés par docking moléculaire avec **Surflex**

N ^o	Complexe	RMSD	Description du RMSD	Description du ligand prédit
1^{er}	1cgh	0.839	Meilleur	Coloré en bleu
2^{ème}	4j0y	2.120	Intermédiaire	Coloré en mauve
3^{ème}	2who	6.041	mauvais	Coloré en rose

1.2. Le test du coefficient de corrélation linéaire

Un deuxième procédé qui permet l'évaluation de la fiabilité des logiciels de docking est le test de corrélation qui quantifie l'intensité de la relation linéaire entre les deux variables suivantes : l'affinité calculée par Surflex et l'IC₅₀ déterminée expérimentalement ; celle-ci relate la concentration qui inhibe 50% du processus enzymatique et juge l'efficacité d'un éventuel médicament.

Nous avons exprimé les IC₅₀ par log IC₅₀ afin de réduire l'échelle de la courbe de corrélation à tracer.

Le test de corrélation est un bon déterminant de la performance des logiciels de docking à travers le calcul du « **coefficient de corrélation r** » après traçage de la courbe, qui est réalisé sur 30 complexes.

Le tableau ci-dessous comprend les données de corrélation :

	Le complexe	IC ₅₀ (nM)	logIC ₅₀	Final-0(M ⁻¹)
01	1ZU4	70	1.84	4.86
02	1W22	176	2.24	4.98
03	1IA2	23	1.36	7.18
04	2Q93	560	2.74	2.20
05	2GG3	580	2.76	5.30
06	2GG5	250	2.39	4.56
07	2BDL	15	1.17	3.2
08	2G9X	45	1.73	4.05
09	2DUV	87	1.93	6.07
10	3FQK	81	1.9	6.62
11	3MPT	251	2.39	5.58
12	2UZB	120	2.07	5.3
13	3OVX	31	1.49	3.52
14	3S7M	10	1	7.17
15	3RK9	520	2.71	5.17
16	3RPU	70	1.84	5.92
17	3NZB	35	1.54	7.13
18	3PKA	540	2.73	3.41
19	3PKB	620	2.79	3.5
20	3QTZ	50	1.69	9.14
21	3S79	300	2.47	7.40
22	4NEU	10	1	12.83
23	4MKB	1	0	7.94
24	4MK8	3	0.47	10.84
25	4BBX	60	1.77	6.84
26	5NQR	70	1.84	3.28
27	5K00	7	0.85	13.19

28	6MDB	64	1.8	9.91
29	6MD7	181	2.25	4.94
30	6COX	50	1.69	7.07

Tableau 5 Données de l'analyse par la corrélation linéaire entre l'affinité obtenue par docking moléculaire (Surflex) et les $\log IC_{50}$

Après avoir regroupé les données expérimentales et obtenu les scores de docking par le logiciel, nous les avons représentés par une courbe de corrélation linéaire entre deux variables qui sont les $\log IC_{50}$ et les affinités en utilisant le programme « Excel » qui nous a affiché : l'équation, le coefficient de détermination « R^2 » et le coefficient de corrélation « r ».

1.2.1. L'analyse de corrélation des affinités obtenues par Surflex

Nous avons représenté ces résultats sur une courbe :

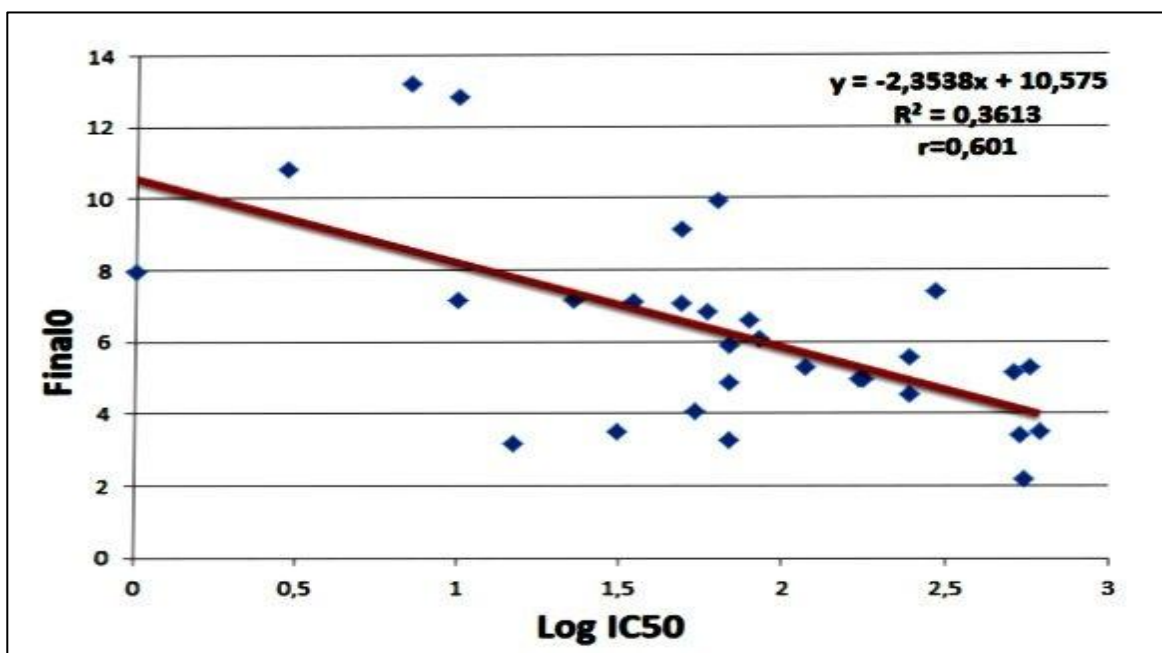


Figure 37 Courbe de corrélation entre $\log IC_{50}$ et final-0 obtenus par Surflex

❖ Commentaire

Le nuage de points se rapproche dans son ensemble de la droite, ce qui justifie la valeur du coefficient de corrélation $r = 0.601$. Il indique ainsi une corrélation négative (variables inversement proportionnelles dont les valeurs évoluent dans des sens opposés)

L'obtention d'un tel coefficient de corrélation prouve la performance du logiciel Surfex et cela en accord avec la valeur $r = 0.71$ obtenue par (Ait Hamoudi. L et Chami. S., 2018)

❖ Conclusion

Les résultats obtenus avec les tests de fiabilité : RMSD, l'analyse visuelle et la corrélation linéaire nous ont prouvé la performance du programme Surfex qui peut donc être utilisé pour l'étude *in silico* de l'inhibition de la 3CLpro par diverses molécules.

2. Etude et recherche de nouveaux inhibiteurs de la 3CLpro

➤ La structure de ligand de référence ebselen

L'ebselen (2-phényl-1,2-benzosélenazol-3 (2H) -one; PZ51; DR-3305) est l'inhibiteur d'une enzyme catalysant la réaction de la glutathion peroxydase [41]. La principale protéase du coronavirus SARS-CoV-2 est une cible médicamenteuse potentielle. L'IC₅₀ de l'ebselen est de 0,67 μM . [80]

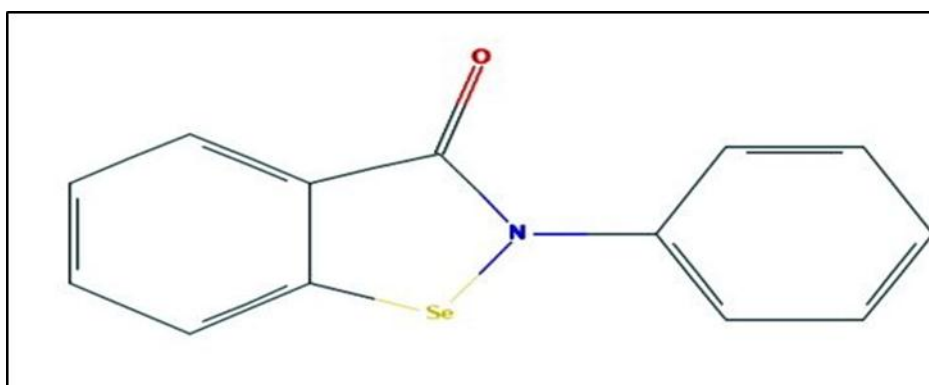


Figure 38 Structure de l'inhibiteur Ebselen

2.1. Etude des interactions enzyme-inhibiteur

2.1 Etude des interactions ebselen /3CLpro

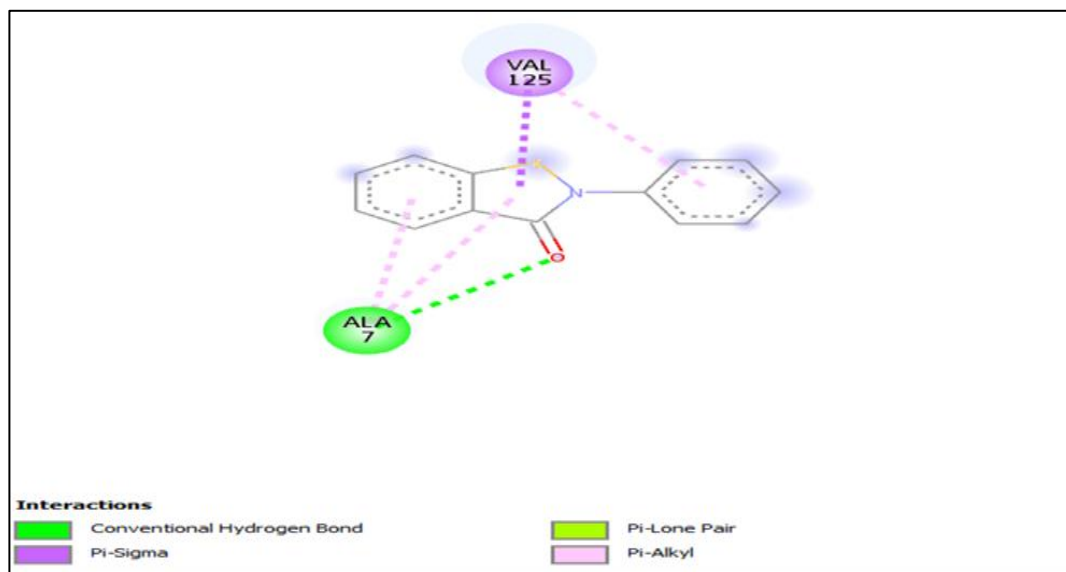


Figure 39 Diagramme 2D des interactions de composé ebselen/3CLpro avec $K_{d0}=3.51M^{-1}$

L'analyse visuelle montre la présence d'une seule liaison hydrogène entre l'oxygène du ligand et le résidu de l'acide aminé Ala-7 du site actif avec une distance de 2.38Å.

De plus, plusieurs interactions hydrophobes participent à la stabilité de l'ebesen au sein du site actif de l'enzyme. Ces interactions font intervenir les résidus : Ala-7 et Val-125.

2.2. Essais de modélisation moléculaire

Pour cela, nous avons effectué 2 types de substitutions : **mono substitution** et **bi substitution**

2.2.1 Mono-substitution

Dans un premier temps, le ligand (ebselen) a subi 63 mono-substitutions sur des différents radicaux (voir **figure 40**) afin d'améliorer son affinité. Les fonctions ajoutées, et les final0 de chaque composé issu de la mono-substitution sont représentés dans le tableau N°6

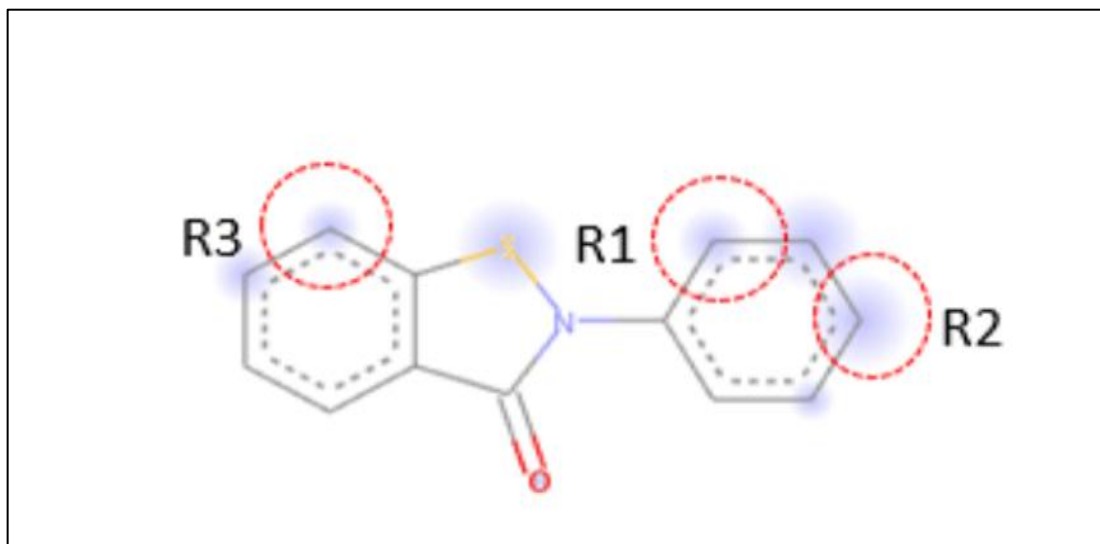


Figure 40 Représentation des radicaux R1.R2.R3 de la structure

De base d'inhibiteur ebselen

Cette étude consiste à faire varier la structure de départ de l'ebselen en vue d'améliorer son affinité ($3,51 M^{-1}$) envers la cible d'une part et son profil ADMET d'autre part. Pour cela, plusieurs groupements fonctionnels présentant la capacité d'engager des liaisons hydrogène avec la cible ont été introduits sur la structure de départ au niveau des différentes positions R1, R2 et R3.

La substitution réalisée par l'ajout d'un groupement méthyle (CH₃), d'un hydroxyle (OH), d'une fonction amine (NH₂), des groupements thiol, Chlore, Nitrosyle, phosphate et iode, des fonctions (CH₂)₂ amide, (CH₂)₂ cétone, (CH₂)₂acide, (CH₂)₂inicine, (CH₂)₂ amine, (CH₂)₂ Phosphatase acide. Des cycles aromatiques benzène, Pyrimidine, Pyridine, 1,3thiazole, Imidazole et un Cyclopentane à la place des hydrogènes liés au niveau des positions R1, R2 et R3 de l'inhibiteur. Le dessin des molécules a été réalisé par l'intermédiaire du programme **Biovia discovery studio**.

Tableau N6 les structures et les affinités des composés issus de la mon-substitution

Composé monosubstitué	R1	R2	R3	Final0 (M ⁻¹)
M1	OH			2.99
M2		OH		4.12
M3			OH	2.19
M4	COOH			1.54
M5		COOH		3.41
M6			COOH	3.49
M7	CH3			3.08
M8		CH3		1.36
M9			CH3	2.76
M10	NH2			2.66
M11		NH2		4.28
M12			NH2	2.58
M13	NO			3.12
M14		NO		3.13
M15			NO	4.69
M16	SH			2.49
M17		SH		3.73
M18			SH	3.47
M19	PH2			2.18
M20		PH2		4.68
M21			PH2	3.46
M22	I			3.56
M23		I		1.34
M24			I	1.61
M25	CL			3.59
M26		CL		2.10
M27			CL	3.03
M28	Benzène			3.51
M29		Benzène		4.18
M30			Benzène	3.03
M31	Cyclopentane			2.26
M32		Cyclopentane		3.64
M33			Cyclopentane	4.04
M34	Pyridine			2.63

M35		Pyridine		3.56
M36			Pyridine	2.72
M37	Pyrimidine			2.46
M38		Pyrimidine		2.61
M39			Pyrimidine	1.96
M40	1,3thiazole			2.59
M41		1,3 thiazole		2.12
M42			1,3thiazole	2.02
M43	Imidazole			2.50
M44		Imidazole		5.30
M45			Imidazole	3.85
M46	(CH ₂) ₂ -CO-NH ₂			3.91
M47		(CH ₂) ₂ -CO-NH ₂		5.32
M48			(CH ₂) ₂ -CO-NH ₂	3.96
M49	(CH ₂) ₂ -NH ₂			3.31
M50		(CH ₂) ₂ -NH ₂		3.20
M51			(CH ₂) ₂ -NH ₂	2.50
M52	(CH ₂) ₂ -COOH			5.24
M53		(CH ₂) ₂ -COOH		4.15
M54			(CH ₂) ₂ -COOH	3.50
M55	(CH ₂) ₂ COCH ₃			4.14
M56		(CH ₂) ₂ COCH ₃		5.05
M57			(CH ₂) ₂ COCH ₃	4.38
M58	(CH ₂) ₂ CH=NH			2.83
M59		(CH ₂) ₂ CH=NH		5.14
M60			(CH ₂) ₂ CH=NH	4.36
M61	(CH ₂) ₂ -O-PO ₃ H ₂			5.18
M62		(CH ₂) ₂ -O-PO ₃ H ₂		4.21
M63			(CH ₂) ₂ -O-PO ₃ H ₂	4.14

➤ Résultats et discussion

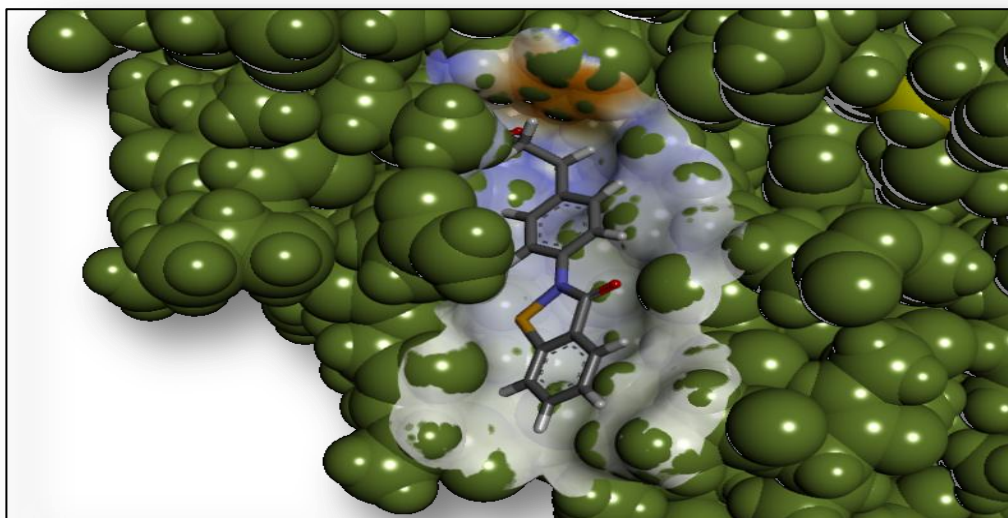


Figure 41 le positionnement du composé M47 dans le site actif de 3CLpro

La mono substitution concerne 63 composés et a permis une amélioration de l'affinité pour 6 d'entre eux : 15, 20, 44, 47, 53, et 61 ; la meilleure (5.32 M^{-1}) étant celle du composé **M47**, qui présente un groupement $(\text{CH}_2)_2\text{-CO-NH}_2$ au niveau de R2.

Le remplacement de l'atome hydrogène en position R2 par un groupement $(\text{CH}_2)_2\text{-CO-NH}_2$ a permis d'améliorer l'affinité du composé M47, provoquant la formation de nouvelles liaisons.

Les liaisons	Mode de liaison de l'inhibiteur avec les résidus du site actif	Le type de liaison	Distance
01	NH (N°34) - - - OH (N°397) de l'AA Tyr54	Conventional hydrogen bond	2.04Å
02	NH (N°34) - - - O (N°1424) de l'AA Asp187	Conventional hydrogen bond	2.07Å

03	NH (N°35) - - - - SD (N°353) de l'AA Met 49	Conventional hydrogen bond	2.48Å
04	Se (N°14)- - - - N (N°1449) de l'AA Thr190	Conventional hydrogen bond	3.68Å
05	Se (N°14) - - - - NE2 (N°1469) de l'AA Gln192	Conventional hydrogen bond	3.50Å
06	Le cycle benzène-----SD (N°1262) de l'AA Met 165	Pi-sulfur	4.10Å

Tableau N°7 les liaisons établies par le composé M47 avec le site actif.

L'analyse visuelle montre la formation de quatre nouvelles liaisons hydrogènes complètement différentes de celles de la molécule de référence, donc au total cinq liaisons hydrogène. La première, la deuxième et la troisième sont formées entre les atomes d'hydrogènes du groupement NH₂ du composé 47 et les résidus : Tyr-54, Asp-187 et Met-49 du site actif avec des distances de 2.04Å, 2.07Å et 2.48Å respectivement. La quatrième et cinquième liaison sont formées entre Se du composé M47 et les résidus de Thr-190 et Gln-192 du site actif avec des distances de 3.68Å et 3.50 Å respectivement.

Il existe également une liaison de type pi-sulfure formée entre le cycle benzène du ligand et les résidus de l'Aa Met-165 à distance de 4.10Å.

En plus des interactions hydrophobes réalisées par l'alkyl hydrocarboné apolaire que nous retrouvons entre : Les cycles du ligand **M47** et les radicaux des aminoacides Pro168 avec deux distances 3.78Å avec le cycle de benzène et la deuxième avec autre cycle de composé **M47** à distance de 5.39Å.

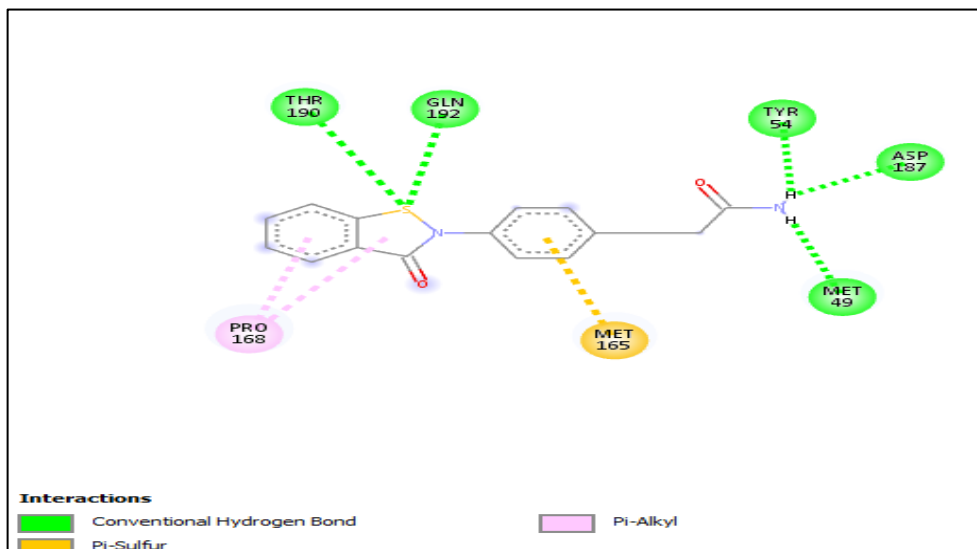


Figure 42 Diagramme 2D des interactions de composé M47/3CLPRO

2.2.2. Bi substitution

Dans la deuxième étape, le composé **M47** a subi 42 bi-substitutions au niveau de différents radicaux (voir figure) afin d'améliorer son affinité.

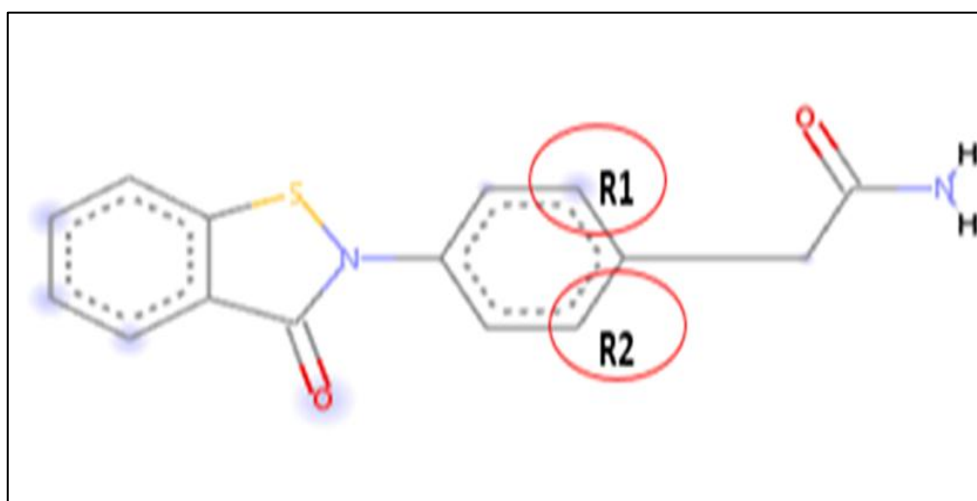


Figure 43 Représentation des radicaux R1.R2 du composé m47 issu
De la mono- substitution

Les fonctions ajoutées, et les final-0 de chaque composé issu de la mono-substitution sont représentés dans le tableau N°8

Tableau N°8 : les structures et les final-0 des composés issus de la bi-substitution

Composé monosubstitué	R1	R2	Final0 (M ⁻¹)
B1	OH		5.53
B2		OH	5.48
B3	COOH		4.57
B4		COOH	4.58
B5	CH3		5.29
B6		CH3	5.73
B7	NH2		3.76
B8		NH2	5.52
B9	NO		5.75
B10		NO	5.89
B11	SH		5.36
B12		SH	3.60
B13	PH2		3.39
B14		PH2	3.60
B15	I		4.45
B16		I	3.74
B17	CL		5.13
B18		CL	4.98
B19	Benzène		5.17
B20		Benzène	4.10
B21	cyclopentane		4.90
B22		cyclopentane	4.82

B23	pyridine		5.92
B24		Pyridine	6.02
B25	Pyrimidine		6.02
B26		Pyrimidine	6.21
B27	1.3thiazole		5.66
B28		1.3thiazole	4.96
B29	imidazole		4.96
B30		Imidazole	4.20
B31	(CH ₂) ₂ -CO-NH ₂		4.86
B32		(CH ₂) ₂ -CO-NH ₂	4.83
B33	(CH ₂) ₂ -NH ₂		5.36
B34		(CH ₂) ₂ -NH ₂	2.99
B35	(CH ₂) ₂ -COOH		5.62
B36		(CH ₂) ₂ -COOH	5.83
B37	(CH ₂) ₂ COCH ₃		4.86
B38		(CH ₂) ₂ COCH ₃	4.66
B39	(CH ₂) ₂ CH=NH		5.29
B40		(CH ₂) ₂ CH=NH	4.21
B41	(CH ₂) ₂ -O- PO ₃ H ₂		5.89
B42		(CH ₂) ₂ -O- PO ₃ H ₂	5.63

➤ Résultats et discussion

Parmi les 42 composés issus de la bi-substitution, neuf ont donné des affinités supérieures à 5.32 M^{-1} , les 2 meilleurs étant le composé **B24** possédant un cycle pyridine au niveau du R2 et le composé **B26** qui présente un cycle pyrimidine au niveau du R1.

➤ Le Composé B24

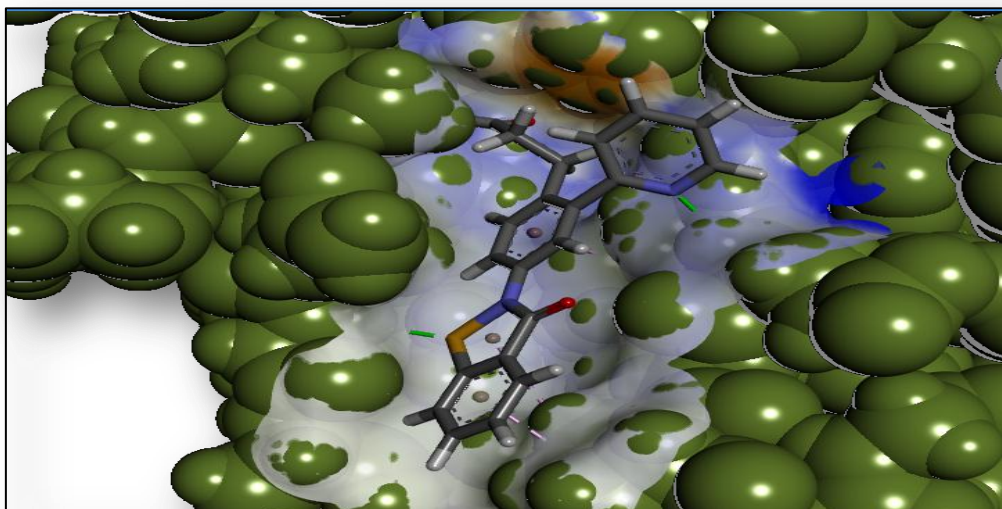


Figure 44 le positionnement du composé B24 dans le site actif de 3CLpro

Les liaisons	Mode de liaison de l'inhibiteur avec les résidus du site actif	Le type de liaison	Distance
01	NH (N°34) - - - -OH (N°397) de l'AA Tyr54	Conventional hydrogen bond	2.14Å
02	NH (N°35) - - - -SD (N°353) de l'AA Met 49	Conventional hydrogen bond	2.50Å
03	NH (N°14) - - - -N (N°1264) de l'AA Glu166	Conventional hydrogen bond	3.12Å
04	Se (N°14) - - - -NE2 (N°1469) de l'AA Thr190	Conventional hydrogen bond	3.56Å

05	Se (N°14) - - -N (N°1449) de l'AA Gln192	Conventional hydrogen bond	3.51Å
06	NH (N°34)- - -N (N°303) de l'AA His41	Van der waals	2.46Å

Tableau N° 9 les liaisons hydrogènes établies par le composé **B24** avec le site actif

L'analyse visuelle montre la présence de cinq liaisons hydrogènes, la première et la deuxième sont formées entre l'atome d'azote N du groupement NH₂ du composé **B24** et les résidus : Tyr-54 et Met-49 du site actif avec des distances de 2.14Å et 2.50Å respectivement. La troisième liaison est formée entre l'azote de cycle pyridine et le résidu Glu-166 du site actif avec distance de 3.12 Å, la quatrième et la cinquième liaison sont formées entre Se du composé **B24** et les résidus de Thr-190 et Gln-192 du site actif avec des distances de 3.56Å et 3.51Å respectivement.

Il existe également deux interactions hydrophobes réalisées par les alkyles hydrocarbonés apolaires que nous retrouvons entre :

Les résidus Met-165 et le cycle de benzène de composé **B24** à distance de 3.85Å. et la deuxième est entre les résidus Pro168 avec deux distances 3.77 Å avec le cycle de benzène et la deuxième avec autre cycle de composé **B24** à distance de 5.33 Å, qui participent dans la stabilité l'inhibiteur au site actif de 3clpro

En plus, un nouveau type de liaison, la force attractive de van der waals a été formée entre l'atome d'hydrogène du groupement NH₂ du composé **B24** et l'acide aminé His-41 à une distance de 2.46Å.

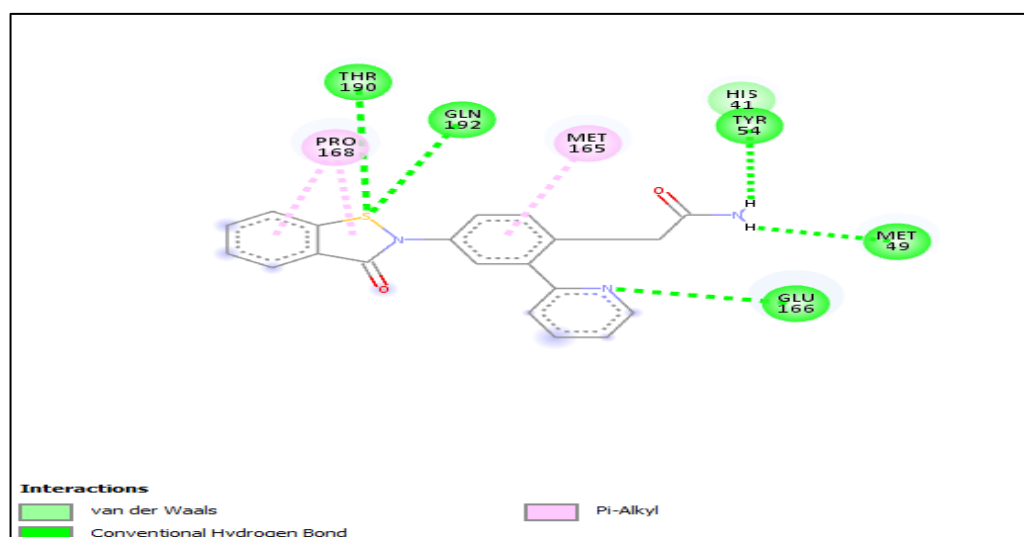


Figure 45 Diagramme 2D des interactions de composé B24/3CLPro

➤ **Le composé B26**

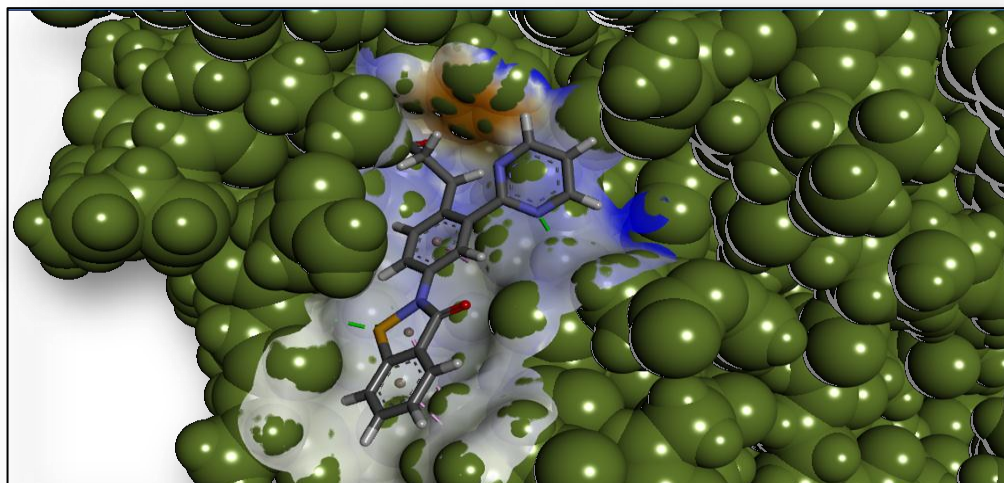


Figure 46 le positionnement du composé B26 dans le site actif de 3CLpro.

Les liaisons	Mode de liaison de l'inhibiteur avec les résidus du site actif	Le type de la liaison	Distance
01	NH (N°34) - - - OH (N°397) de l'AA Tyr54	Conventional hydrogen bond	2.14Å
02	NH (N°34) - - - O (N°1424) de l'AA Asp187	Conventional hydrogen bond	2.46Å
03	NH (N°35) - - - - SD (N°353) de l'AA Met 49	Conventional hydrogen bond	2.50Å
04	NH (N°14) - - - - N (N°1264) de l'AA Glu166	Conventional hydrogen bond	3.12Å
05	Se (N°14) - - - - NE2 (N°1469) de l'AA Thr190	Conventional hydrogen bond	3.56Å

06	Se (N°14) - - - N (N°1449) de l'AA Gln192	Conventional hydrogen bond	3.51Å
----	---	----------------------------	-------

Tableau N° 10 les liaisons hydrogènes établies par le composé B26 avec le site actif

L'analyse visuelle montre la présence de six liaisons hydrogènes donc il y a une formation d'une nouvelle liaison par rapport au **M47** et **B24**. La première, la deuxième et la troisième sont formées entre l'atome d'azote du groupement NH₂ du composé **B26** et les résidus : Tyr-54, Asp-187 et Met-49 du site actif avec des distances de 2.09Å, 2.12 Å et 2.50Å respectivement. La quatrième liaison est formée entre l'azote de cycle pyrimidine et le résidu de Glu-166 du site actif avec une distance de 3.04 Å, la cinquième et la sixième liaison sont formées entre Se du composé **B26** et les résidus de Thr-190 et Gln-192 du site actif avec des distances de 3.61Å et 3.53Å respectivement.

Il existe également deux liaisons hydrophobiques de type alkyl formées avec les résidus : Met-165 à distance de 3.81Å et Pro168 avec deux distances 3.52Å avec le cycle de benzène et la deuxième avec autre cycle de composé **B26** à distance de 5.33 Å, qui participent dans la stabilité de l'inhibiteur au site actif de 3CLpro.

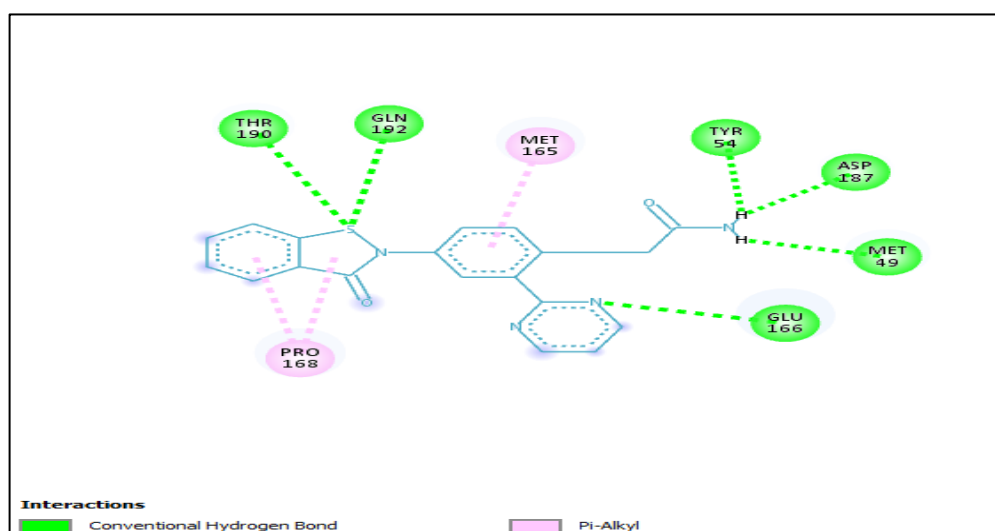


Figure 47 Diagramme 2D des interactions de composé B26/3CLpro

Remarque :

Après analyse visuelle des quatre composés nous avons remarqué un changement total dans le nombre et le type des liaisons formées par les trois composés **M47**, **B24** et **B26** par rapport à la molécule de référence.

3. Prédiction des propriétés ADMET

Il est estimé que près de 90% des candidats-médicaments échouent en raison d'une efficacité inacceptable ou en raison de leur toxicité. Il est donc devenu évident que les études ADMET jouent un rôle crucial dans le succès d'un candidat-médicament. C'est pourquoi nous avons jugé utile de prédire les propriétés ADME des composés **B24** et **B26** issus de la bi-substitution.

Les résultats obtenus ont été comparés à ceux d'une molécule de référence en l'occurrence l'ebseen

3.1. propriétés Physico-chimique

propriétés	Ebselen	B24	B26
Formule	C ₁₃ H ₉ NOSe	C ₂₁ H ₁₇ N ₃ O ₂ Se	C ₂₀ H ₁₆ N ₄ O ₂ Se
PM g /mol	274.18	422.346	423.334
Log P	2.048	2.528	1.923
Liaisons rotative	1	5	5
Accepteur	2	4	5
Donneur	0	1	1
TPSA²	22.00	77.98	90.87
Réfractivité molaire	66.48	107.231	105.026
Nombre totale d'atome	25	44	43

Nombre de liaisons Rigide	17	25	25
Règle de lipinski	100	100	100
Règle de veber	100	100	100
Règle Ghose	100	100	100
Règle Oprea	33.33	66.67	66.67

Tableau N° 11 propriétés physico-chimiques des composés obtenus à partir de serveur ADMETLab

À l'aide du serveur ADMETLab, nous avons simulé les propriétés physicochimiques des composés **B24** et **B26** et de l'abselen, la molécule de référence.

Les résultats du tableau N°11 montrent que, comme le ligand de référence, les composés **B24** et **B26** répondent bien aux règles de Lipinski et Veber, concernant le nombre de liaisons rotatives, accepteurs et donneurs d'hydrogène, la surface polaire TPSA...etc. indiquant qu'ils peuvent être administrés par voie orale sans soucis de biodisponibilité, ainsi qu'aux critères des règles proposées par le serveur ADMET Lab.

Le tableau N°11 montre aussi les bonnes valeurs de Log P et du poids moléculaire, de la réfractivité molaire de la molécule de référence et des deux composés, donc ces composés sont solubles et peuvent traverser facilement la membrane cellulaire.

Enfin une augmentation notable du pourcentage de règle d'oprea pour les deux molécules **B24** et **B26** (66.67%) par rapport à la molécule de référence (33.33%).

3.2. Propriétés pharmacocinétiques

Catégorie	Propriété	abselen	B24	B26
absorption	HIA	0.991	0.823	0.765
		(1)	(1)	(1)
distribution	BBB	0.991	0.8	0.905
		(1)	(1)	(1)

Métabolisme	CYT1A2inhibition	0.787 (1)	0.477 (0)	0.591 (1)
Métabolisme	CYT2C19inhibition	0.255 (0)	0.323 (0)	0.409 (0)
Métabolisme	CYT2C9inhibition	0.399 (0)	0.301 (0)	0.198 (0)
Métabolisme	CYT2D6inhibition	0.167 (0)	0.36 (0)	0.386 (0)

Tableau N°12 propriétés pharmacocinétiques des composés obtenus à partir
De serveur ADMETLab

➤ **Absorption**

Les propriétés ADMET, dérivées du serveur ADMETLab, révèlent que les composés B24 B26 et la molécule de référence, ont une HIA élevée (catégorie 1) indiquant que ces composés peuvent être bien absorbés par le tractus gastro-intestinal lors d'une administration orale. Les valeurs de HIA des composés se situent dans la plage de 0.991 à 0.765, cependant l'ebesen à une absorption intestinale légèrement meilleure.

➤ **Distribution**

Le test BBB présente un résultat positif pour les composés **B24** et **B26** (catégorie1), ce qui signifie que ces composés, comme la molécule de référence, ont une grande capacité à traverser la barrière hémato-encéphalique.

➤ **Métabolisme**

Le test d'inhibition des Cyt P450 présente un résultat positif pour les composés B26 et la molécule de référence (catégorie 1) concernant le cytochrome p450 (Cyt 1A2) par contre le composé **B24** n'inhibe aucune isoforme du cytochrome P450.

3.3. Les tests de toxicité

Paramètre	Ebselen	B24	B26
Teste d'Ames	Mutagène	mutagène	mutagène
hERG_inhibition	Risque moyen	Risque moyen	Risque moyen
Algae_at	0.0801787	0.0331885	0.0514568
Daphnia_at	0.111967	0.0326604	2.30793
Medaka_at	0.0194144	0.00242778	3.99508
Minnow_at	0.0159324	0.0066676	2.67469

Tableau N°13 le profil de toxicité relative des composés candidats obtenus

À partir de serveur PreADMET.

À l'aide du serveur PreADMET, nous avons simulé la toxicité potentielle des composés **B24** et **B26** et l'ebsele.

La prédiction du test d'Ames montrent que les trois composés ebselen, **B24** et **B26** sont inducteurs de mutation génétique. Il s'est avéré également que les trois composés possèdent un risque moyen pour l'inhibition du gène hERG qui, rappelons-le, peut causer des problèmes cardiaques. Il est également à souligner que les faibles valeurs obtenues de la stimulation de la toxicité aiguë sur les quatre espèces marines suggèrent que le composé **B24** possède une plus faible toxicité aiguë relativement à la molécule de référence.

➤ Conclusion :

L'étude théorique que nous avons menée sur l'inhibition de la 3clpro a révélé après une série de substitution et analyse des différents paramètres ADMET, que les composés

B24 et **B26** peuvent être proposés comme des nouveaux inhibiteurs de l'enzyme protéase principale. Ces composées ont présenté les meilleurs affinités ainsi qu'un profil ADMET acceptable pour une application biologique. Cependant, d'autres tests *in vitro* et *in vivo* sont nécessaires pour leur validation en tant que médicaments potentiels.

Conclusion

L'objectif primordial de notre travail est de développer *in silico* par le biais du docking moléculaire de nouvelles structures agissant comme inhibiteurs plus puissants de la main protéase 3CLpro, une cible attractive pour le développement de médicaments antiviraux efficaces contre le SARS-Cov-2.

Afin d'affirmer que Surfex est un programme adapté à notre étude, nous avons essayé de tester la capacité de ce programme à prédire l'emplacement correct d'un ligand au niveau de son récepteur. Pour cela, nous avons fait appel au test du RMSD (RootMean Square Deviation) qui sert à évaluer la déviation de la géométrie du ligand simulée par Surfex par rapport à sa conformation de référence qu'on retrouve dans la PDB. Sur les 100 complexes pris arbitrairement de la PDB, 61% ont présenté un RMSD inférieur à 2 Å. ce test a été accompagné par une analyse visuelle des complexes formés.

Le deuxième test utilisé est celui du coefficient de corrélation linéaire qui existe entre les valeurs expérimentales de l'IC₅₀ des molécules étudiées et leurs affinités simulées par le docking moléculaire. Les résultats de la régression linéaire donnent un coefficient de corrélation $r = 0.601$

Ces tests nous apportent la preuve que le programme Surfex est suffisamment performant et peut-être utilisé pour simuler les interactions d'autres inhibiteurs de la 3CLpro.

Dans un deuxième temps, nous avons réalisé 63 mono substitutions au niveau de la molécule de référence, l'ebsele et ensuite 42 bi-substitutions en vue d'augmenter l'affinité de diverses molécules envers notre cible. Ces essais de modélisation structurale ont permis une amélioration conséquente de l'affinité de plusieurs d'entre elles, les meilleures étant le composé mono substitué (**N°47**) qui a atteint une affinité de 5.32 M^{-1} et les deux composés bi substitués (**N°24** et **N°26**) avec des affinités de 6.02 M^{-1} et 6.21 M^{-1} respectivement.

Enfin, l'étude des propriétés physicochimiques, pharmacocinétiques et de toxicité potentielle montre que le composé **B24** présente le meilleur profil ADMET, par son métabolisme et sa faible toxicité aiguë, contrairement au composé de référence et au composé qui a la plus grande valeur d'affinité (**B26**).

Bien qu'ils s'avèrent intéressants, les résultats obtenus dans cette étude restent des prédictions théoriques réalisées *in silico*. Il serait toutefois souhaitable de réaliser

une application expérimentale *in vitro* et/ou *in vivo* pour vérifier davantage leur fiabilité et efficacité.

L'amélioration de l'affinité par d'autres modifications additionnelles telles qu'une tri-substitution ou l'utilisation d'autres outils informatiques plus performants pour l'optimisation et la vérification de la crédibilité de ces résultats est une piste assez intéressante qu'il faudrait explorer dans l'avenir.

Références
Bibliographiques

Les références

- [1]. Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; et *al.* A geometric approach to macromolecule-ligand interactions. *J Mol Biol*, 1982. **161**(2): p. 269-88.
- [2]. A. Arrault, “Stratégies de docking-scoring assistées par analyse de données. Application au criblage virtuel des cibles thérapeutiques COX-2 et PPAR gamma,” Thèse de doctorat d’université : Modélisation Moléculaire et Chimoinformatique. Université d’Orléans., 2008.
- [3]. Aurélien G, “Conception d’un logiciel de docking et applications dans la recherche de nouvelles molécules actives,” Thèse de doctorat d’université: Joseph Fourier, Grenoble, France., 2015.
- [4]. J. Cui, F. Li, Z.-L. Shi Origine et évolution des coronavirus pathogènes *Nat Rev Microbiol*, 17 (3) (2019), pp. 181 – 192.
- [5]. C.-C. Lai, T.-P. Shih, W.-C. Ko, H.-J. Tang, P.-R. Hsueh Syndrome respiratoire aigu sévère coronavirus 2 (SRAS-CoV-2) et maladie à virus corona-2019 (COVID-19): l’épidémie et les défis *Agents antimicrobiens Int J*, 105924 (2020).
- [6]. Organisation WH. Tests de laboratoire pour la maladie à coronavirus 2019 (COVID-19) dans les cas humains suspects: directives provisoires, 2 mars 2020. Organisation mondiale de la Santé, 2020.
- [7]. C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao une nouvelle épidémie de coronavirus préoccupante pour la santé mondiale *Le Lancet* (2020).
- [8]. Y. Shu, J.J.E. McCauley GISAID: Initiative mondiale sur le partage de toutes les données sur la grippe - de la vision à la réalité *Euro Surveill.* , 22 ans (2017).
- [9]. Chen, Y. , Liu, Q. , et Guo, D. (2020 , avril). Coronavirus émergents: structure du génome, réplication et pathogénèse. *Journal of Medical Virology*, 92 (4), 418 - 423. doi: 10.1002 / jmv.25681.
- [10]. Arden, KE , Nissen, MD , Sloots, TP et Mackay, IM (2005). Nouveau coronavirus humain, HCoV-NL63, associé à une grave maladie des voies respiratoires

inférieures en Australie. *Journal of Medical Virology*, 75 (3), 455 - 462. doi: 10.1002 / jmv.20288.

[11]. Wu, F., Zhao, S., Yu, B., Chen, Y.-M. , Wang, W., Song, Z.-G. , Hu, Y., Tao, Z.-W. , Tian, J.-H., Pei, Y.-Y. , Yuan, M.-L., Zhang, Y.-L. , Dai, F.

H., Liu, Y., Wang, Q.-M. , Zheng, J.-J., Xu, L., Homles, EC et Zhang, Y.-Z. (2020, mars). Un nouveau coronavirus associé à une maladie respiratoire humaine en Chine *Nature*, 579 (7798), 265 - 269 doi: 10.1038 / s41586-020-2008-3.

[12]. Juckel D, Dubuisson J, Belouzard S. les coronavirus, ennemis incertains. *Med sci (Paris)* 2020, 36 : 633-41.

[13]. Organisation WH. Transmission du SARS-Cov-2 : implications pour les précautions visant à prévenir l'infection : document d'information scientifique, 9 juillet 2020.

[14]. Organisation WH. Coronavirus disease covid-19 how is it transmitted, 13 December 2020.

[15]. Coronavirus : des recommandations sur l'utilisation des installations CVC en zones infectées. 23/03/2020.

[16]. Bourouiba L. Turbulent Gas Clouds and Respiratory Pathogen Emissions: Potential Implications for Reducing Transmission of COVID-19. *JAMA*. 2020. Doi : 10/ggqtj4.

[17]. Autorités américaines: le SARS-CoV-2 peut être transmis sur de plus longues distances. Mardi 6 octobre 2020.

[18]. Transmission par aérosol de Covid-19: une pièce, un bar et une salle de classe: comment le coronavirus se propage dans l'air | Société.29 octobre 2020.

[19]. COVID-19 : Modes de transmission et mesures de prévention et de protection contre les risques, incluant le rôle de la ventilation, 13 janvier 2021.

[20]. Jean-Luc Gala, Omar Nyabi, Jean-François Durant, Nawfal Chibani, Mostafa Bentahir. Méthodes diagnostiques du COVID-19. *Louvain Med* 2020 mai-juin ; 139(05-06) : 228-235.

[21]. Organisation WH. Maladie à coronavirus 2019 (COVID-19) : ce qu'il faut savoir. 10 novembre 2020.

[22]. P. C. Deluzarche Journaliste, « PCR », *Futura*. <https://www.futura-sciences.com/sante/definitions/genetique-pcr-91/>. (Consulté le juill. 09, 2021).

[23]. Eric Muraille. Covid-19 : comment fonctionnent les tests et quelles sont leurs utilités ? April 22, 2020.

[24]. Dr Akli BOUAZIZ. Test antigénique : définition, comment ça se passe ? 22 octobre 2020.

[25]. Organisation WH. Algeia Country Overview | world health Organization <https://www.who.int/countries/kaz> consulter le 16/6/2021

[26]. Organisation WH. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>. Consulter le 15/06/2021.

[27]. Annie Ladoux, Stéphane Azoulay et Christian Dani. Cibler la protéase majeure du SARS-Cov-2 pour fabriquer un médicament efficace contre ce coronavirus. 18 novembre 2020.

[28]. Li G., De Clercq E. Options thérapeutiques pour le nouveau coronavirus 2019 (2019-nCoV) *Nat Rev Drug Discov.* 2020; 19 (3): 149-150. doi: 10.1038 / d41573-020-00016-0.

[29]. Khan SA, Zia K., Ashraf S., Uddin R., Ul-Haq Z. Identification des inhibiteurs de protéase de type chymotrypsine de SRAS-CoV-2 via une approche informatique intégrée. *J. Biomol. Struct. Dyn.* 2020: 1–10.

[30]. Alessandro Grottesi, Neva Besker, Andrew Emerson, Candida Manelfi, Andrea R, Beccari, Francesco Frigerio, Erik Lindahl, Carmer Cerchia, Carmine Talarico. Etude informatiques de SARS-Cov-2 3CL pro : aperçus des simulations MD. 28 juillet 2020. Doi. Org/ 10.3390/ ijms 21155346.

[31]. Jager S, Cimermanic P, Gulbahce N, et al. Global landscape of HIV-human protein complexes. *Nature* 2011 ; 481 : 365–370.

[32] Hilgenfeld R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J* 2014 ; 281 : 4085–4096.

- [33]. Thiel V., Ivanov KA, Putics Á. Mécanismes et enzymes impliqués dans l'expression du génome du coronavirus du SRAS. *J Gen Virol.* 2003; 84 (9): 2305-2315. doi: 10.1099 / vir.0.19424-0.
- [34]. Wu C., Liu Y., Yang Y. Analyse des cibles thérapeutiques pour le SRAS-CoV-2 et découverte de médicaments potentiels par des méthodes informatiques. *Acta Pharm Sin B.* 2020; 10 (5): 766–788. doi: 10.1016 / j.apsb.2020.02.008.
- [35]. Juckel D, Dubuisson J, Belouzard S. Les coronavirus, ennemis incertains. *Med Sci (Paris)* 2020, 36: 633–41.
- [36]. Zhang L., Lin D., Sun X. La structure cristalline de la protéase principale du SRAS-CoV-2 fournit une base pour la conception d'inhibiteurs améliorés de l' α -cétamide. *La science.* 2020; 368 (6489): 409–412. doi: 10.1126 / science.abb3405.
- [37]. bioRxiv. <https://www.biorxiv.org/content/10.1101/2020.02.26.964882v2>
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), 926–935. Doi: 10. 1063/144586.
- [38]. Lu, I.-L., Mahindroo, N., Liang, P.-H., Peng, Y.-H., Kuo, C.-J., Tsai, K.-C., Hsieh, H.-P., Chao, Y.-S., & Wu, S.-Y. (2006, August 24). Structure-based drug design and structural biology study of novel nonpeptide inhibitors of severe acute respiratory syndrome coronavirus main protease. *Journal of Medicinal Chemistry*, 49(17), 5154–5161. doi:10.1021/ jm060207o.
- [39]. Cintia A.Menéendez, Fabian Byléhn, Gustavo R.perez-lemus, Walter Alvarado et Juan J. de Pablo Caractérisation mléculaire de l'activité de liaison d'ebseen à la protéase principale du SRAS-cov-2 DOI :10.1126 /sciadv.abd0345.
- [40]. Gi UK Jeong, Hanra Chanson, Gun Jeune Yoon, Doyoun Kim, Young-Chan Kwon. Stratégies thérapeutiques contre le COVID-19 et caractérisation structurelle du SARS-Cov-2 : un examen. 14 juillet 2020. Doi.org/10.3389/fmicb.2020.01723.
- [41]. Müller A., Cadenas E., Graf P., Sies H. Un nouveau composé séléno-organique biologiquement actif - I. Activité de type glutathion peroxydase in vitro et capacité antioxydante de PZ 51 (Ebselen) *Biochem. Pharmacol.* 1984; 33 : 3235–3239.

[42]. PubChem [En ligne] <http://pubchem.ncbi.nlm.nih.gov>.

[43]. Chunlong Ma , Michael Dominic Sacco , Brett Hurst , Julia Alma Townsend , Yanmei Hu , Tommy Szeto , Xiujun Zhang , Bart Tarbet , Michael Thomas Marty , Yu Chen , et Jun Wang . Le bocéprévir, le GC-376 et les inhibiteurs de la calpaïne II, XII inhibent la réplication virale du SARS-CoV-2 en ciblant la principale protéase virale. 2020 août; 30 (8): 678–692. Doi: 10.1038 / s41422-020-0356-z.

[44]. NC Pedersen , Y. Kim , H. Liu , AC Galasiti, Kankanamalage , C. Eckstrand , WC Groutas , M. Bannasch , JM Meadows , K.-O. Chang, efficacité d'un inhibiteur de protéase de type 3C dans le traitement de diverses formes de péritonite infectieuse féline acquise. *J. Feline Med. Surg.* 20, 378 - 392 (2018).

[45]. Y. Kim , H. Liu , AC Galasiti, Kankanamalage , S. Weerasekara , DH Hua , WC Groutas , K.-O. Chang, NC Pedersen, Inversion de la progression de l'infection mortelle à coronavirus chez le chat par un inhibiteur de protéase de coronavirus à large spectre. *PLOS Pathog.* 1, e1005531 (2016).

[46]. Dementiev A .; Joachimiak A .; Nguyen H .; Gorelik A .; Illes K .; Shabani S .; Gelsomino M .; Ahn EE; Nagar B .; Doan N. (2019) Mécanisme moléculaire d'inhibition de l'acide céramidase par Carmofur . *J. Med. Chem.* 62, 987–992. 10.1021 / acs.jmedchem.8b01723.

[47]. Empereur-Mot C. Développement d'outils statistiques d'évaluation de méthodes de criblage virtuel : courbes de prédictivité & Screening Explorer. Thèse de doctorat d'université : Bioinformatique. Paris : Conservatoire National Des Arts et Métiers. France. 2017. 206p.

[48]. G.L. Warren, C. Webster Andrews, A.M. Capelli, B. Clark, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff, M.S. Head. A critical assessment of docking programs and scoring functions. *Journal of Medicinal Chemistry* 49 (2006) 5912-5931.

[49]. I. Helperin, B. Ma, H. Wolfson, R. Nussinov. Principles of dockings : an overview of search algorithms and a guide to scoring functions. *PROTEINS: Structure, Function and Bioinformatics* 47 (2002) 409-443.

[50]. Mahdjoub Youcef. Développement d'une application distribuée en utilisant la plateforme Jini « Application docking moléculaire ». Département d'informatique. Université des Science et des Technologies d'Oran Mohamed Boudiaf, Date, 100.

[51]. L. Hoffer, "Développement et validation du logiciel S4MPLE : application au docking moléculaire et à l'optimisation de fragments assistée par ordinateur dans le cadre du fragment-based drug design. " Thèse de doctorat d'université : Chimie/Chémoinformatique. Université de Strasbourg, France., 2013.

[52]. El Haj Said K. Contribution à l'étude de l'inhibition d'enzyme par des tripodes pyrazoliques par modélisation moléculaire. Mémoire de Master en chimie Organique Université de Tlemcen. Algérie, 2016, 50 p.

[53]. Lebbad F. Etude par modélisation moléculaire des mécanismes de complexation. Thèse de Doctorat en chimie Théorique et Modélisation moléculaire Université Abou Bekr Belkaid. Tlemcen. Algérie, 2016, 96 p.

[54]. Bouchagra S. Modélisation des interactions protéine-petites molécules : Etude de la relation structure-fonction dans le cas des lipases. Thèse de doctorat en chimie organique et Bioorganique : Université Badji Mokhtar-Annaba. Algérie, 2018, 145 p.

[55]. Daoud I., "Modélisation de l'interaction des ions de métaux lourds avec des colorants," Thèse de doctorat d'université : Chimie Théorique et Modélisation Moléculaire. Université Abou Bekr Belkaïd Tlemcen, Algérie., 2015.

[56]. Lalout W, "Contribution a la modélisation des interactions dans les biomolécules : Cas de la Maladie de Parkinson," Thèse de doctorat d'université : Chimie Théorique et Modélisation Moléculaire. Université Abou Bekr Belkaïd Tlemcen, Algérie., 2016.

[57]. Renvez G, "Application des modes statiques à l'étude de la flexibilité des protéines : vers un processus de docking," Thèse de doctorat d'université : Nanophysique. Université Toulouse III - Paul Sabatier, France., 2010.

[58]. Mokrani E. H., "Contribution à l'amélioration de l'activité biologique des inhibiteurs de la dipeptidyl-peptidase 4 dans le diabète de type 2 par simulation informatique .," Thèse de magister d'université: Technologie des explorations biochimiques. Université Mentouri Constantine, Algérie., 2012.

- [59]. RCSB PDB : Homepage [En ligne] <http://www.rcsb.org/>.
- [60]. Zouaghi K., Hamdi A. Contribution à la mise en évidence de nouveaux inhibiteurs de l'ECA par criblage virtuel. Mémoire de Master d'Université : Biochimie Moléculaire et Sant
- [61]. Chimiotheque-nationale.enscm.fr [En ligne] <http://chimiotheque-nationale.enscm.fr>.
- [62]. Welcome to ZINC, a free database of commercially-available compounds for virtual screening [En ligne] <http://zinc.docking.org>.
- [63]. Grosdidier A. Conception d'un logiciel de docking et applications dans la recherche de nouvelles molécules actives. Thèse de doctorat d'université : Pharmacie. Grenoble : Université Joseph Fourier-Grenoble. France. 2007. 90p.
- [64]. Lagarde N. Méthodes de criblage virtuel in silico : importance de l'évaluation et application à la recherche de nouveaux inhibiteurs de l'interleukine 6. Thèse de doctorat d'université : Bioinformatique. Paris : Conservatoire National Des Arts et Métiers. France. 2015. 316p.
- [65]. Bessadok M.A. La multiplicité de transport de la P-glycoprotéine : Etudes de modélisation comparative et de docking au sein de la famille des protéines ABC. Thèse de doctorat d'université : Bioinformatique Structurale. Paris : université pierre et marie curie. France.2011. 189 p.
- [66]. Morris G. M. Lim-Wilby M. Molecular Docking. Methods in molecular Biology (Clifton, N.j.). 2008 ;443 : 365-382. Doi : 10.1007/978-1-59745-177-2_19.
- [67]. A. Fortun, "Techniques de Modélisation Moléculaire appliquées à l'Etude et à l'Optimisation de Molécules Immunogènes et de Modulateurs de la Chimiorésistance.," Thèse de doctorat d'université : Sciences du Médicament. Université Joseph Fourier, France., 2008.
- [68]. Veith M., Hirst J D., Kolinski A et al. 1998. Assessing energy functions for flexible docking. J. Comput. Chem. 1998 ; 19 : 1612-1622.

- [69]. Gabb J., Jackson R.M., Sternberg M.J.E. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* 1997 ; **272** : 106-120.
- [70]. jayA., jainN. Surflex-Dock 2.1 : Robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J. Comput. Aided Mol. Des.* 2007, 21:281-306
- [71]. Open Babel: Main Page [En ligne] <http://openbabel.org/api/2.0.2/>.
- [72]. ArgusLab. [En ligne]. <http://www.arguslab.com/arguslab.com/ArgusLab.html>. Consulter le [25/6/2021]
- [73]. ViewerLite 4.2. [En ligne] <http://viewerlite.software.informer.com/4.2/>. Consulté le : [25-06-2021].
- [74]. Discovery Studio Visualizer – weizmann institute of science. [En ligne] <http://bip.weizmann.ac.il/toolbox/structure/ds.htm>. Consulter le [25/06/2021]
- [75]. Home-ADMETlab : ADMET Prédiction/ADMET Predictor/QSAR/ADMET Database. [En ligne] <http://admet.scbdd.com/>. Consulté le [25/06/2021]
- [76]. PreADMET. – OmicX. [En ligne]. <http://omictools.com/preadmet-tool>. Consulté le [25/06/2021].
- [77]. [En ligne] <http://spss.espaceweb.usherbrooke.ca/pages/stat-inferentielles/correlation.php>. Consulté le : [25/06/2021].
- [78]. Ulrike Held. 2010. Pièges des corrélations: les coefficients de corrélation de Pearson et de Spearman. *BIOSTATISTIQUE. Forum Med Suisse*; 10(38):652-653.
- [79]. Sari Mohamed S. Etude de l'interaction protéine-protéine par Modélisation Moléculaire Application aux Processus Biologiques. Thèse de doctorat en Chimie Physique. Université Abou bekr belkaid. Telemcen. 2013. 115p.
- [80]. Riddhidev Banerjee. Lalith Perera² et L. M. Viranga Tillekeratne. Principaux inhibiteurs potentiels de la protéase principale du SARS-CoV-2. 26 novembre, 3 mars 2021.

- [81]. Teniou S. Conception in silico de nouveaux inhibiteurs de la topoisomérase I de cellules cancéreuses, Technologie des explorations biochimiques. Constantine : Université Mentouri Constantine. Algérie. 2012. 95 p.
- [82]. Hazan C. Recherche d'inhibiteurs de haute affinité de l'ADN polymérase Beta par criblage virtuel et RMN. Thèse de doctorat d'université : biologie structurale. Université de Toulouse III – paul Sabatier. France. 2008. 249p.
- [83]. Gupta P. K. Illustrated Toxicology: With Study Questions. Elsevier, 1st Edition. 2018: 67-106. Doi : 10.1016/B978-0-12-813213-5.00002-X.
- [84]. Lipinski C. A. Lombardo F. Dominy B. W. Feeney P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Advanced Drug Delivery Reviews*. 1997 ; 23 ; 3-26. Doi : 10.106/s0169-409x(00)00129-0
- [85]. Veber D. F. Johnson S. R, Cheng H. Y, Smith B. R, Ward K. W, Kopple K. D. Molecular properties that influence the Oral Bioavailability of Drug Candidates. *Journal of medical chemistry*. 2002: 45: 2615-2623. Doi: 10.1021/jm020017n.
- [86]. Bodor N, Buchwald P. Recent advances in the brain targeting of neuropharmaceuticals by chemical delivery systems. *Advanced Drug Delivery Reviews*. 1999 ; 36 (2-3) : 229-254. Doi : 10.1016/s0169-409x(98)00090-8.
- [87]. Interactions médicamenteuses et cytochromes – agence nationale de sécurité du médicament (ansm) et de produits de santé. [En ligne]. [http://ansm.sante.fr/Dossiers/Interactions-medicamenteuses/Interactions-medicamenteuses-et-cytochromes/\(offset\)/1](http://ansm.sante.fr/Dossiers/Interactions-medicamenteuses/Interactions-medicamenteuses-et-cytochromes/(offset)/1). Consulté le: [25/06/2021].
- [88]. Yang H., Sun L., Li W. Liu G. Tang Y. In silico Prediction of chemical Toxicity for Drug Design Using Machine Learning Methods and Structural Alerts. *Frontiers in chemistry*. 2018 ; 6 : 30. Doi : 10.3389/fchem.2018.00030.
- [89]. Biotechnologie : le test d'Ames pour évaluer le pouvoir cancérogène d'une substance. – Technobio. [En ligne]. <http://www.technobio.fr/article-le-test-d-ames-43438456.html>. Consulté le [25/06/2021]
- [90]. De Meo M. Test d'Ames. – Gazette labo [En ligne].

<http://www.gazettelabo.fr/archives/pratic/1997/20mutatest.htm>. Consulté le [25/06/2021]

[91]. Déroulement d'un test d'Ames. Copyright © 2006 Pearson Education ... [En ligne] https://www.researchgate.net/figure/Deroulement-dun-test-dAmes-Copyright-C-2006-Pearson-Education-Inc_fig3_285061897. Consulté le [25/06/2021].

[92]. Y. Jing, A. Easter, D. Peters, N. Kim, and I. J. Enyedy, "In silico prediction of hERG inhibition," *Future Med. Chem.*, vol. 7, no. 5, pp. 571–586, 2015.

[93]. G. E. Kirsch *et al.*, "Variability in the measurement of hERG potassium channel inhibition: Effects of temperature and stimulus pattern," *J. Pharmacol. Toxicol. Methods*, vol. 50, no. 2, pp. 93–101, 2004.

[94]. F. Zeman, "Toxicité d'un mélange binaire sur la daphnie *Daphnia magna*," thèse de doctorat d'université: Evolution, Ecologie, Ressources génétiques, Paléontologie. Université Montpellier II, France., 2008.

[95]. D. M. M. Adema, "Daphnia magna as a test animal in acute and chronic toxicity tests," *Hydrobiologia*, vol. 59, no. 2, pp. 125–134, 1978.

[96]. Claudine Rast, "Etude de la toxicité d'effluents complexes à l'aide d'un test d'inhibition de luminescence bactérienne," Thèse de doctorat d'université: Centre des sciences de l'environnement. Université de Metz, Lorraine, France, 1983.

[97]. A. Tabata *et al.*, "Estrogenic influences of estradiol-17 β , p-nonylphenol and bisphenol-A on Japanese Medaka (*Oryzias latipes*) at detected environmental concentrations," *Water Sci. Technol.*, vol. 43, no. 2, pp. 109–116, 2001.

[98]. I. Joon Kang *et al.*, "Effect of 17 β -estradiol on the reproduction of Japanese medaka (*Oryzias latipes*)," *Chemosphere*, vol. 47, no. 1, pp. 71–80, 2002.

[99]. J. Zha, Z. Wang, N. Wang, and C. Ingersoll, "Histological alternation and vitellogenin induction in adult rare minnow (*Gobiocypris rarus*) after exposure to ethynylestradiol and nonylphenol," *Chemosphere*, vol. 66, no. 3, pp. 488–495, 2007.

Annexe

Annexe

Annexe : Les valeurs de RMSD des 100 complexes protéine/ligand étudiés

	Le complexe	Rmsd surflex	Rmsd optimisé
01	1bdr	3.449	1.228
02	1bp4	2.436	1.252
03	1btu	0.854	/
04	1cgh	0.799	/
05	1c1s	0.986	/
06	1eed	1.940	/
07	1fxy	0.731	/
08	1gfw	2.963	2.307
09	1gs4	0.356	/
10	1ia2	0.537	/
11	1ida	3.722	2.930
12	1jbd	2.131	/
13	1jkl	0.903	/
14	1js3	0.443	/
15	1k1n	3.473	2.605
16	1k1o	0.888	/
17	1kil	0.058	/
18	1kim	1.841	/
19	1mfi	1.111	/
20	1nhu	3.075	1.477
21	1os5	0.750	/
22	1s1w	2.621	2.000
23	1scw	2.859	0.977
24	1u9v	3.561	0.639
25	1u9x	1.929	/
26	1ypl	0.937	/
27	1yvz	0.384	/
28	1zvx	0.850	/
29	2aim	1.414	/
30	2aux	1.566	/
31	2bdl	3.166	2.066
32	2bnh	3.214	0.436
33	2brk	0.486	/
34	2fdv	2.706	1.122
35	2ghl	1.033	/
36	2rgp	0.482	/
37	2r4b	0.440	/
38	2o7v	1.169	/
39	2tnf	2.201	0.618

40	2who	6.120	/
41	2w12	1.161	/
42	2ypu	1.981	/
43	3bla	1.153	/
44	3c56	0.841	/
45	3c6u	0.693	/
46	3c52	0.391	/
47	3ftz	1.590	/
48	3ful	1.642	/
49	3kfz	2.994	1.777
50	3mpt	1.055	/
51	3of8	3.730	1.694
52	3of9	1.989	/
53	3ovx	2.244	0.306
54	3s7m	0.969	/
55	3tgg	2.745	/
56	3tso	0.540	/
57	3usn	3.924	1.491
58	3l5u	0.696	/
59	4acu	3.488	2.432
60	4btu	1.511	/
61	4ith	0.611	/
62	4j0y	2.120	0.045
63	4jpe	2.377	0.046
64	4ly1	1.931	/
65	4mk8	2.241	1.704
66	4mvh	0.907	/
67	4rnv	1.731	/
68	4v05	1.647	/
69	4z1h	1.749	/
70	5bwz	3.999	3.008
71	5dc5	0.980	/
72	5git	2.369	1.832
73	5hvt	0.956	/
74	5nkh	3.248	1.205
75	5n9k	0.820	/
76	5o1a	2.507	1.736
77	5o1d	3.084	1.920
78	5q00	0.814	/
79	5pzq	2.222	1.126
80	5uab	1.868	/
81	5uad	1.460	/
82	5uuy	1.992	/
83	5v5y	2.612	1.823
84	5vc5	1.526	/
85	5vc6	2.924	0.128
86	5v58	2.527	1.195

87	5wh5	2.511	0.341
88	6bbv	1.121	/
90	6hsk	0.983	/
91	6ind	2.801	2.094
92	6ink	0.842	/
93	6inm	2.948	2.012
94	6mbd	2.836	0.080
95	6mda	0.884	/
96	6mdb	0.427	/
97	6mdc	0.810	/
98	6md7	0.441	/
99	6md9	2.485	2.212
100	6oy5	1.911	/

Thème : ETUDE IN SILICO DE L'INHIBITION DE LA « 3-CHYMOTRYPSIN LIKE PROTEASE » (3CLpro) EN VUE DE COMBATTRE LE SYNDROME RESPIRATOIRE AIGUE SEVERE (SRAS)

Mémoire de fin de cycle pour l'obtention du diplôme de Master en **Biochimie Appliquée**

Résumé

La pandémie actuelle de **COVID-19** a déclenché des efforts mondiaux pour l'identification rapide des vaccins et des traitements antiviraux spécifiques. **3CLpro** joue un rôle clé dans le cycle de vie de la réplication et de la maturation du coronavirus. Le **Docking moléculaire** est une méthode souvent utilisée dans les phases initiales du processus de développement de nouveaux médicaments. Nous l'avons utilisé dans le but de concevoir de nouveaux composés à meilleur effet inhibiteur que « **Ebselen** » qui agit sur la cible **3CL pro**. Le programme de docking moléculaire **Surflex** a été développé pour aider à la mise au point des molécules à activité thérapeutique. Avec un taux de valeurs de **RMSD** inférieures ou égales à **2 Å** de **61%** et un **coefficient de corrélation** de **0,601**, les performances du programme **Surflex** sont avérées. Ce test a été suivi d'une analyse visuelle de la superposition des ligands de référence et ceux obtenus par docking moléculaire. Dans un deuxième temps, nous avons complété nos essais par une modélisation structurale en effectuant **63 mono substitutions** et **42 bi-substitutions**, parmi lesquelles quelques composés substitués ont eu des affinités améliorées. Le meilleur composé **n°47** parmi les mono substitués a atteint une affinité de **5.32 M⁻¹** et les meilleurs composés bi substitués **n°26** et **n°24** avec des affinités respectives de **6.21 M⁻¹** et **6.02M⁻¹**. Enfin, l'étude des propriétés physicochimiques, pharmacocinétiques et de toxicité potentielle montrent que le composé **B24** a présenté un meilleur profil **ADMET**, par son métabolisme, et faible toxicité aigüe, contrairement au composé de référence et au composé qui a une plus grande valeur d'affinité **B26**. Nous le proposons donc comme nouvel inhibiteur éventuel de la **3CL pro**.

Mot clé : COVID-19, 3CLpro, Docking moléculaire, Inhibiteur, Surflex, RMSD, ADMET

L laboratoire de recherche : Biochimie Appliquée

Jury d'évaluation:

Président du jury: Mr BENSEGUENI. A Pr. Université des Frères Mentouri Constantine1

Rapporteur : Mr CHIKHI. A Pr. Université des Frères Mentouri Constantine1

Examineur : Mr. MOKRANI E. H (Maitre de conférences ' B ' - UFM Constantine-1)

Date de soutenance : 14/07/2021

