



N° Ref :.....

**Centre Universitaire Abdelhafid BOUSSOUF- Mila**  
**Institut des Sciences et de la Technologie**  
**Département des Sciences de la Nature et de la Vie**  
**Mémoire préparé en vue de l'obtention du diplôme de**  
**Master**

**Domaine : Sciences de la Nature et de la Vie**  
**Filière : Sciences Biologiques**  
**Spécialité : Biochimie Appliquée**

**Thème :**

**Identification de nouveaux inhibiteurs de la PLprotéase  
en tant qu'une stratégie thérapeutique pour traiter le  
COVID-19 : criblage *in silico* et filtrage ADME-Tox**

**Présenté par :**

-BOUSSELLAH Ikram  
-GHERDA Lamis

**Devant le jury :**

<b>Présidente :</b> MERZOUG Amina	M.C.A Centre.Univ.A.Boussouf - Mila
<b>Promotrice :</b> BOUCHERIT Hanane	M.C.B Centre.Univ.A.Boussouf - Mila
<b>Examinatrice :</b> LALAOUI Meryem	M.A.B Centre.Univ.A.Boussouf - Mila

# *Remerciements*

*Avant tout, nous remercions Allah qui nous a donné la force et la puissance pour réaliser et achever ce travail.*

*Après un long parcours notre mémoire touche à sa fin. Nous tenons à exprimer nos sincères remerciements à ceux qui ont contribué à l'élaboration de ce modeste travail.*

*Nous tenons à remercier notre promotrice **Dr. BOUCHERIT Hanane**, Maître de conférences classe «B» au Centre Universitaire Abdelhafid Boussouf de Mila, Nous le remercions pour tout le soutien et l'encouragement qu'elle nous a donné, pour sa patience, sa simplicité et sa disponibilité durant notre préparation de ce mémoire.*

*Nos profonds remerciements à **Dr. MERZOUG Amina**, Maître de conférences classe «A» au Centre Universitaire Abdelhafid Boussouf de Mila Pour avoir acceptée de présider le jury.*

*Au **Dr. LALAOUI Meryem**, Maître-assistant classe «B» au Centre Universitaire Abdelhafid Boussouf de Mila, pour avoir acceptée d'examiner ce travail.*

*Et au final, à nos chers parents, famille, amies et proches qui nous ont Soutenues.*

*Merci pour vos encouragements.*

*Lamis et Ikram*

# *Dédicace*

*Je dédie ce travail :*

*A mon très cher père **GHERDA Mahiedine** :*

*Mon support dans ma vie, qui m'a appris m'a supporté et m'a dirigé, tu as toujours été à mes côtés pour me soutenir et m'encourager*

*A ma très chère mère **BOUROUSSA Fatîha** :*

*Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes côtés a toujours été ma source de force Pour affronter les différents obstacles Que ce travail traduit ma gratitude et mon affection.*

*A mes très cher frères **DJIHAD** et **AYOUB** et ma belle-sœur **DOUAA**.*

*Et mon cher binôme*

***BOUSSELLAH Ikram**. Puisse Dieu vous donne santé, bonheur, courage et surtout réussite.*

*A mon encadreur **BOUCHERIT Hanane**.*

*Et à tous mes collègues de la promotion 2021/2022 Biochimie Appliquée avec lesquels j'ai partagé mes moments de joie et de bonheur.*

*A toutes les personnes de ma grande famille **GHERDA**, et tous mes amis.*

## *Lamis*



# *Dédicace*

*Je dédie ce modeste travail en signe de reconnaissance et de respect, le premier amour de ma vie*

*A mon père Ziadî*

*A ma chère maman "Fatîma" qui m'a fait protéger toute ma vie et a été mon soutien et ma force*

*A mes très chers frères Tarek, Abdalhak et Islam a pour leur soutien et encouragement*

*A l'âme de ma chère grand-mère MADANI Chaa « Allah yarhemha »*

*A Mes cousines Halîma et Ibtîsam*

*A mon enseignant BOUBANDIR Tarek pour votre soutien et votre aide*

*A ma chère Dr. BOUCHERIT Hanane et mon binôme GHERDA Lamis, qui a donné tous les efforts pour terminer ce travail, Je te souhaite une vie remplie d'amour et de succès*

*A tous mes enseignants ainsi que tous mes collègues de promotion de master 2022.*

*A toute la famille BOUSSELLAH et la famille LAGUEB*

*A toutes les personnes qui m'ont encouragé à l'élaboration de ce mémoire.*

# *Akram*

# **Thème : Identification de nouveaux inhibiteurs de la PL protéase en tant qu'une stratégie thérapeutique pour traiter le COVID-19 : criblage *in silico* et filtrage ADME-Tox**

## **Résumé**

Le COVID-19 est une maladie respiratoire potentiellement mortelle causée par le virus SARS-COV-2. Il s'est propagé rapidement, d'abord à travers la Chine, puis à l'étranger, provoquant une épidémie mondiale.

Notre travail s'inscrit dans le cadre de rechercher *in silico* de nouveaux traitement du coronavirus en inhibant l'enzyme PL protéases (PLpro) avec de nouveaux inhibiteurs. Nous avons effectué un criblage virtuel basé sur la structure de la cible enzymatique PLpro de SARS-COV-2. Cette approche permet de simuler les interactions entre une protéine et des milliers de molécules dans le but de découvrir des composés ayant une activité inhibitrice plus élevée contre une cible thérapeutique.

Le programme Surflex-dock a été utilisé pour découvrir de nouveaux inhibiteurs plus puissants de la PLpro. Avec un pourcentage important de valeurs RMSD inférieures ou égales à 2 Å (88%), on peut dire que ce programme fonctionne bien. On peut donc s'en servir pour se plonger dans le mécanisme d'inhibition de PLpro sans trop de risque d'erreur. En effet, une collection de 555 molécules de la chimiothèque PubChem a été testée contre le site actif de la PLpro. Le criblage virtuel par Surflex-dock ressortir les composés CID149925965 et CID142139076 comme meilleurs inhibiteurs potentiellement plus actif envers la PLpro avec des affinités égales à  $9.52\text{M}^{-1}$  et  $8.20\text{M}^{-1}$ , respectivement.

Enfin, l'application de la règle de *Lipinski* nous renseigné de manière positive sur les propriétés pharmacocinétiques (ADME-Tox) de ces deux inhibiteurs proposés.

**Mots clés :** SARS-COV-2, PL protéases, Criblage Virtuel, RMSD, Surflex-dock, Règle de *Lipinski*.

# **Title: Identification of new PL protease inhibitors as a therapeutic strategy to treat COVID-19: *in silico* screening and ADME-Tox filtering**

## **Abstract**

COVID-19 is a potentially fatal respiratory disease caused by the SARS-COV-2 virus. It spread rapidly, first across China and then overseas, causing a global epidemic.

Our work is part of the *in silico* research of new coronavirus treatment by inhibiting the enzyme PL proteases (PLpro) with new inhibitors. We performed a virtual screening based on the structure of the SARS-COV-2 enzyme target PLpro. This approach makes it possible to simulate the interactions between a protein and thousands of molecules with the aim of discovering compounds with higher inhibitory activity against a therapeutic target.

The Surflex-dock program was used to discover new, more potent PLpro inhibitors. With a significant percentage of RMSD values less than or equal to 2 Å (88%), it can be said that this program works well. It can therefore be used to delve into the PLpro inhibition mechanism without too much risk of error. Indeed, a collection of 555 molecules from the PubChem chemical library was tested against the active site of PLpro. The virtual screening by Surflex-dock revealed the compounds CID149925965 and CID142139076 as the best inhibitors potentially more active towards PLpro with affinities equal to 9.52 M<sup>-1</sup> and 8.20M<sup>-1</sup>, respectively.

Finally, the application of Lipinski's rule provides us with positive information on the pharmacokinetic properties (ADME-Tox) of these two proposed inhibitors.

**Keywords:** SARS-COV-2, PL Proteases, Virtual Screening, RMSD, Surflex-dock, Rule of *Lipinski*.

# الموضوع: تحديد مثبطات الأنزيم PL protéase الجديدة كإستراتيجية علاجية لـ COVID-19 في فحص السيليكو وتصفية ADME-Tox

## ملخص

COVID-19 هو مرض تنفسي قاتل يسببه فيروس SARS COV-2 انتشر بسرعة ، أولاً عبر الصين ثم في الخارج ، مما تسبب في وباء عالمي.

عملنا هو جزء من البحث في السيليكو عن علاج فيروس كورونا الجديد عن طريق تثبيط إنزيم PL protéases (PLpro) بمثبطات جديدة. أجرينا فحصاً افتراضياً بناءً على هيكل إنزيم SARS-COV-2 المستهدف PLpro. هذا النهج يجعل من الممكن محاكاة التفاعلات بين البروتين وآلاف الجزيئات بهدف اكتشاف المركبات ذات النشاط المثبط العالي ضد الهدف العلاجي.

تم استخدام برنامج Surfex-dock لاكتشاف مثبطات PLpro الجديدة الأكثر فاعلية. مع وجود نسبة مئوية كبيرة من قيم RMSD أقل من أو تساوي 2 (88%)، يمكن القول أن هذا البرنامج يعمل بشكل جيد. لذلك يمكن استخدامه للتعلم في آلية تثبيط PLpro دون مخاطر كبيرة للخطأ. في الواقع، تم اختبار مجموعة من 555 جزيء من مكتبة المواد الكيميائية PubChem مقابل الموقع النشط لـ PLpro. يضيء الفحص الافتراضي بواسطة Surfex-dock المركبين CID142139076 و CID149925965 كأفضل مثبطات يحتمل أن تكون أكثر نشاطاً تجاه PLpro مع تقاربات تساوي  $M^{-1} 9.52$  و  $M^{-1} 8.20$  ، على التوالي.

أخيراً ، يوفر لنا تطبيق قاعدة ليبينسكي معلومات إيجابية عن الخصائص الحركية الدوائية (ADME-Tox) لهذين المثبطين المقترحين.

الكلمات المفتاحية: SARS COV-2، PL protéase، RMSD، Surfex-dock، الفحص الافتراضي قاعدة Lipinski

<b>TABLE DES MATIERS</b>	<b>Page</b>
Remerciements	
Dédicace	
Résumé	
Table des matières	
Liste des figures	
Liste de tableaux	
Liste des graphes	
Liste des abréviations	
<b>INTRODUCTION</b>	<b>01</b>
<b>Partie bibliographique</b>	
<b>CHAPITRE I. CORONAVIRUS</b>	
I. Historique	<b>03</b>
II. Définition	<b>03</b>
II.1. SRAS-COV et MERS-COV : les épidémies de 2003 et 2012	<b>03</b>
II.2. COVID-19	<b>04</b>
III. La classification de COVID-19	<b>04</b>
IV. Structure de COVID-19	<b>05</b>
V. Physiopathologie d'une infection à SARS-COV-2	<b>06</b>
V.1. Rôle de la réponse inflammatoire	<b>06</b>
V.2. Une situation à risque thrombotique	<b>07</b>
VI. Symptômes	<b>08</b>
VI.1. SRAS-COV et le MERS-COV	<b>08</b>
VI.2. Le COVID-19	<b>08</b>
VII. Critères diagnostique du COVID-19	<b>09</b>
VII.1. Tests virologiques (RT-PCR)	<b>09</b>
VII.2. Tests sérologiques	<b>09</b>
VII.3. Tests antigéniques	<b>09</b>
VII.4. Tests salivaires	<b>09</b>
VII .5. Autre tests	<b>10</b>
VIII. Transmission	<b>10</b>
IX. Traitement	<b>10</b>
IX.1. Médecine chimique	<b>10</b>
IX.2. Médecine biologique	<b>12</b>
IX.3. Les vaccins COVID-19	<b>13</b>
IX.4. Stratégie thérapeutique basée sur l'inhibition de la protéase du SARS-COV-2	<b>14</b>
<b>CHAPITRE II. PAPAÏNE-LIKE PROTEASE</b>	
I. Protéases du SARS-COV2	<b>15</b>
II. Papaïne-Like protéase	<b>15</b>
III. Structure de PLpro	<b>16</b>
IV. Site actif du PLpro	<b>18</b>
V. Mode d'action dePLpro	<b>19</b>

VI. Mode d'inhibition de PLpro	20
VI.1. Les inhibiteurs covalents	20
VI.2. Les inhibiteurs non covalents	21
<b>CHAPITRE III. CRIBLAGE VIRTUEL</b>	
I. Généralité	24
II. Outils du criblage virtuel	25
II.1. Structure de la cible protéique	25
II.2. Importance des chimiothèques dans la recherche de nouvelles substances thérapeutiques	25
II.3. Logiciels de criblage virtuel	26
III. Différentes stratégies du criblage virtuel	27
III.1. Criblage virtuel « <i>ligand-based</i> »	27
III.2. Criblage virtuel « <i>structure-based</i> »	28
IV. Docking moléculaire	29
IV.1. Définition	29
IV.2. Principe	30
IV.3. Méthodes de docking moléculaire	30
IV.4. Protocole générale de docking	31
V. Programme de docking moléculaire Surflex-dock	32
<b>Partie pratique</b>	
<b>CHAPITRE IV. MATERIEL ET METHODES</b>	
I. Tests de fiabilité de programme Surflex-dock	34
I.1. Le RMSD ( <i>Root Mean Square Deviation</i> )	34
I.2. Analyses visuelles	35
II. Criblage virtuel	35
II.1. Préparation de la cible enzymatique	35
II.2. Préparation des ligands	36
III. Banques de données	36
III.1. PDB ( <i>Protein Data Bank</i> )	36
III.2. PubChem	36
IV. Programmes utilisés	37
IV.1. Surflex-dock (1.3 .2005)	37
IV.2. Open Babel	40
IV.3. Discovery studio visualizer version 2021	41
V. Filtrage ADME-Tox	42
<b>CHAPITRE V. RESULTATS ET DISCUSSIONS</b>	
I. Tests de fiabilité du programme de docking Surflex-dock	43
I.1. Test RMSD	43
I.2. Analyses Visuelles	45
I.3. Conclusion	47
II. Etude des interactions intervenant dans l'inhibition de la PL protéase par diverses molécules provenant de la PDB	47
II.1. Choix du complexe 7JN2	48
II.2. Site actif de la 7JN2	48
II.3. Interaction 7JN2-inhibiteurs	48
II.4. Analyse visuelle des interactions « 7JN2-inhibiteur »	52
II.4.1. Complexe 7JN2-TTT	52
II.4.2. Complexe 7JN2-JW9	54

III. Proposition de nouveaux inhibiteurs de la PLpro	<b>56</b>
III.1. Criblage virtuel basé sur la structure de l'enzyme PLpro	<b>56</b>
III.2. Molécules sélectionnées comme nouveaux inhibiteurs	<b>62</b>
III.3. Prédiction du mode d'interaction des meilleurs inhibiteurs	<b>63</b>
III.3.1. Interaction de complexe 7JN2- CID 149925965	<b>63</b>
III.3.2. Interaction de complexe 7JN2-CID 142139076	<b>65</b>
IV. filtrage ADME/Tox de composés proposés	<b>67</b>
CONCLUSION ET PERSPECTIVE	<b>70</b>
BIBLIOGRAPHIE	<b>72</b>
ANNEXES	

## LISTE DES FIGURES

Figure	Page
<b>Chapitre I. Coronavirus</b>	
<b>Figure 1.</b> Représentation schématique de la taxonomie des Coronaviridae (selon l'ICTV)	<b>05</b>
<b>Figure 2.</b> Structure d'un Coronavirus	<b>06</b>
<b>Figure 3.</b> Physiopathologie du Coronavirus Covid-19	<b>07</b>
<b>Chapitre II. Papaine-Like Protease</b>	
<b>Figure 1.</b> Représentation de polyprotéines du SARS-COV2	<b>16</b>
<b>Figure 2.</b> La superposition de PLpro entre SARS-COV2 (couleur bleue, PDB ID : 7D7K) et SARS-COV (couleur rose ; PDB ID : 2FE8)	<b>17</b>
<b>Figure 3.</b> Structure de PLpro de SARS-COV2	<b>18</b>
<b>Figure 4.</b> Représentation du SARS-CoV2 PLpro	<b>19</b>
<b>Figure 5.</b> (A) Structure 2D de VIR251 ; (B) Coordonnées aux rayons X de VIR251 dans le site de liaison PLpro (PDB ID : 6WX4)	<b>21</b>
<b>Figure 6.</b> A : Structures chimiques du GRL-0617. B : site actif de PLpro avec l'inhibiteur GRL-0617	<b>22</b>
<b>Figure 7.</b> Inhibiteurs SARS-COV-2 PLpro	<b>23</b>
<b>Chapitre III. Criblage virtuel</b>	
<b>Figure 1.</b> Criblage virtuel <i>in silico</i>	<b>25</b>
<b>Figure 2.</b> Criblage virtuel « Structure-based »	<b>29</b>
<b>Figure 3.</b> Docking protéine-ligand	<b>30</b>
<b>Figure 4.</b> Les étapes de docking moléculaire	<b>32</b>
<b>Figure 5.</b> La méthode de construction incrémentale	<b>33</b>
<b>Chapitre IV. Matériel et méthodes</b>	
<b>Figure 1.</b> Page internet de la PDB	<b>36</b>
<b>Figure 2.</b> Page internet de la chimiothèque PubChem	<b>37</b>
<b>Figure 3.</b> Page internet de programme Surflex-dock	<b>38</b>
<b>Figure 4.</b> Génération du protomol pour la PLpro autour de l'inhibiteur TTT	<b>39</b>

<b>Figure 5.</b> Exemple d'un résultat de docking moléculaire donné par Surfex-dock	<b>40</b>
<b>Figure 6.</b> Interface graphique de logiciel Discovery studio visualizer	<b>41</b>
<b>Chapitre V. Résultats et Discussions</b>	
<b>Figure 1.</b> Superposition du ligand donné par rayon-X (coloré en rose) et du finale-0 calculé par Surfex-dock (coloré en vert) du complexe 2B5J	<b>46</b>
<b>Figure 2.</b> Superposition du ligand donné par rayon-X (coloré en rose) et du finale-0 calculé par Surfex-dock (coloré en vert) du complexe 1VEY	<b>46</b>
<b>Figure 3.</b> Superposition du ligand donné par rayon-X (coloré en rose) et du finale-0 calculé par Surfex-dock (coloré en vert) du complexe 1ICJ	<b>47</b>
<b>Figure 4.</b> Analyse visuelle par discovery studio visualizer du mode d'interaction du ligand TTT avec le site actif de la 7JN2	<b>52</b>
<b>Figure 5.</b> Représentation des interactions hydrophobiques formées par le composé TTT et le site actif de la PLpro.	<b>53</b>
<b>Figure 6.</b> Mode de liaison de l'inhibiteur TTT avec le site actif de la PLpro.	<b>54</b>
<b>Figure 7.</b> Analyse visuelle par discovery studio visualizer de mode d'interaction du complexe 7JN2-JW9	<b>54</b>
<b>Figure 8.</b> Représentation des interactions hydrophobiques formées par le composé JW9	<b>55</b>
<b>Figure 9.</b> Mode de liaison de l'inhibiteur JW9 avec le site actif de la 7JN2	<b>56</b>
<b>Figure 10.</b> Représentation des interactions hydrophobiques formées par l'inhibiteur CID149925965 avec la 7JN2	<b>64</b>
<b>Figure 11.</b> Représentation des interactions hydrogène formées par le similaireCID149925965dans le site actif de 7JN2	<b>64</b>
<b>Figure 12.</b> Représentation des interactions hydrophobiques formées par le composé CID142139076 avec la 7JN2	<b>66</b>
<b>Figure 13.</b> Mode de liaison de l'inhibiteur CID142139076 avec le site actif de la 7JN2	<b>66</b>

## LISTE DES TABLEAUX

Tableau	Page
<b>Chapitre I. Coronavirus</b>	
<b>Tableau 1.</b> Comparaison de certains des vaccins Covid-19	<b>14</b>
<b>Chapitre III. Criblage virtuel</b>	
<b>Tableau 1.</b> Quelques exemples de logiciels de criblage virtuel	<b>27</b>
<b>Chapitre V. Résultats et Discussions</b>	
<b>Tableau 1.</b> Valeurs de RMSD des 3 complexes étudiés	<b>45</b>
<b>Tableau 2.</b> Principales caractéristiques du code 7JN2	<b>48</b>
<b>Tableau 3.</b> Les valeurs d’Affinité des complexes 7JN2-inhibiteur étudiés	<b>49</b>
<b>Tableau 4.</b> Résultats de docking des similaires de composé TTT	<b>57</b>
<b>Tableau 5.</b> Résultats de docking des similaires de composé JW9	<b>61</b>
<b>Tableau 6.</b> Données chimiques et structurales des composés sélectionnés comme candidats.	<b>63</b>
<b>Tableau 7.</b> Les quatre ponts hydrogène formés par l’inhibiteur CID149925965	<b>65</b>
<b>Tableau 8.</b> Les trois ponts hydrogène formés par l’inhibiteur CID142139076	<b>67</b>
<b>Tableau 9.</b> Critères de la règle de <i>Lipinski</i> pour les différents inhibiteurs étudiés	<b>68</b>
<b>Tableau 10.</b> Autres critères	<b>69</b>

## LISTE DES GRAPHES

Graphe	Page
<b>Chapitre V. Résultats et Discussions</b>	
<b>Graphe 1.</b> Résultat global du RMSD en %	<b>43</b>
<b>Graphe 2.</b> Résultats en % obtenus par Surflex-dock à divers intervalles de RMSD (Å)	<b>44</b>

## LISTE DES ABREVIATIONS

**Å**: Angström

**ACE2** :Angiotensin Converting Enzyme 2

**ADN** : Acide Désoxyribo Nucléique

**ARN**: Acide ribonucléique

**COVID-19**:Coronavirus disease 2019

**CV** : Criblage virtuelle

**ICTV** : International comité de taxonomie virale

**IFN** : interfère de type I

**KDa** : Unité de masse atomique en Kilodalton

**LogP**: Coefficient de Partage

**MERS-COV**: Middle –East Respiratory Syndrome Related Coronavirus

**Mpro**: SARS coronavirus main proteinase

**Nsp3** : Non-structural protein 3

**PDB**: Protein Data Bank

**PLpro**: Papain-like protease

**RMSD**: Root-Mean-Square Deviation

**SARS-COV-2**:Severe Acute Respiratory Syndrome Coronavirus 2

**UbL**: Ubiquitin-like proteins

**UsP**: Ubiquitin-specific- protease

**2D**: Deux dimensionns

**3D** : Trois dimensions

**3CLpro** : 3C-like Protease

# INTRODUCTION



### INTRODUCTION

Le SARS-COV-2 est un nouveau coronavirus (COV) humain. Il a émergé en Chine fin 2019 et est responsable de la pandémie mondiale de COVID-19 qui a causé plus de 540 000 décès en six mois. La compréhension de l'origine de ce virus est une question importante et il est nécessaire de déterminer les mécanismes de sa dissémination afin de pouvoir se prémunir de nouvelles épidémies [1].

Compte tenu du grand nombre d'essais cliniques en cours et des nombreuses recherches sur la physiopathologie de l'infection virale du SARS COV-2, on peut espérer qu'un nouveau traitement sera trouvé. Cependant, il n'y a malheureusement aucune certitude d'arriver à un traitement efficace, si l'on se réfère aux nombreuses infections virales qui sont encore orphelines de traitement médicamenteux [2]. Le développement d'inhibiteurs de PL protéase (PLpro) devrait bénéficier de moyens conséquents pour pouvoir disposer rapidement d'un traitement efficace contre le COVID-19 [1].

La PLpro est l'une des cibles antivirales intéressantes car elle est essentielle à la réplication du coronavirus. La PLpro assure l'élimination de l'ubiquitine et l'ISG15 des protéines de la cellule hôte pour aider les coronavirus à échapper aux réponses immunitaires innées de l'hôte. De ce fait, le ciblage des protéases de type papaine avec des médicaments antiviraux peut être bénéfique non seulement pour inhiber la réplication virale, mais également en inhibant de signalisation [3].

Le criblage virtuel par docking moléculaire est une nouvelle approche visant à simuler et prédire l'affinité d'un très grand nombre de ligands pour le site actif d'une cible thérapeutique donnée, ce qui est assez facile à mettre en place, plus rapide, pour un coût bien moindre que les criblages expérimentaux [4].

Dans le but de découvrir des nouveaux traitements du COVID-19, des études de criblage *in silico* et de simulation moléculaire ont été réalisées pour identifier de nouveaux médicaments contre les protéases du ce virus comme la PL-protéase.

Le but de cette étude est de :

- Tester la performance du programme de criblage virtuel Surflex-dock grâce à quelques critères ;

- Etudier l'inhibition de l'enzyme PLpro par la méthode de criblage virtuel assisté par ordinateur. Pour cela, nous nous intéressons à déterminer le mode d'interaction, lors de la fixation de l'inhibiteur à l'enzyme durant la formation du complexe PLpro-inhibiteur, avec une meilleure complémentarité par le calcul de l'affinité du complexe formé par le logiciel Surflex-dock ;
- Proposer de nouveaux inhibiteurs plus puissants envers notre cible enzymatique. Cette recherche permet de découvrir de nouveaux candidats médicament du SARS COV-2.

Nous présentons nos travaux dans ce mémoire au cours de deux parties:

**1.** Une partie bibliographique avec trois chapitres :

- Le premier chapitre parle sur le coronavirus.
- Le deuxième donne un aperçu sur notre cible, la « PLpro » et ses inhibiteurs.
- Le troisième chapitre décrit le criblage virtuel et ses principales stratégies.

**2.** Une partie pratique divisée en deux chapitres :

- Le premier chapitre présente le matériel et les méthodes utilisés dans cette étude.
- Le second expose les résultats et leurs interprétations suivis d'une conclusion qui résume le travail et des perspectives.

# CHAPITRE I : CORONAVIRUS



## CHAPITRE I. CORONAVIRUS

### I. Historique

Les coronavirus (COV) consistent en des virus qui infectent plusieurs espèces. Les premiers COV se rapportent aux animaux et n'ont d'abord pas reçu l'appellation «coronavirus», apparue plus tardivement dans le premier rapport du comité international de taxonomie virale (ICTV) en 1971[5].

En 1968, le terme « coronavirus » fait officiellement son apparition dans la revue Nature. Ce nouveau groupe de virus se définit alors, à partir de critères essentiellement morphologiques. Il faut attendre mars 2003, et l'identification du coronavirus le SARS-COV qui est l'agent infectieux responsable du Syndrome Respiratoire Aigu Sévère (SARS) pour que les coronavirus suscitent l'intérêt de la communauté médicale et scientifique. En effet, le SARS-COV se trouve à l'origine de la première pandémie infectieuse du XXIe siècle [5].

En septembre 2012, un nouveau coronavirus, le MERS COV émerge au Moyen-Orient (*Middle-East respiratory syndrome-related coronavirus*, MERS-COV), plus précisément en Arabie Saoudite, responsable d'un syndrome respiratoire sévère, confirmant le haut potentiel d'émergence de ces virus [5].

Enfin, le dernier coronavirus humain en date, le SARS-COV-2 apparaît à Wuhan en Chine, il s'avère responsable d'une sévère pandémie en 2020 [5].

### II. Définition

#### II.1. SRAS-COV et MERS-COV : les épidémies de 2003 et 2012

Les Coronavirus forment une immense famille de virus possédant un génome à ARN extrêmement long (plusieurs milliers de nucléotides). Ils sont entourés d'une capsule de protéines en forme de couronne qui leur vaut leur nom. Il existe de nombreux sous-types de coronavirus infectant différentes espèces animales. L'homme peut en héberger au moins cinq, parmi lesquels les plus courants sont HCoV-229 et HCoV-OC43 [6].

Très répandus, ces virus sont associés à des rhumes et des syndromes grippaux bénins. Ils peuvent également infecter l'homme sans déclencher de symptômes ou à l'inverse, être impliqués dans des complications respiratoires de type pneumonie chez des personnes immunodéprimées ou des nourrissons [6].

Ces virus se transmettent facilement d'homme à homme par voie aérienne, au contact de sécrétions ou à celui d'objets contaminés, particulièrement en période hivernale. La période d'incubation qui précède l'apparition des symptômes dure 3 à 6 jours et les traitements, s'ils sont nécessaires, sont symptomatiques (traitement de la fièvre, des congestions ou des douleurs éventuelles). Néanmoins, les infections à coronavirus ne sont habituellement pas diagnostiquées en raison de leur caractère bénin et de leur guérison spontanée [6].

## II. 2. COVID-19

COVID-19 fait référence à « Coronavirus Disease 2019 », la maladie provoquée par un virus de la famille des Coronaviridae, le SARS-COV-2. Cette maladie infectieuse est une zoonose, dont l'origine est encore débattue, qui a émergé en décembre 2019 dans la ville de Wuhan, dans la province du Hubei en Chine. Elle s'est rapidement propagée, d'abord dans toute la Chine, puis à l'étranger provoquant une épidémie mondiale [7].

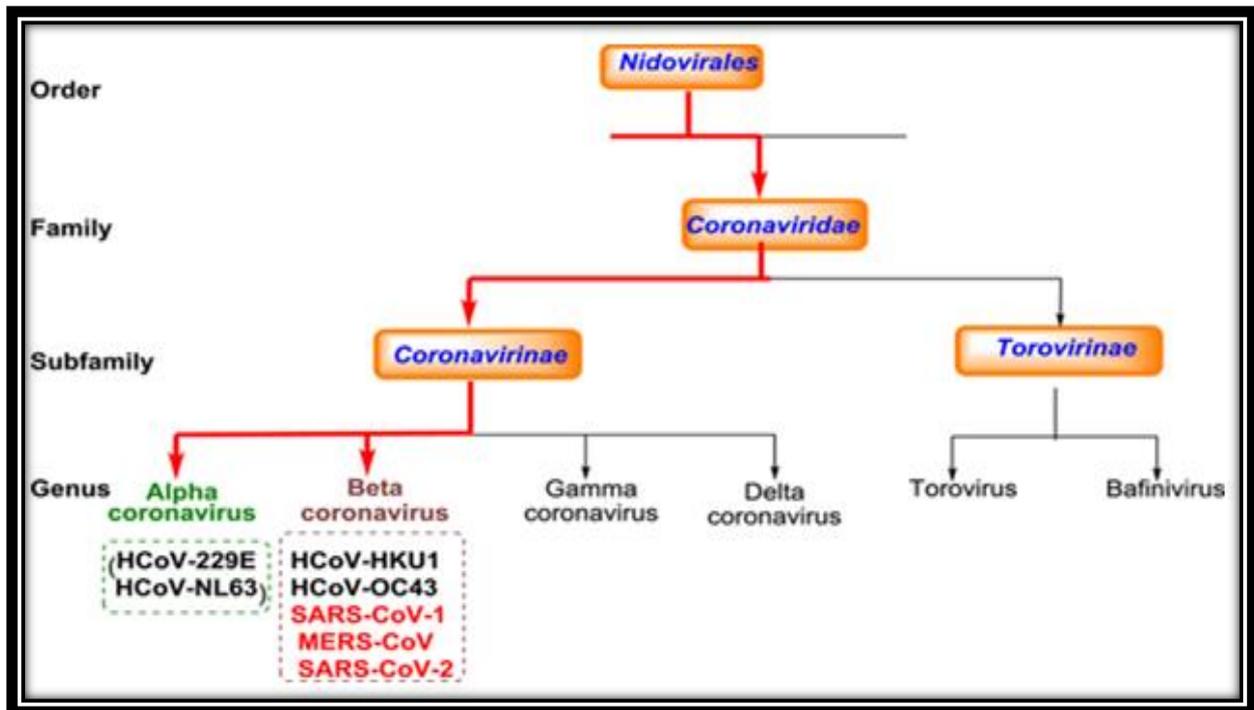
Le COVID-19 est une maladie respiratoire pouvant être mortelle chez les patients fragilisés par l'âge ou une autre maladie chronique. Elle se transmet par contact rapproché avec des personnes infectées. La maladie pourrait aussi être transmise par des patients asymptomatiques mais les données scientifiques manquent pour en attester avec certitude [7].

Le COVID-19 signifie que [8] :

- "Co" pour "corona" ;
- "vi" pour "virus" ;
- "D" pour "disease" (maladie).
- 19 pour l'année de son apparition : 2019.

## III. La classification de COVID-19

Les COV sont des virus qui appartiennent à l'ordre des Nidovirales et à la famille des Coronaviridae, elle-même subdivisée en 2 sous-familles, les Coronavirinae et les Torovirinae. Dans la taxonomie actuelle, la famille des Coronavirinae comprend 4 genres appelés Alpha -, Beta -, Gamma - et Delta -coronavirus. Tandis que les Alpha coronavirus et Beta coronavirus infectent principalement les mammifères, ainsi que les chauves-souris, les Gamma coronavirus et les Delta coronavirus touchent surtout les oiseaux [5].



**Figure I .1.** Représentation schématique de la taxonomie des Coronaviridae (selon l'ICTV). Les sept coronavirus infectant l'homme appartiennent au genre  $\alpha$ - ou  $\beta$ -coronavirus (les agents pathogènes hautement infectieux sont surlignés en rouge) [10].

#### IV. Structure de COVID-19

Les COV prennent la forme de virus enveloppés, plutôt sphériques et mesurent 80 à 220 nanomètres de diamètre. Le SARS-COV-2 constitue un virus à ARN monocaténaire non segmenté de polarité positive. Sa taille s'élevant à près de 30 kb, le classifie le plus grand des génomes des virus à ARN. Il comporte un grand transcrite se traduisant en 16 protéines coupées par des protéases. Ces protéines réalisent la structure du virus et permettent la réplication virale. Outre ce grand transcrite, le génome comporte également quatre gènes de l'extérieur vers l'intérieur la glycoprotéine Spike (S), la protéine d'enveloppe (E), la protéine de matrice (M) et la nucléocapside (N). L'enveloppe virale porte à sa surface de hautes projections formées de protéines de surface S associées en trimère donnant un aspect en couronne (d'où le préfixe latin « corona ») à la particule virale [5].

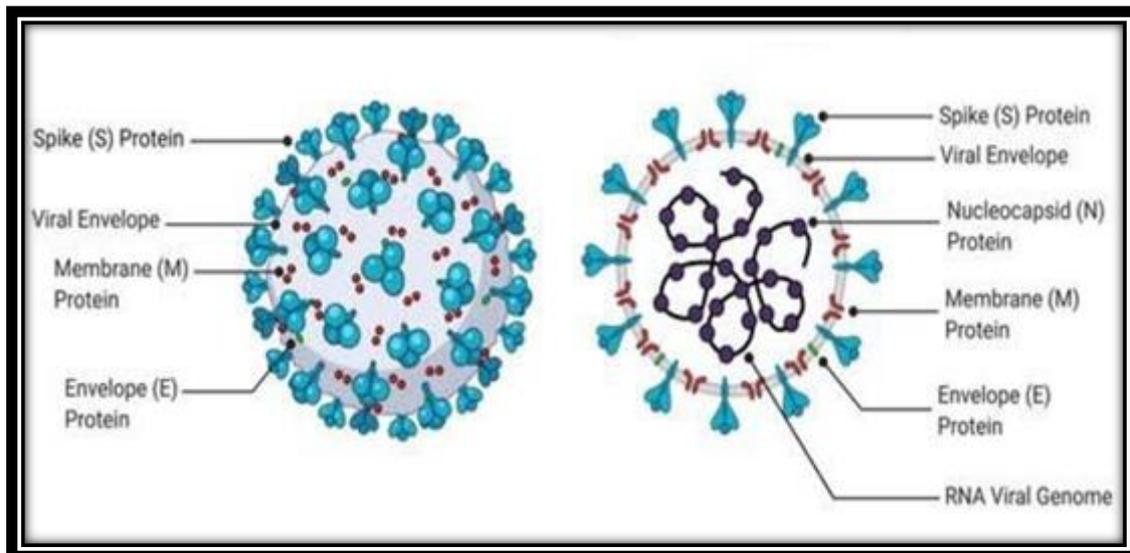


Figure I.2. Structure d'un coronavirus [12].

## V. Physiopathologie d'une infection à SARS-COV-2

Après pénétration du virus dans l'organisme, une phase d'incubation de quelques jours est observée avant l'apparition de signes cliniques. L'infection cellulaire est possible grâce à la protéine Spike, qui interagit avec le récepteur cellulaire à l'angiotensine II, nommé ACE2 (*Angiotensin Converting Enzyme 2*). Cette liaison protéique permet l'endocytose du virus par la cellule hôte puis la fusion membranaire. La nucléocapside virale est larguée dans le cytosol de la cellule infectée. L'ARN de la réplicase virale est traduit par la machinerie cellulaire aboutissant à la formation d'un complexe protéique de transcription et de réplication. De L'ARN et des protéines virales structurales sont alors synthétisées et assemblés abolissant à la production de nouveaux virions. La distribution de ce récepteur ACE2 explique en partie la symptomatologie d'infection à SAR-COV-2. Son expression est en effet importante au niveau de la membrane des cellules épithéliales pulmonaires et intestinales mais également au sein du système vasculaire à la surface des cellules endothéliales artérielles et veineuses et des cellules musculaires lisses vasculaires [16].

### V.1. Rôle de la réponse inflammatoire

La réponse immunitaire, parfois disproportionnée, est à l'origine de manifestations sévères observées lors de la COVID-19. En effet, l'infection des cellules épithéliales des voies respiratoires génère une sécrétion de cytokines pro-inflammatoires permettant normalement d'inhiber la réplication virale et de stimuler l'activité lymphocytaire antivirale. Or, cette réponse immunitaire initiale peut être inefficace et s'intensifier nettement environ huit jours après

l'apparition des premiers symptômes. Cette hypersécrétion cytokinique, parfois appelée orage cytokinique, entraîne notamment une augmentation de la perméabilité des capillaires sanguins et l'attraction de cellules immunitaires, résultant en une infiltration pulmonaire inflammatoire altérant significativement les échanges gazeux cliniquement, une dyspnée de gravité variable est observée, jusqu'au syndrome de détresse respiratoire aigu [16].

## V.2. Une situation à risque thrombotique

Différents mécanismes intriqués sont à l'origine de thromboses vasculaires. Un premier mécanisme en jeu est la réponse inflammatoire. La libération intense de cytokines peut induire une activation de la coagulation par différentes voies. Le facteur tissulaire (ou facteur III de la coagulation) est sécrété par les cellules mononuclées sanguines aboutissant à activation plaquettaire et à la formation de thrombine, pour finalement provoquer la formation d'un thrombus. Par ailleurs, une agression virale directe de l'endothélium vasculaire (exprimant le récepteur ACE2) et indirecte via les molécules de l'inflammation, provoque une dysfonction et une inflammation des cellules endothéliales. Celles-ci vont alors produire le facteur tissulaire, activant la cascade de coagulation. Il est également suspecté que l'hypoxémie profonde entraînée par l'atteinte pulmonaire ait un rôle sur l'expression de facteurs pro-thrombotiques et l'inhibition de la synthèse de facteurs anticoagulants [16].

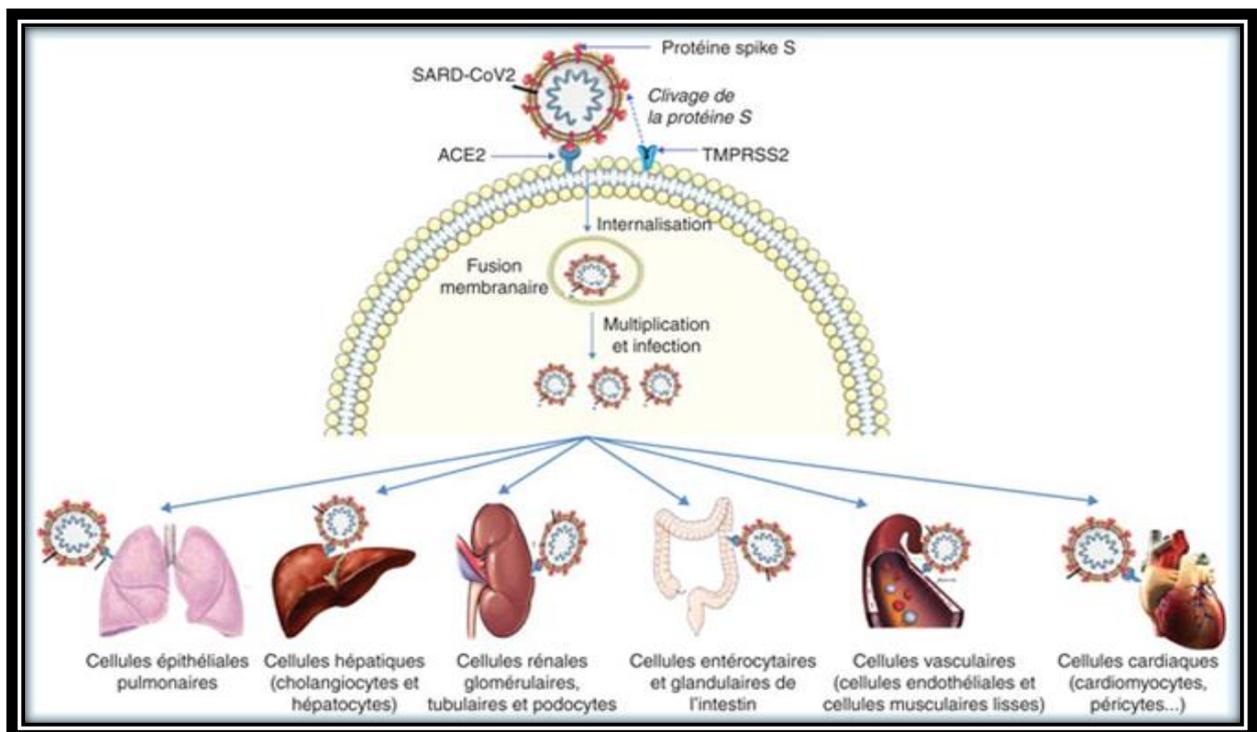


Figure I.3. Physiopathologie du Coronavirus Covid-19 [11].

## VI. Symptômes

### VI.1. SRAS-COV et le MERS-COV

Les infections par le SRAS-COV et le MERS-COV sont asymptomatiques chez certaines personnes. Elles peuvent induire une simple fièvre accompagnée de toux chez d'autres. Mais, elles peuvent aussi entraîner une détresse respiratoire aigüe et le décès de la personne infectée. Les premiers symptômes sont peu spécifiques comme une fièvre, des douleurs musculaires, des céphalées, de la fatigue. De 2 à 7 jours plus tard pour le SRAS-COV et de 5 à 15 jours pour le MERS-COV, apparaissent des symptômes respiratoires comme une toux sèche, des difficultés à respirer et un manque d'oxygène, ainsi qu'une dégradation de l'état général avec notamment des diarrhées fréquentes, mais également des troubles hépatiques ou urinaires, des étourdissements ou encore des problèmes neuromusculaires. La plupart des malades ont une radiographie thoracique ou une tomodensitométrie anormale dès les premiers jours de la maladie, même en l'absence de signes respiratoires. Les cas sévères évoluent en détresse respiratoire qui nécessite des soins intensifs (oxygénothérapie, ventilation assistée...etc.) [6].

### VI.2. Le COVID-19

La plupart des personnes infectées développent une forme légère à modérée de la maladie et se rétablissent sans avoir besoin d'être hospitalisées.

- La durée de l'incubation est en moyenne de 5 jours, avec des extrêmes de 2 à 12 jours. L'installation des symptômes se fait progressivement sur plusieurs jours, contrairement à la grippe qui débute brutalement. Les premiers symptômes sont peu spécifiques : maux de tête, douleurs musculaires, fatigue. La fièvre et les signes respiratoires arrivent secondairement, souvent deux ou trois jours après les premiers symptômes [7].
- Dans les premières études descriptives provenant de Chine, il s'écoule en moyenne une semaine entre l'apparition des premiers symptômes et l'admission à l'hôpital à la phase d'état de la maladie. A ce stade, les symptômes associent fièvre, toux, douleurs thoraciques et gêne respiratoire et la réalisation d'un scanner thoracique montre presque toujours une pneumonie touchant les deux poumons. D'autres signes cliniques ont été décrits depuis les premières études : des signes d'atteinte du système nerveux central s'exprimant en particulier chez les personnes âgées sous la forme d'une désorientation ; une perte du goût ou/et de l'odorat, qui survient chez 30 à 50% des adultes infectés, avec une prédominance féminine, est très évocatrice du diagnostic de COVID-19 [7].

- La gravité des signes cliniques nécessite le maintien à l'hôpital d'environ 20% des malades et 5% nécessitent une admission en réanimation. Les formes les plus graves sont observées principalement chez des personnes vulnérables en raison de leur âge (plus de 70 ans) ou de maladies associées, en particulier l'obésité [7].
- Des études observationnelles privilégiées ainsi que des travaux de modélisation ont montré que l'infection peut être asymptomatique ou pauci symptomatique (entraîner pas ou peu de manifestations cliniques) chez 30 à 60 % des sujets infectés, en particulier chez les jeunes enfants (moins de 12 ans) [7].

## **VII. Critères diagnostique du COVID-19**

### **VII.1. Tests virologiques (RT-PCR)**

Ils permettent de déterminer si une personne est porteuse du virus au moment du test, avec un prélèvement par voie nasale ou salivaire. Le résultat est en général disponible dans les 24 heures qui suivent [17].

### **VII.2. Tests sérologiques**

Ils permettent de rechercher si une personne a développé une réaction immunitaire après avoir été en contact avec le virus. Ces tests détectent la présence d'anticorps au moyen d'une prise de sang [17].

### **VII.3. Tests antigéniques**

Ils détectent la présence du virus ou de fragments de virus SARS-COV-2. Comme les tests par RT-PCR, ils permettent le diagnostic précoce des maladies dès la phase aiguë. Ils répondent à la question « le patient est-il oui ou non porteur du Covid-19 ? ». Comme le test de référence actuel, le RT-PCR, les tests antigéniques sont à ce jour réalisés à partir de prélèvements dans le nez, par écouvillon. Mais alors que le RT-PCR nécessite une analyse parfois de plusieurs heures en laboratoire, pour détecter le matériel génétique du coronavirus, le test antigénique repère des protéines du virus en moins de 30 minutes [17].

### **VII.4. Tests salivaires**

Ils correspondent à un examen de la salive, ils ne font pas mal. Ils sont déployés dans les écoles maternelles et élémentaires pour renforcer le dépistage au COVID-19 et rompre au plus vite les chaînes de contamination. Leur prélèvement est plus acceptable par les jeunes enfants que celui du test nasopharyngé (par le nez) [17].

### VII.5. Autre tests

Un biocapteur peut avoir un impact significatif pour détecter et mesurer le virus COVID-19 dans l'air. Il répond rapidement aux défis des barrières biologiques et des limitations technologiques. Cette technologie aide à mesurer comment un virus peut rester dans l'air lorsqu'il se propage principalement d'une personne à l'autre. Il est également utile pour le diagnostic des maladies liées à l'infection. Il analyse automatiquement divers tests effectués dans les laboratoires et la cause de certaines maladies infectieuses dans l'air [18].

## VIII. Transmission

La transmission se fait essentiellement par voie aérienne (gouttelettes de postillons émises au cours des efforts de toux mais aussi lors de la parole) et passe par un contact rapproché (moins d'un mètre) et durable (au moins 15 minutes) avec un sujet contagieux. Des particules de plus petite taille peuvent aussi être émises sous formes d'aérosols au cours de la parole, ce qui expliquerait que le virus puisse persister en suspension dans l'air dans une pièce non ventilée (et justifie dans ces circonstances le port du masque). Enfin le virus peut conserver une infectiosité pendant quelques heures sur des surfaces inertes d'où il peut être transporté par les mains ce qui justifie une bonne hygiène des mains [7].

## IX. Traitement

### IX.1. Médecine chimique

#### ❖ Lopinavir/Ritonavir

Les comprimés de lopinavir/ritonavir (nom commercial : Clevee) sont des médicaments antiviraux à large spectre. Il s'agit d'un type d'inhibiteur de la protéase du virus de l'immunodéficience humaine (VIH) qui peut bloquer le clivage de la protéine précurseur Gag-Pol en protéine mature, ce qui conduit finalement à la production de particules virales immatures et non infectieuses. Il a été confirmé que le ritonavir est un inhibiteur peptidomimétique actif contre les aspartyl protéases du VIH-1 et du VIH-2. L'enzyme ne peut pas cliver la protéine précurseur Gag-Pol en inhibant la protéase du VIH, ce qui pourrait conduire à la formation de formes immatures de particules de VIH et empêcher un nouveau cycle d'infection [15].

#### ❖ Ribavirine

La ribavirine est un médicament antiviral nucléosidique synthétique doté d'une activité antivirale à large spectre, qui inhibe à la fois les virus à ADN et à ARN, et largement utilisé pour le traitement du SRAS et du MERS. Il a été recommandé que l'utilisation combinée de ribavirine et d'interférona ou de lopinavir, 500 mg, le temps pour les adultes avec 2 à 3 injections intraveineuses par jour pourrait contrôler et améliorer l'état du patient, mais le processus thérapeutique devrait être limité à 10 jours [15].

#### ❖ **Arbidol**

Arbidol est un médicament antiviral non nucléosidique. L'arbidol exerce un effet antiviral en inhibant la fusion de la membrane lipidique virale avec la cellule hôte pour bloquer la réplication du virus. Il est actif contre de nombreux virus enveloppés et non enveloppés et a un effet inducteur d'interféron. Il a été constaté que 10 à 30  $\mu\text{mol}$  d'arbidol pourraient inhiber efficacement la prolifération du SRAS-COV-2 [15].

#### ❖ **Remdesivir**

Le remdesivir est un médicament nucléosidique et un inhibiteur de l'ARN polymérase (RdRp). Le remdesivir pourrait inhiber la réplication du virus et jouer un rôle antiviral en bloquant la synthèse de l'ARN à brin négatif en inhibant le SARS-COV-2 RdRp, ARNm sous-génomique et ARN génomique sous-viral [15].

#### ❖ **Favipiravir**

Le favipiravir est un médicament antiviral nucléosidique à large spectre. Il a un effet inhibiteur sur le virus à ARN RdRp. L'indication initiale du médicament était contre la grippe. Ainsi, le favipiravir est devenu un médicament potentiel pour le traitement des patients COVID-19 [15].

#### ❖ **Darunavir**

Le darunavir est un inhibiteur de la dimérisation de la protéase et de l'activité catalytique du VIH-1. Il peut inhiber sélectivement le clivage de la protéine précurseur gag-Pol codée par le VIH dans les cellules infectées par le virus, bloquant ainsi la formation de particules virales infectieuses matures [15].

#### ❖ **Emtricitabine et denofoviralafénamide**

L'emtricitabine est un inhibiteur nucléosidique de la transcriptase inverse. Il bloque le processus de transcription inverse du VIH en inhibant la transcription inverse. Il a été rapporté

que la combinaison d'emtricitabine et de denofoviralafénamide dans un essai clinique pour le traitement du COVID-19 a été appliquée avec de grands effets thérapeutiques [15].

#### ❖ **Chloroquine Phosphate**

De nombreuses études ont présenté qu'il a un effet antiviral à large spectre. Récemment, il a été découvert que le phosphate de chloroquine pouvait efficacement inhiber l'infection du coronavirus au niveau cellulaire. Il a également été recommandé pour le traitement antiviral du COVID-19 [15].

#### ❖ **l'hydroxychloroquine**

Récemment, il a été rapporté que certains patients COVID-19 bénéficiaient du traitement à l'hydroxychloroquine. De plus, son effet d'élimination virale pourrait être renforcé par l'azithromycine [15].

### **IX.2. Médecine biologique**

#### ❖ **Interféron**

L'interféron (IFN) est une sorte de famille de cytokines multifonctionnelles ayant des activités antivirales, antiprolifératives et immunomodulatrices à large spectre. Selon les différents récepteurs de liaison, il peut être divisé en type I, II et III. Parmi eux, l'IFN de type II (principalement l'IFN $\beta$ ) joue un rôle important dans le contrôle de l'infection virale [15].

#### ❖ **Tocilizumab injectable**

Le tocilizumab injectable est un anticorps monoclonal humanisé recombinant dirigé contre le récepteur humain de la cytokine interleukine-6 (IL-6), qui se lie spécifiquement au récepteur soluble et membranaire de l'IL-6 pour bloquer la transduction du signal, inhibant ainsi l'activité du récepteur de l'IL-6. Il a été constaté qu'en tant que teneur de l'IL-6, l'injection de tocilizumab est susceptible de bloquer la tempête de cytokines et ainsi d'empêcher les patients COVID-19 de se tourner vers des maladies graves et critiques [15].

#### ❖ **Glucocorticoïdes**

Le glucocorticoïde joue un rôle important dans la régulation du développement, de la croissance, du métabolisme et de la croissance de l'organisme. Pendant le traitement du COVID-19, il est recommandé de l'utiliser pour les cas graves avec une forte réponse inflammatoire ou les enfants atteints du syndrome de détresse respiratoire aiguë [15].

### ❖ Thérapie au plasma convalescent

Actuellement, la thérapie au plasma convalescent est en cours de promotion ; certains indicateurs cliniques et symptômes de patients ont été améliorés. Un grand nombre de personnes se sont remises d'une infection à coronavirus, faisant don d'un précieux sérum contenant de l'immunoglobuline, ce qui semble mieux que d'essayer de produire de grandes quantités d'anticorps monoclonaux en laboratoire [15].

### Autres traitements

- L'oxygène supplémentaire est la norme de soins pour les patients ayant de faibles niveaux d'oxygène, et les anticoagulants sont le traitement standard pour réduire les risques de coagulation du sang [14].
- D'autres traitements se sont révélés au fil du temps inefficaces, comme l'hydroxychloroquine très médiatisée [14].
- Au début, le plasma convalescent, une ancienne thérapie où les patients reçoivent du plasma sanguin de personnes qui se sont rétablies du COVID-19 et ont généré des anticorps contre le SRAS-COV-2. Il a été utilisé par des prestataires désespérés de proposer une thérapie, mais il existe actuellement peu de bonnes données d'essais pour soutenir son utilisation et les essais sont toujours en cours [14].
- De même, les suppléments en vente libre comme la vitamine C, D et le zinc n'ont actuellement aucune preuve à l'appui de leur utilisation [14].

### IX.3. Les vaccins COVID-19

Les vaccins COVID-19 sont très efficaces pour prévenir les maladies graves, les hospitalisations et les décès. Se faire vacciner est le meilleur moyen de ralentir la propagation du SRAS-COV-2, le virus qui cause la COVID-19 [13]. Dans le tableau I.1 ci-dessous nous avons résumé les différents vaccins de Covid-19.

Tableau I.1. Comparaison de certains des vaccins Covid-19[9].

Laboratoire	Type	Doses	Stockage
 Oxford Uni-AstraZeneca	Vecteur viral (virus génétiquement modifié)	x2 	 2 à 8°C (6 mois)
 Moderna	ARN (partie du code génétique du virus)	x2 	 De -25 à -15°C (7 mois)
 Pfizer-BioNTech	ARN	x2 	 De -80 à -60°C (6 mois)
 Gamaleya (Sputnik V)	Vecteur viral	x2 	 -18,5°C (forme liquide) 2 à 8°C (forme sèche)
 Sinovac (CoronaVac)	Virus inactivé	x2 	 2 à 8°C
 Sinopharm	Virus inactivé	x2 	 2 à 8°C
 Novavax	À base de protéines	x2 	 2 à 8°C
 Janssen Johnson & Johnson	Vecteur viral	x2 	 2 à 8°C (3 mois)

#### IX.4. Stratégie thérapeutique basée sur l'inhibition de la protéase du SARS-COV-2

Une autre stratégie est de cibler les protéases virales indispensables à la réplication. Si des études structurales permettront le développement de molécules spécifiques, des traitements développés contre la protéase du VIH ont montré une certaine efficacité *in vitro* contre le SARS-COV-1 et le SARS-COV-2. En particulier, le lopinavir/ritonavir a été utilisé comme traitement du SARS-COV-1 et des données rétrospectives l'ont associé à une diminution de la mortalité. Dans le cas de la COVID-19, un essai prospectif randomisé ouvert incluant 199 patients n'a pas retrouvé d'efficacité clinique ni de différence sur la clairance virale. Un autre traitement anti-protéase du VIH, le darunavir, fait l'objet de plusieurs essais cliniques bien qu'une étude menée par son fabricant ne rapporte aucune efficacité notable *in vitro* [19].

# CHAPITRE II :

## PAPAÏNE-LIKE PROTEASE



## CHAPITRE II. PAPAÏNE-LIKE PROTEASE

Les protéases virales sont des enzymes (endopeptidases EC 3.4.2) codées par le matériel génétique (ADN ou ARN) de pathogènes viraux. Le rôle de ces enzymes est de catalyser le clivage de liaisons peptidiques spécifiques dans des précurseurs de polyprotéines virales ou dans des protéines cellulaires. Dans la plupart des cas, ces événements protéolytiques sont essentiels à l'achèvement du cycle infectieux viral. La reconnaissance sélective de ces motifs de séquence par un site de liaison au substrat complémentaire de l'enzyme assure un degré élevé de reconnaissance spécifique et de clivage [20].

### I. Protéases du SARS-COV2

Les protéases du coronavirus, comme le SARS-COV2, la *papain-like protease* (PLpro) et la *3C-like protease* (3CLpro), sont des cibles antivirales intéressantes car elles sont essentielles à la réplication du coronavirus. La PLpro est une protéase présente dans la protéine non structurale 3 (NSp3) du polypeptide viral. La PLpro assure l'élimination de l'ubiquitine et l'ISG15 des protéines de la cellule hôte pour aider les coronavirus à échapper aux réponses immunitaires innées de l'hôte. En effet, l'action de la PLpro perturbe les voies de signalisation INF bêta et NF-kB (*nuclear factor-kappa B*). Le ciblage des protéases de type papaine avec des médicaments antiviraux peut être bénéfique non seulement pour inhiber la réplication virale, mais également en inhibant de signalisation. PLpro supprime l'activation NF-kB et, par la suite, interfère avec la production d'interféron- $\beta$  et supprime le système immunitaire réponse [3]. Les protéases à cystéine de type papaine (EC 3.4.2) appartiennent à la superfamille des papaines. La famille est constituée de protéases d'origine végétale, mammifère, parasitaire et virale. Cette enzyme est généralement un monomère de masse moléculaire relativement faible de l'ordre de 20 à 35 kDa [21].

### II. Papaine-Like protéase

La PLpro appartient au clan des protéases à cystéine mais elle est affiliée à la famille C16, qui contient des endopeptidases polyprotéiques issues des coronavirus. Il existe sous forme de monomère (environ 300 résidus) dans les milieux biologiques et possède le repli USP (*ubiquitin-specific proteases*) [25].

Lors de l'entrée dans les cellules hôtes, le génome d'ARN simple brin (sens positif) est libéré dans le cytoplasme et les polyprotéines pp1a et pp1ab sont produites à partir de la traduction des cadres de lecture ouverts (ORF1a et ORF1b). Les polyprotéines sont clivées par

deux protéases à cystéine, la protéase principale (Mpro) et PLpro, pour former 16 protéines non structurales (nsp) et le domaine PLpro est codé dans nsp3 [24].

Le SARS-COV2 PLpro reconnaît la séquence consensus P4-P1 LXGG (X = n'importe quel acide aminé) et hydrolyse la liaison peptidique qui se trouve entre nsp1-nsp2, nsp2-nsp3 et nsp3-nsp4. Cela entraîne la libération de nsp1, nsp2 et nsp3 [24].

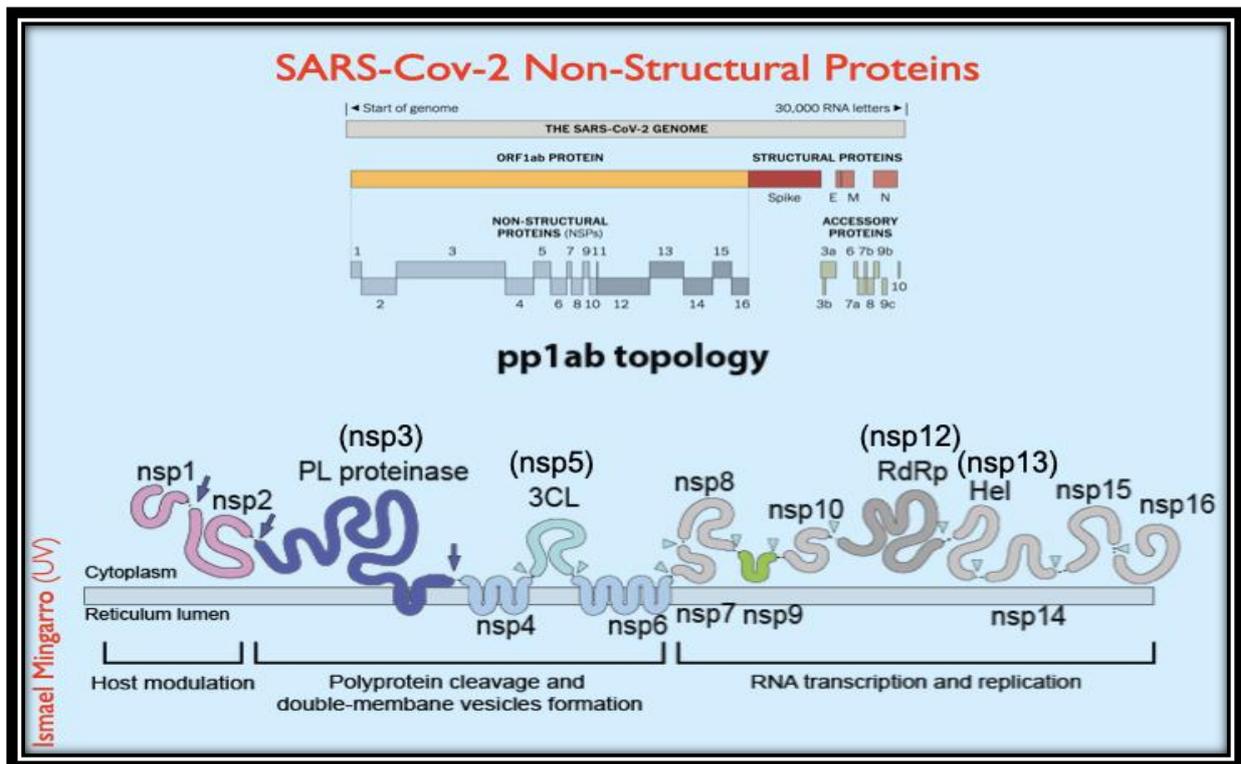
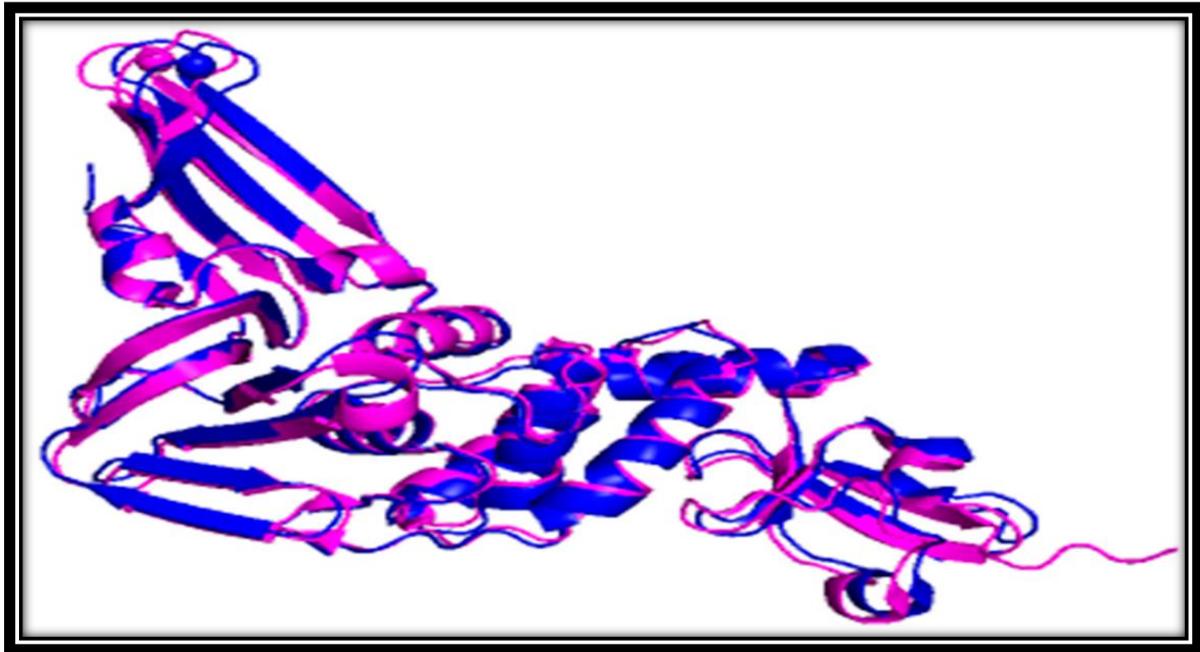


Figure II.1. Représentation de polyprotéines du SARS-COV2 [26].

### III. Structure de PLpro

La séquence de SARS-COV2 PLpro est assez similaire à SARS-COV PLpro comme en témoigne une identité de séquence de 83 % et similaire à 90%. De plus, la superposition entre la structure tridimensionnelle des deux enzymes donne 0,792 Å comme valeur de RMSD. Montrant que les protéines sont structurellement similaires [25].



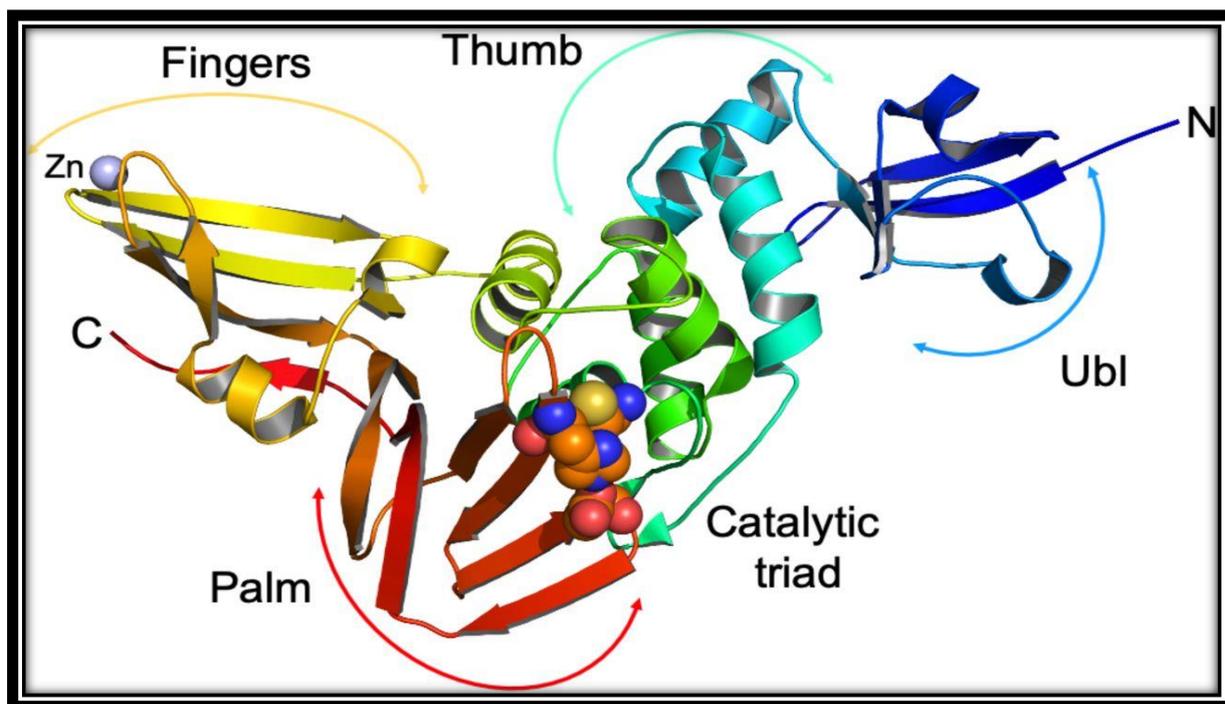
**Figure II.2.** La superposition de PLproentre SARS-COV2 (couleur bleue, PDB ID : 7D7K) et SARS-COV (couleur rose ; PDB ID : 2FE8) [25].

La PLpro-COV2 se compose de quatre domaines distincts. Les 60 premiers résidus forment un domaine indépendant de type ubiquitine (Ubl) N-terminal qui est bien séparé des trois autres domaines (domaines de la paume, du pouce et des doigts), qui forme un domaine de protéase spécifique de l'ubiquitine (Usp) C-terminal [25].

- ❖ Le domaine PLpro Ubl a cinq brins  $\beta$ , de deux hélices  $\alpha$ . Il adopte un pli  $\beta$ -grasp qui est similaire aux domaines ubiquitine et Ubl de plusieurs protéines, dont ISG15. La fonction de ce domaine n'est pas bien comprise et certaines études suggèrent qu'il n'a aucun effet sur la fonction de PLpro. Cependant, la transposition d'Ubl vers le domaine du pouce a entraîné des interactions hydrophobes entre les résidus Pro59 du domaine Ubl et Pro77 et Thr75. L'acide aminé Thr75 interagit alors avec Phe69 de l'hélice "ridge" et modifie ainsi la conformation de ces derniers résidus [25] ;
- ❖ Les plis C-terminaux dans une structure canonique de type pouce-paume-doigts avec le domaine d'ubiquitine ancré au pouce. Dans l'architecture "main ouverte" de PLpro, l'ubiquitine repose sur le domaine "paume" et est maintenue en place par le domaine "doigts" de liaison au zinc [25] ;
- ❖ Le domaine "pouce" (résidus 107–113 et 162–168) est formé de quatre hélices proéminentes ( $\alpha 4-7$ ), et la paume (résidus 269–279) est constituée d'un feuillet  $\beta$  à six

brins ( $\beta 8-13$ ) qui descend dans le site actif, et logé dans une fente exposée au solvant entre les domaines du pouce et de la paume [25] ;

- ❖ Une feuille  $\beta$  antiparallèle à quatre brins constitue le domaine des "doigts" et, dans la région du bout des doigts, quatre résidus de cystéine se coordonnent à un ion zinc [25] ;
- ❖ L'ion Zn est labile et coordonné de manière tétraédrique avec la Cys (189-192-224-226) conservé, il est essentiel pour le processus de catalyse car il détient l'intégrité structurale du PLpro-CoV2 [25].

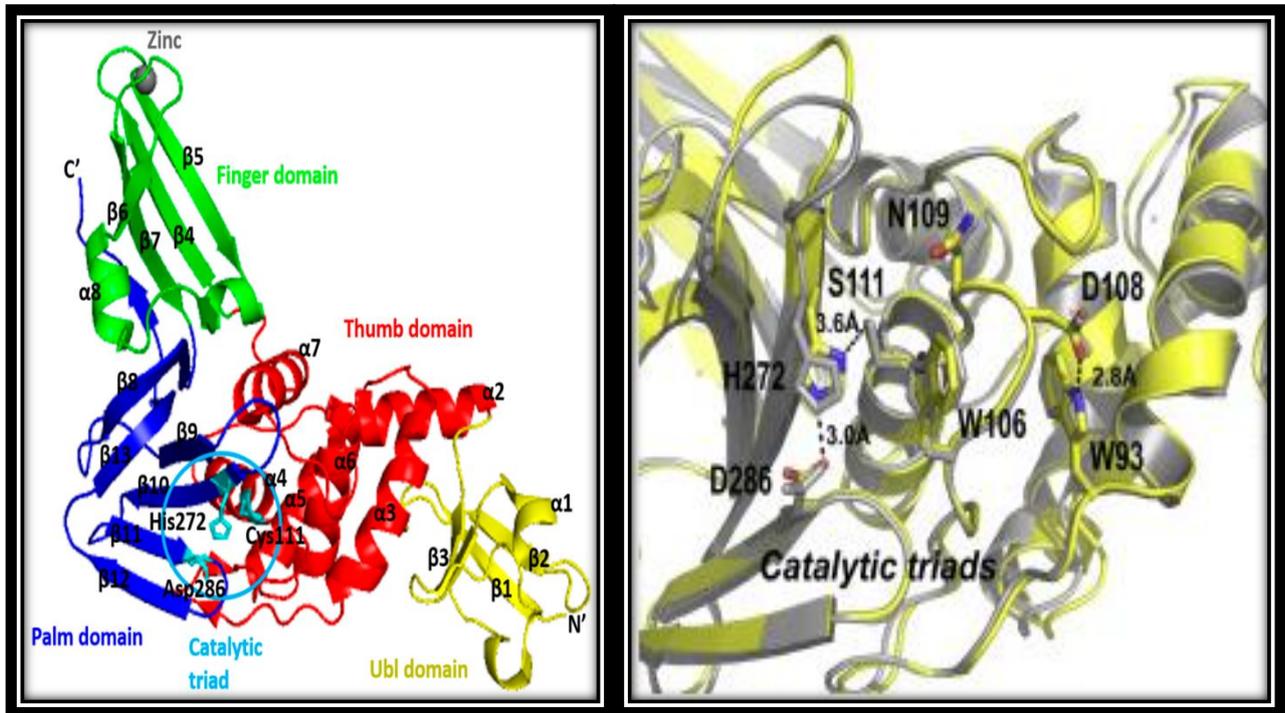


**Figure II.3.** Structure de PLpro de SARS-COV2 montrant la structure secondaire, des sous-domaines avec des résidus de site actifs représentés sous forme de sphères, l'ion zinc est en bleu [23].

#### IV. Site actif du PLpro

Le site actif de PLpro est constitué de Cys111, His272 et Asp286 et d'autres résidus catalytiquement importants Trp93, Trp106, Asp108 et Asn109 sont représentés avec le modèle de bâton (figure II.4), formant la triade catalytique qui a catalysé la liaison peptidique. La cystéine nucléophile (Cys111) est située au pied (extrémité N-terminale) de l'hélice  $\alpha 4$  dans le domaine du pouce. L'His272 est situé au pied du domaine palm et adjacent à la boucle flexible en épingle à cheveux  $\beta$  appelée BL2 (une boucle flexible qui peut se traduire par une conformation ouverte ou fermée). Le cycle catalytique proposé implique le Cys111 catalytique en tant que nucléophile, His272 en tant qu'acide-base général et Asp286 associé à His272 pour

aligner et favoriser la déprotonation de Cys111. En conséquence, les résidus du site actif sont situés à l'interface entre le sous-domaine du pouce et de la paume [25].



**Figure II.4.** Représentation du SARS-CoV2 PLpro. L'interface entre le pouce et la paume forme le site de liaison au substrat menant à la triade catalytique du site actif composé de Cys111, His272 et Asp286 et d'autres résidus importants [25, 30].

## V. Mode d'action de PLpro

L'enzyme se trouve dans nsp3 entre le domaine unique du SARS (SUD) et un domaine de liaison aux acides nucléiques (NAB). Il est hautement conservé et présent dans tous les coronavirus [25].

Cette enzyme reconnaît un motif de clivage consensus, LXGG (X = n'importe quel acide aminé, L = leucine et G = glycine), qui est présent dans deux grandes polyprotéines. Le processus de clivage libérant les protéines nsp1, nsp2 et nsp3 qui s'assembleront et entraîneront la génération d'un complexe de réplicase multifonctionnel associé à la membrane sur les membranes hôtes, initiant la réplication et la transcription du génome viral [25].

La PLpro reconnaît et hydrolyse les protéines cellulaires de l'ubiquitine et la protéine de type ubiquitine (Ubl) ISG15 (gène 15 induit par l'interféron) comme portant le motif de reconnaissance LXGG à leur extrémité C-terminale pour les éliminer des protéines de la cellule hôte [25]. De plus, la PLpro possède également des capacités de désubiquitination et

la déconjugaison d'ISG15 qui interfèrent avec les voies de signalisation critiques menant à l'expression des interférons de type I [25].

## VI. Mode d'inhibition de la PLpro

Dans la recherche d'antiviraux efficaces pour prévenir et traiter l'infection par le coronavirus 2 (SARS-COV2) du syndrome respiratoire aigu sévère, les scientifiques ont étudié la possibilité d'inhiber la principale protéase virale (3CLpro) et la protéase de type papaïne (PLpro), qui est la clé de sa réplication. De nombreux composés qui inhibent ces enzymes ont été décrits et progressent dans la voie du développement de médicaments [27].

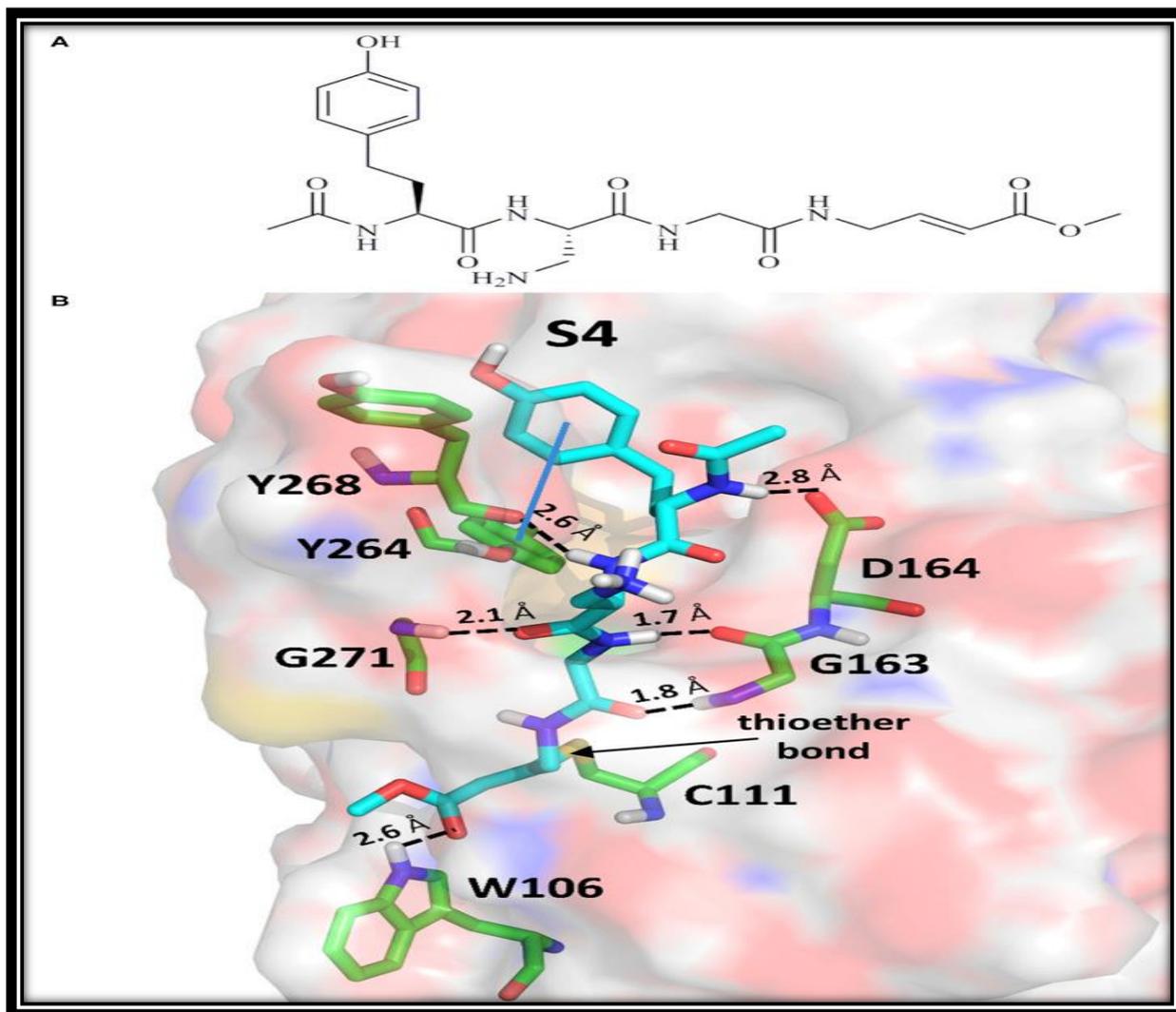
Il existe deux classes principales d'inhibiteurs du SARS-COV2 PLpro, avec à la fois des structures tertiaires résolues de complexes protéine-inhibiteur et des résultats d'études in vitro concernant leurs affinités de liaison [27].

### VI.1. Les inhibiteurs covalents

L'inhibition de la protéase semblable à la papaïne (PLpro) du SARS-COV2 est une cible réussie pour empêcher la propagation du coronavirus dans le corps infecté. À cet égard, les inhibiteurs covalents, tels que le ligand VIR251 récemment proposé, peuvent inactiver la PLpro de manière irréversible en formant une liaison covalente avec un résidu spécifique du site catalytique (Cys111). Un mécanisme d'inhibition peut donc être proposé, comportant quatre étapes [28] :

- Ligand entré dans la poche de protéase ;
- Déprotonation du résidu Cys111 du groupe thiol ;
- Addition de Cys111-S- au ligand ;
- Le transfert de protons de la base protonée vers le ligand lié par covalence.

Évaluer les changements énergétiques et conformationnels de PLpro à chacun de ces étapes pourraient aider à la conception d'inhibiteurs covalents plus efficaces et sélectifs. L'inhibition covalente est généralement un processus irréversible, assurant une désactivation de l'enzyme, entravant ainsi sa fonction [28].



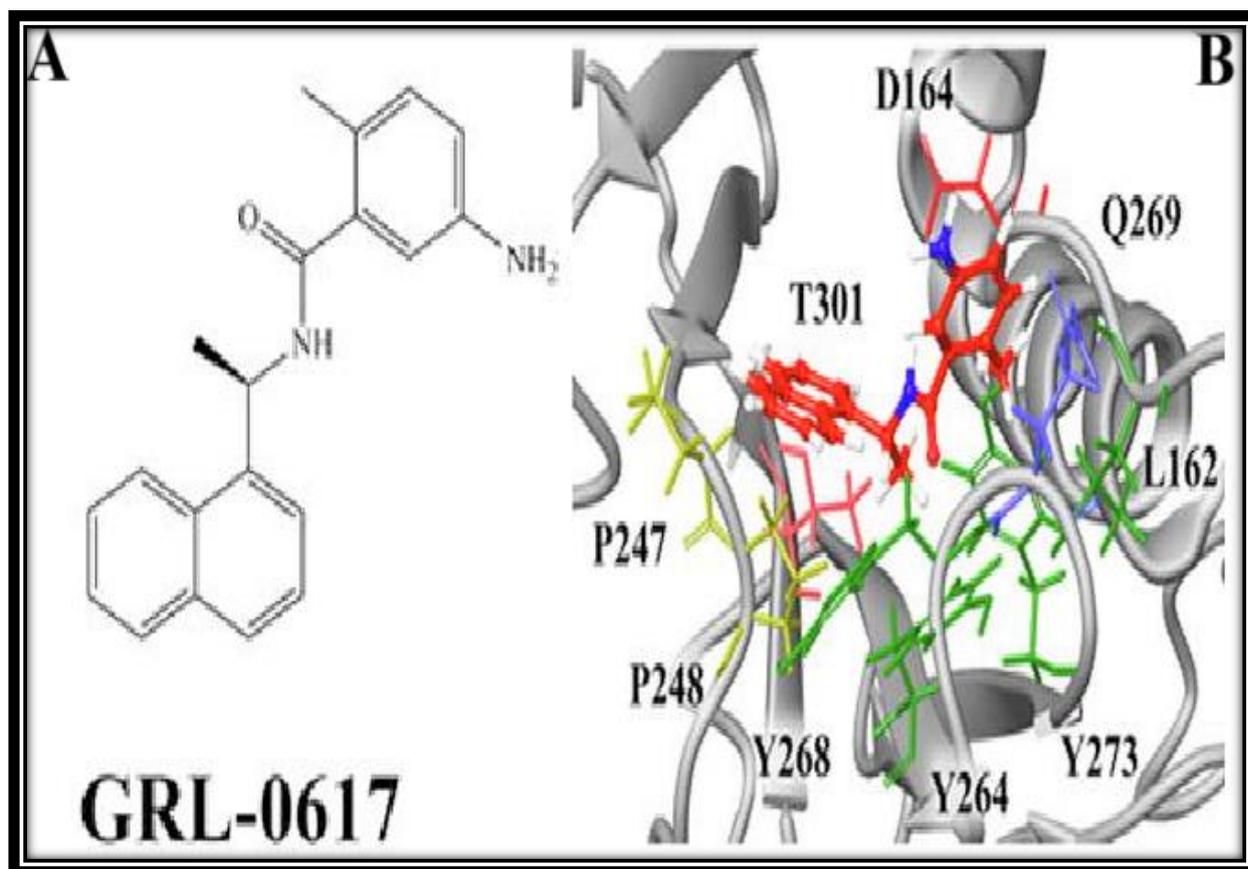
**Figure II.5.** (A) Structure 2D de VIR251 ; (B) Coordonnées aux rayons X de VIR251 dans le site de liaison PLpro (PDB ID : 6WX4). VIR251 et les résidus importants sont rendus sous forme de bâtons, tandis que la protéine est représentée sous forme de surface. Les liaisons H sont représentées par des lignes noires en pointillés [29].

## VI.2. Les inhibiteurs non covalents

La majorité des inhibiteurs de PLpro signalés sont des inhibiteurs non covalents (un phénomène réversible), qui devrait diminuer l'activité enzymatique sans la bloquer complètement [28]. L'inhibiteurs de type médicamenteux non covalents, maintenant connus sous le nom d'inhibiteurs de naphthalène PLpro, ont été découverts. Certains membres de cette classe d'inhibiteurs ont présenté une inhibition nanomolaire contre la PLpro de SRAS-COV et pourraient contrecarrer la réplication virale dans la gamme micromolaire inférieure. Ces inhibiteurs de PLpro à base de naphthalène se sont révélés efficaces pour arrêter l'activité PLpro du SARS-COV-2 ainsi que la réplication du SRAS-CoV-2. [31].

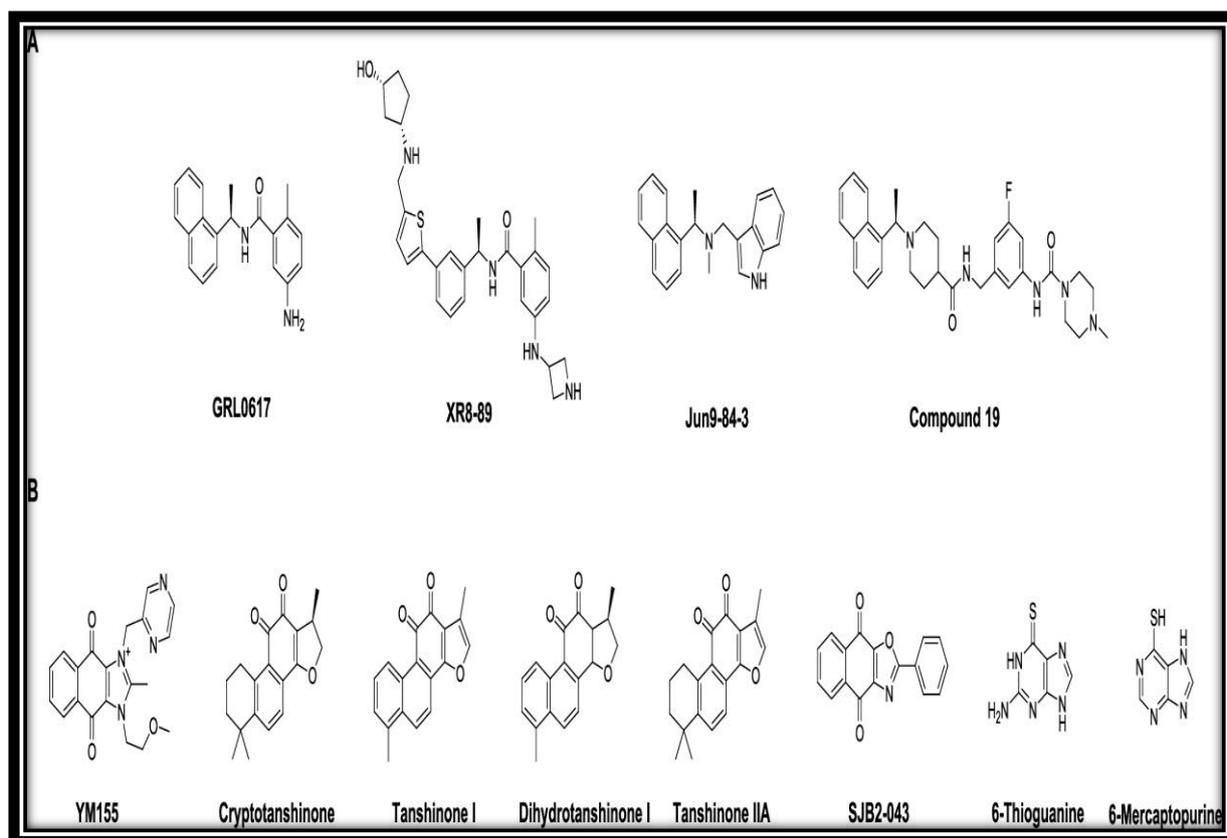
## ❖ L'inhibiteur GRL-0617

Est l'inhibiteur non covalent le plus puissant inhibiteur de la PLpro du SARS-CoV et SARS-COV2. GRL-0617 se lie de manière réversible à SCOV2-PLpro et perturbe son interaction avec le gène 15 stimulé par l'interféron (ISG15). GRL-0617 affecte faiblement (> 100  $\mu$ M) les activités de MERS PLpro, une série de protéases à cystéine, y compris les enzymes humaines de déubiquitination HAUSP, USP18, UCH-L1, UCH-L3, et une protéase de type papaïne (PLP2) du coronavirus humain NL63. Le GRL-0617 supprime l'infection par le SARS-COV2 et favorise l'immunité antivirale [32].



**Figure II.6.A :** Structures chimiques du GRL-0617. **B :** site actif de PLpro avec l'inhibiteur GRL-0617 [24].

Le développement des inhibiteurs de PLpro en est encore à ses balbutiements. Les inhibiteurs de PLpro les plus prometteurs sont la série de composés à base de naphthalène GRL0617 (figure II.7) [33].



**Figure II.7.** Inhibiteurs SARS-COV-2 PLpro. (A) Analogues GRL0617. (B) Inhibiteurs de PLpro structurellement disparates identifiés à partir d'un criblage à haut débit ou d'un criblage de repositionnement de médicaments de composés bioactifs [33].

En conclusion, l'inhibition de l'activité de PLpro peut arrêter la réplication virale et perturber son rôle dans l'évasion de la réponse immunitaire de l'hôte, ce qui en fait une excellente cible antivirale [25].

# **CHAPITRE III : CRIBLAGE VIRTUEL**



## CHAPITRE III. CRIBLAGE VIRTUEL

### I. Généralité

La découverte d'une seule molécule médicamenteuse nécessite l'investissement d'énormes quantités de ressources par les sociétés biotechnologiques et pharmaceutiques. Dans la découverte de médicaments modernes, l'ordinateur joue un rôle vital en économisant à la fois du temps et de l'argent, car le processus suit une approche biologique chimique *in silico*. Le processus de développement de nouveaux médicaments est un processus en plusieurs étapes qui comprend essentiellement le développement préclinique et les essais cliniques, Le processus de découverte de médicaments est très coûteux. De nos jours, des millions de produits chimiques sont disponibles et peuvent être examinés par criblage à haut débit (*HTS*) *in vitro* afin de déterminer leur capacité à se lier, à inhiber ou à activer les cibles biomoléculaires. En raison des grandes campagnes *HTS* coûteuses, les chercheurs concentrent leurs efforts sur le développement des outils informatiques pour effectuer un criblage virtuel (CV) avant *l'HTS*. L'objectif du développement préclinique est l'identification et l'optimisation d'un composé principal pour une action biologique ciblée dans laquelle l'investigation pharmacocinétique et pharmacodynamique du médicament est impliquée et commence généralement par des examens *in silico* qui conduisent ensuite à études *in vitro* et *in vivo*[34].

Le criblage virtuel est devenu une partie intégrante du processus de découverte de médicaments ces dernières années. En rapport avec le concept plus général et poursuivi de longue date de recherche dans les bases de données. Le CV a été en grande partie un jeu de nombres axé sur des questions telles que la façon dont nous pouvons filtrer l'énorme espace chimique de  $> 10^{60}$  composés concevables à un nombre gérable qui peut être synthétisé, acheté et testé[35]. Généralement, le criblage virtuel comporte trois phases [37] :

- La phase de préparation de la chimiothèque incluant la mise à disposition (ou pas) de vrais actifs/inactifs dans le cas d'un apprentissage supervisé ;
- La phase d'exécution proprement dite qui varie en fonction de la taille de la chimiothèque et de la connaissance de la structure 3D d'une protéine cible ;
- La phase d'analyse des résultats conduisant à la définition de touches.

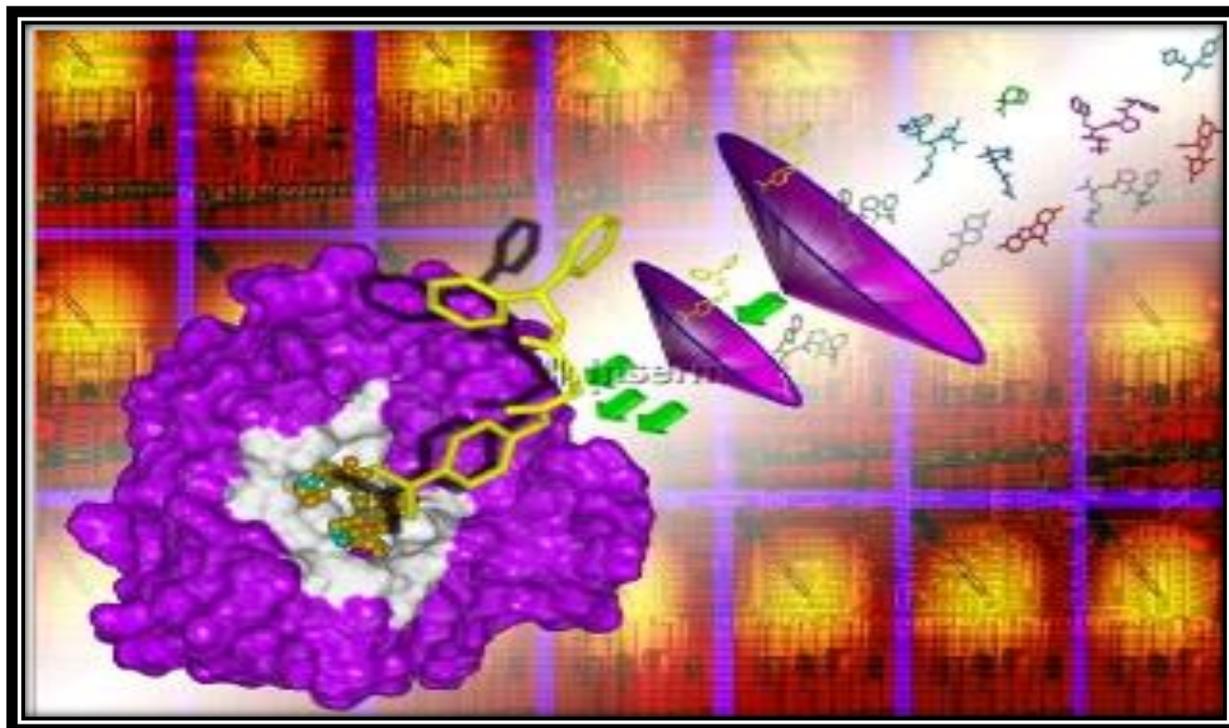


Figure III.1. Criblage virtuel *in silico*[36].

## II. Outils du criblage virtuel

### II.1. Structure de la cible protéique

Trois méthodes expérimentales permettent désormais de déterminer la structure des protéines : la résonance magnétique nucléaire (RMN), la microscopie électronique et la cristallographie aux rayons X. Cette dernière technique est responsable de la plupart des structures produites à partir d'une base de données structurale librement accessible, connue sous le nom de "*Protein Data base*" (PDB). La diversité des ligands cristallisés dans une protéine est une information précieuse qui permet d'observer les résidus clés des interactions d'une part, mais aussi la variabilité structurale 3D de la protéine d'une molécule à l'autre [39].

### II.2. Importance des chimiothèques dans la recherche de nouvelles substances thérapeutique

La sélection des molécules cribler est une étape importante pour réduire le nombre de composés à tester, plusieurs filtres peuvent être utilisés pour que les composés sélectionnés valident certains critères. Ceux-ci peuvent être divisés en deux catégories : les premiers consistent à réduire l'espace chimique étudié, en se basant sur l'exclusion des composés avec de mauvaises propriétés physico-chimiques et des caractéristiques incompatibles avec celles de

molécules d'intérêt pharmaceutique, les deuxièmes permettent de choisir les composés à cribler selon la structure et les propriétés des ligands de références [40].

Les fournisseurs de produits chimiques possèdent des chimiothèques, des bases de données moléculaires qui contiennent toutes les informations sur les molécules qu'ils vendent. Certains fournisseurs fournissent des bibliothèques prétraitées pour le criblage, telles que des groupes de molécules présentant des caractéristiques physico-chimiques de médicaments "drug-like", des molécules sélectionnées par diversité, des produits naturels et leurs dérivés, ou se concentrant sur certaines cibles thérapeutiques comme la  $\beta$ -lactamase, la neuraminidase et l'acétylcholinestérase [40].

Les chimiothèques virtuelles stockent les molécules qui la composent dans différents formats chimioinformatiques. Il existe deux principaux types de formats : 2D et 3D (sdf, mol2, pdb...etc), qui se différencient des premiers en fournissant des informations sur la conformation des molécules dans l'espace. Choisissez le format approprié en fonction de la capacité de stockage disponible et de l'utilisation prévue de la chimiothèque. Cependant, ce choix n'est pas nécessairement limitatif, car il existe des logiciels disponibles pour convertir les bibliothèques chimiques d'un format à un autre [45].

### II.3. Logiciels de criblage virtuel

De nombreuses publications ont proposé des expériences parallèles menées avec un bon nombre de programmes de criblage virtuel (qui sont commercialisés ou non). Les plus fréquemment utilisés sont : Gold, Flex X, Glide et Surflex (voir tableau III.1).

Ces programmes reposent sur l'ajustement de fragments, commençant par construire un modèle «en négatif» du site actif dit « *protomol* » dans le cas du programme Surflex. Ce modèle est constitué par le volume accessible du site et les points d'interaction tels que les sites de liaison hydrogène, les charges ou les sites lipophiles. Ces algorithmes sont dits aussi de fragmentation / reconstruction 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. Cette approche qui permet notamment un criblage rapide de vastes bibliothèques de composés[49].

Tableau III.1. Quelques exemples de logiciels de criblage virtuel [49].

Logiciel	Site internet	Algorithme de recherche	Fonction d'affinité
<b>FlexX</b>	<a href="http://www.biosolveit.de/FlexX/">http://www.biosolveit.de/FlexX/</a>	Fragmentation / reconstruction	Basée sur les connaissances
<b>Glide</b>	<a href="http://www.schrodinger.com/Products/glide.html">http://www.schrodinger.com/Products/glide.html</a>	Monte Carlo	Empirique
<b>Gold</b>	<a href="http://www.ccdc.cam.ac.uk/products/life_sciences/gold/">http://www.ccdc.cam.ac.uk/products/life_sciences/gold/</a>	Génétique	Basée sur les champs de force
<b>Surflex-Dock</b>	<a href="http://www.biopharmics.com/products.html">http://www.biopharmics.com/products.html</a>	Hammerhead modifié	Empirique

### III. Différentes stratégies du criblage virtuel

Les techniques de criblage virtuel sont basées sur la nature des informations expérimentales disponibles concernant le système étudié, qui peuvent être la structure tridimensionnelle de la cible biologique « *structure-based* » ou la structure et les propriétés d'un ou de plusieurs composés bioactifs de référence « *ligand-based* ». La combinaison de ces deux stratégies dans une même étude permet d'augmenter les chances d'identifier de nouvelles touches ou d'optimiser de molécules connues [40].

#### III.1. Criblage virtuel « *ligand-based* »

Dans le processus de criblage virtuel basé sur un ligand, la molécule principale biologiquement active la plus efficace est détectée à l'aide d'une recherche de similarité structurelle ou topologique ou de similarité pharmacophorique. La prise en compte de plusieurs critères tels que la structure ainsi que la forme du fragment individuel ou les propriétés électrostatiques de la molécule effectue les comparaisons de similarité. Les *leads* générés sont classés en fonction de leur score de similarité, obtenu à l'aide de différentes méthodes ou algorithmes [48].

### III.2. Criblage virtuel « *structure-based* »

Le criblage virtuel basé sur la structure (*Structure-based virtual screening* ou *SBVS*) est une technique *in silico* robuste, utile et prometteuse pour la conception de médicaments. Ce criblage tente de prédire la meilleure interaction entre les ligands contre une cible moléculaire pour former un complexe. En conséquence, les ligands sont classés en fonction de leur affinité avec la cible, et les composés les plus prometteurs sont présentés en tête de liste. Les méthodes *SBVS* nécessitent que la structure 3D de la protéine cible soit connue afin que les interactions entre la cible et chaque composé chimique puissent être prédites *in silico*. Dans cette stratégie, les composés sont sélectionnés dans une base de données et classés selon leur affinité pour le site récepteur. L'utilisation du *SBVS* présente des avantages et des inconvénients [44].

Parmi les avantages [44] :

- Il y a une diminution du temps et des coûts impliqués dans le criblage de millions de petites molécules.
- L'existence physique de la molécule n'est pas nécessaire, elle peut donc être testée par ordinateur avant même d'être synthétisée.
- Il existe plusieurs outils disponibles pour aider le *SBVS*.

Les inconvénients peuvent être mis en évidence comme suit [44] :

- Certains outils fonctionnent mieux dans des cas spécifiques, mais pas dans des cas plus généraux.
- Il est difficile de prédire avec précision la position de liaison correcte et la classification des composés en raison de la difficulté de paramétrer la complexité des interactions de liaison ligand-récepteur.
- Il peut générer des faux positifs et des faux négatifs.

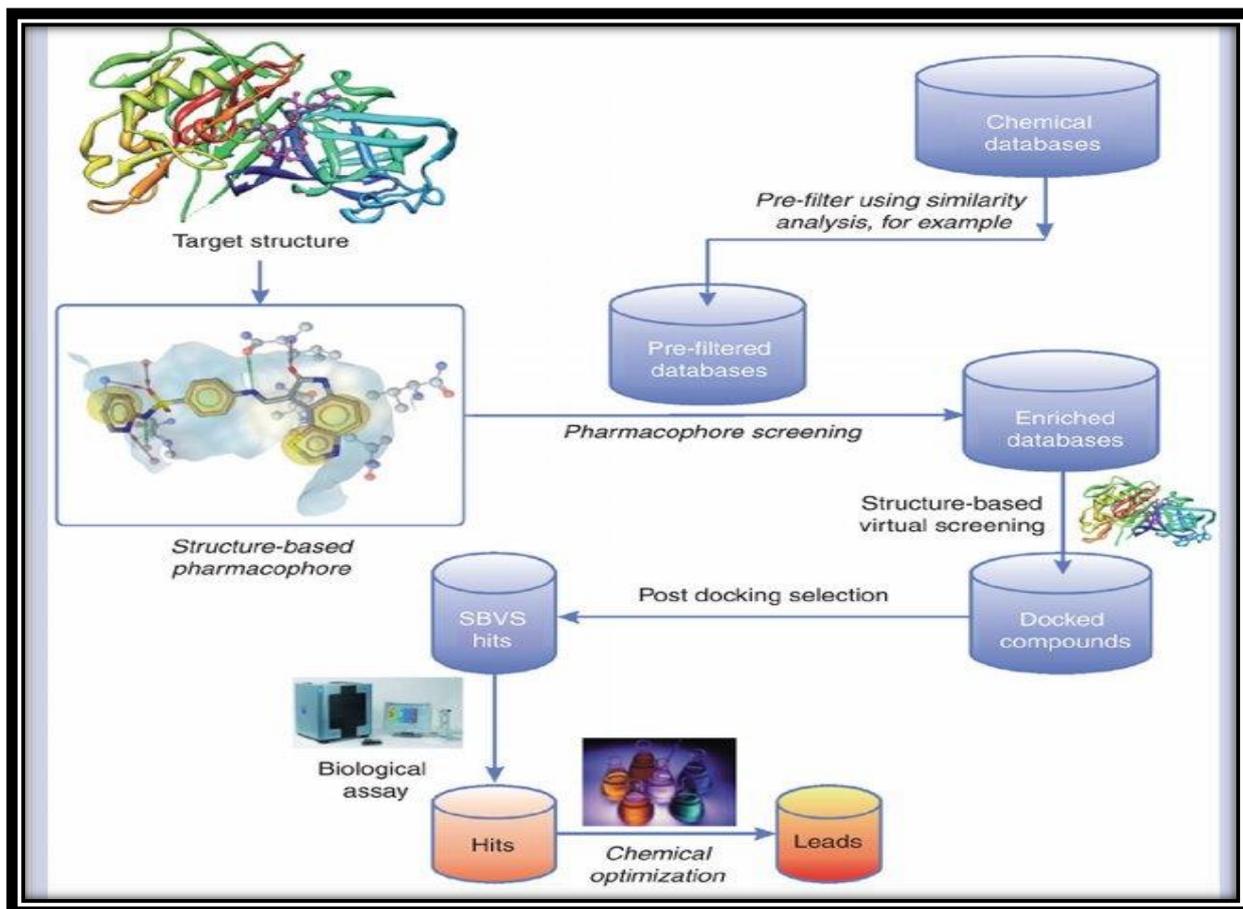


Figure III. 2. Criblage virtuel « *Structure-based* » [46].

## IV. Docking moléculaire

### IV.1. Définition

Le docking moléculaire est une technique utilisée pour prédire la meilleure correspondance entre deux molécules lorsqu'elles sont liées l'une à l'autre afin de générer un complexe stable. Ces informations sur l'orientation préférée peuvent en outre être utilisées pour prédire l'affinité de liaison ou la force d'association entre deux molécules à l'aide de fonctions de notation. Par conséquent, docking est utile pour prédire la force ainsi que le type de signal produit. La méthode docking moléculaire permet la recherche d'une manière automatique et plus objective pour évaluer l'adéquation entre deux molécules [34].

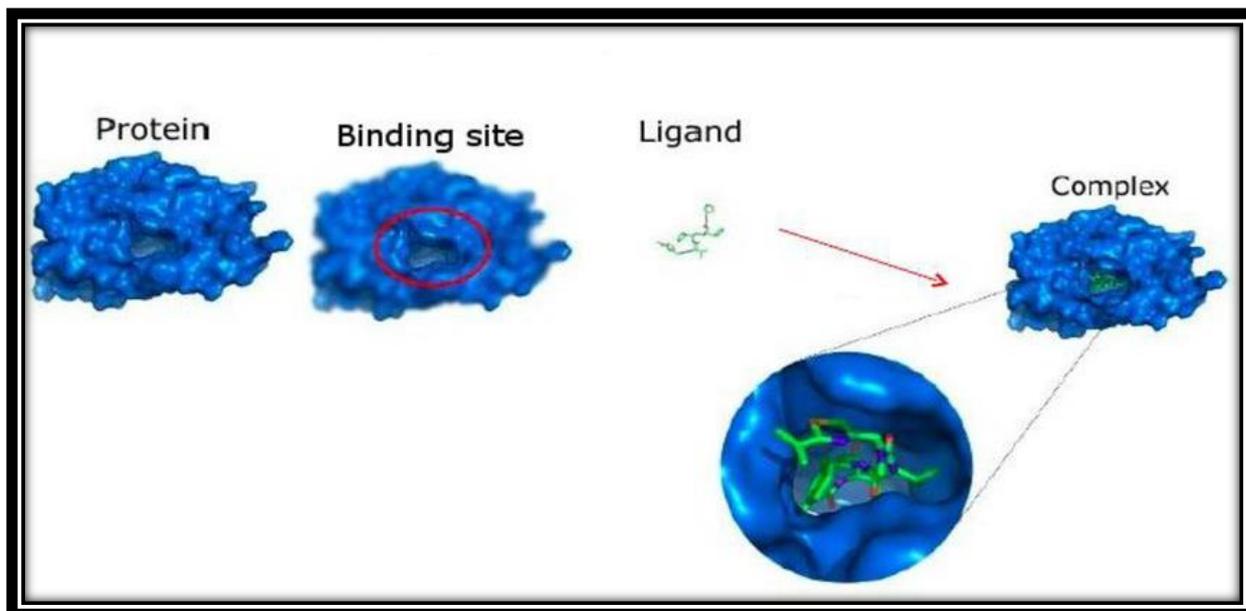


Figure III.3. Docking protéine-ligand [50].

## IV.2. Principe

Un programme de docking doit être capable de générer les modes de liaison attendus pour des ligands dont la position adoptée au sein du site actif est connue dans un temps raisonnable. Pour cela, leur protocole se compose de deux étapes [41].

- La première (*docking*) est l'étape de sélection, qui consiste à placer le ligand sur le site actif d'une protéine, avec échantillonnage de la conformation, de la position et de l'orientation (pose) possible, ne gardez que celles qui représentent des modes interactifs le plus favorable.
- La seconde (*scoring*) est l'étape de classification, qui comprend l'évaluation de l'affinité entre ligands et protéines, et notation des poses obtenues à ce stade. Ce score tiendra probablement la meilleure posture de tous suggéré.

## IV.3. Méthodes de docking moléculaire

La méthodologie de base de l'amarrage moléculaire peut être classée en trois catégories [38] :

- **Docking par ajustement induit** : le ligand et le récepteur sont flexibles. Le ligand se lie de manière flexible au site actif du récepteur pour maximiser les forces de liaison entre eux. Il met en œuvre le concept de complémentarité entre protéine et ligand.

- **Docking de verrouillage et de clé** : sur la base de la théorie de verrouillage et de clé, le ligand et le récepteur sont rigides et présentent une liaison étroite. Il définit le concept de base de la complémentarité tridimensionnelle.
- **Ensemble docking** : cette approche explique la flexibilité et la complexité des états conformationnels des protéines. Structures protéiques multiples utilisées comme un ensemble pour l'amarrage avec le ligand.

#### IV.4. Protocole générale de docking

Nous décrivons la procédure en quatre étapes adoptée dans cette étude pour effectuer le docking moléculaire [43].

- **Sélection cible** : la structure cible doit être déterminée expérimentalement. Après avoir sélectionné le modèle, il doit être préparé en retirant les molécules d'eau de la cavité, en stabilisant les charges, en remplissant les résidus manquants et en générant les chaînes latérales, le tout en fonction des paramètres disponibles. Le récepteur doit être à ce point biologiquement actif et dans un état stable.
- **Sélection et préparation des ligands** : le type de ligands choisis pour l'amarrage dépendra de l'objectif. Il peut être obtenu à partir de diverses bases de données, tel que ZINC et PubChem. Il est souvent nécessaire d'appliquer des filtres pour réduire le nombre de molécules à ancrer.
- **Docking** : c'est la dernière étape, où le ligand est ancré sur le récepteur et les interactions sont vérifiées. La fonction de scoring génère un score en fonction du meilleur ligand sélectionné.
- **Évaluation des résultats docking** : le succès des algorithmes d'amarrage dans la prédiction d'une pose de liaison de ligand est normalement mesuré en termes RMSD entre les positions des atomes lourds observées expérimentalement des ligands et celles prédites par l'algorithme.

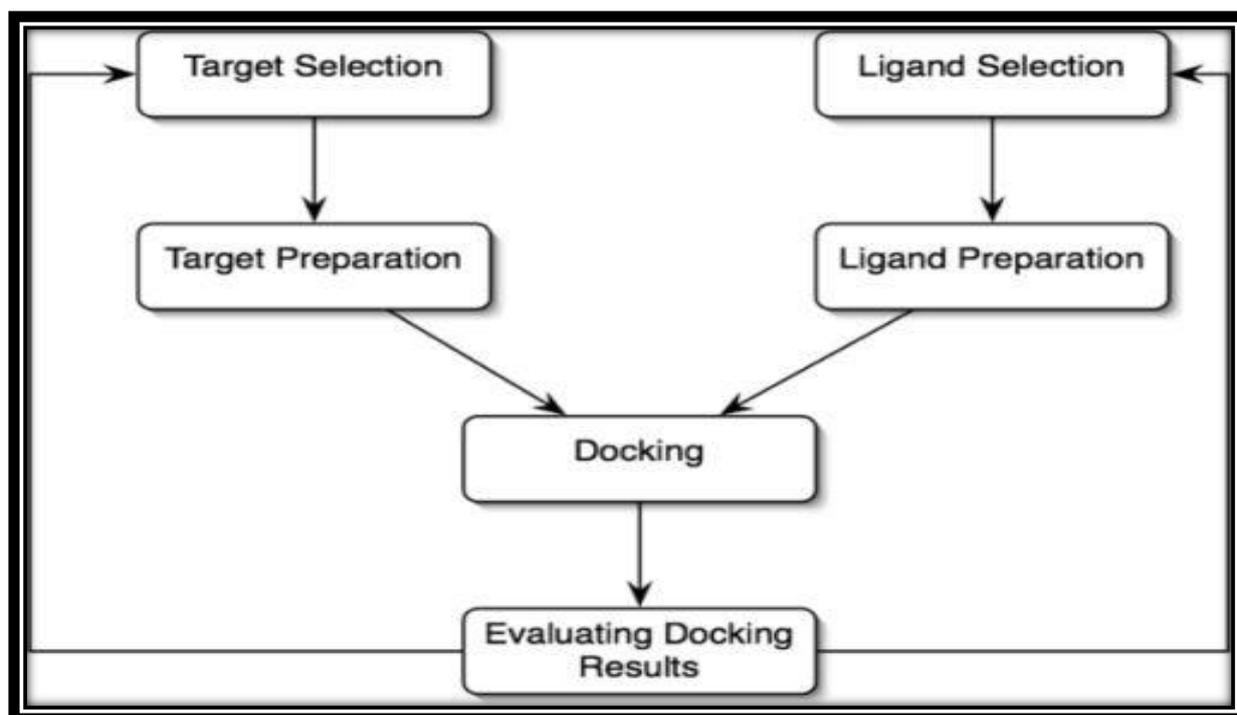


Figure III.4. Les étapes de docking moléculaire [53].

## V. Programme de docking moléculaire Surflex-dock

Le programme Surflex comprend deux parties. L'une pour des études de similarité tridimensionnelle des molécules, l'autre pour des études de docking. Seule la partie docking est abordée ici. Elle se base sur la construction d'une pseudo-molécule comme cible sur laquelle aligner le ligand. Il s'agit de fragments de molécules placés dans le site actif de manière à occuper idéalement et de manière redondante celui-ci à partir de critères morphologiques. Il est important de préciser que cette pseudo molécule peut être construite à partir du ligand dans le site, à condition de bénéficier d'une structure[47].

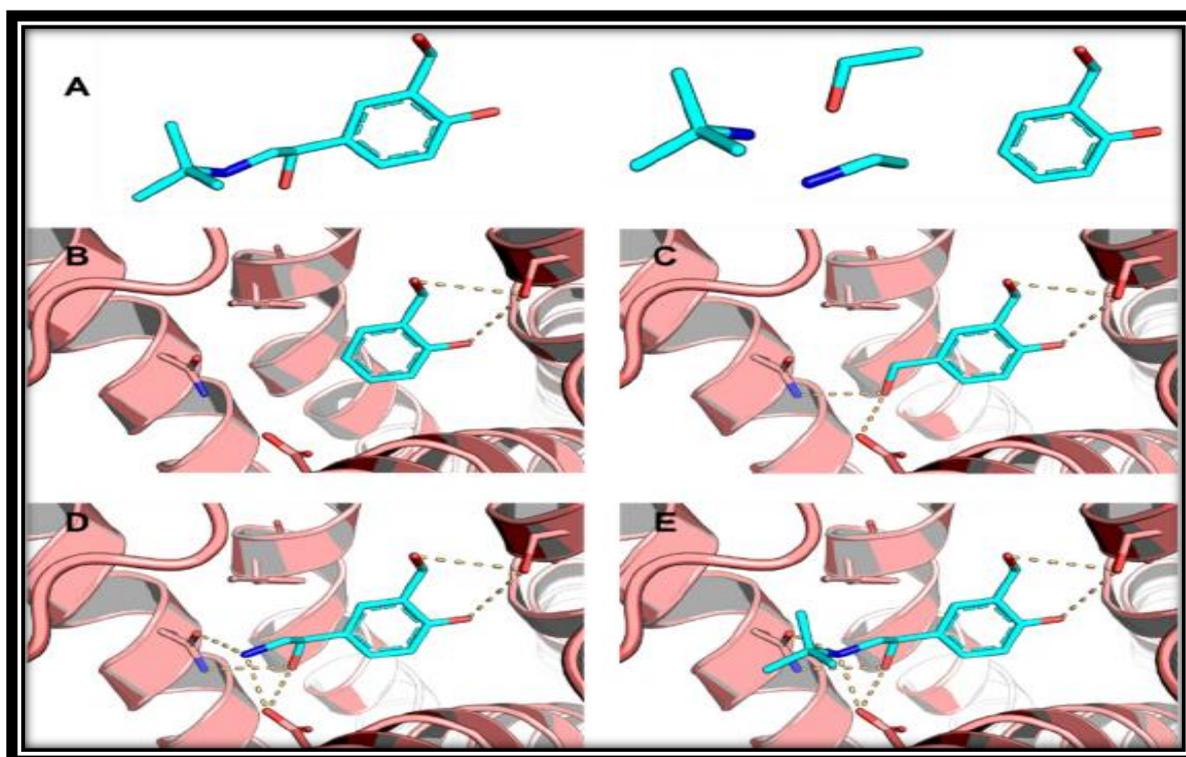
L'emploi du logiciel Surflex-dock passe par trois étapes :

- Choisir de quelle manière définir le site actif, soit à partir d'un ligand soit à partir du récepteur ;
- Construire la pseudo-molécule ou le *protomol* sur lequel docker le ligand ;
- Enfin lancer le processus de docking.

Plusieurs approches ont été utilisées pour chercher les conformations les plus stables des deux composants d'un complexe. Nous avons détaillé l'approche incrémentale qui est la stratégie utilisée par surflex-dock, le programme de docking moléculaire utilisé dans notre étude [51].

- **Approche incrémentale**

Le ligand est fragmenté à partir de liaisons rotatives en divers segments. L'un des segments est ancré à la surface du récepteur. Le fragment d'ancrage est généralement considéré comme le fragment qui montre le maximum d'interactions avec la surface du récepteur, a un nombre minimum de conformations alternées et assez rigide comme le système d'anneaux. Une fois la base/l'ancrage établie, l'étape suivante consiste à ajouter chacun des fragments étape par étape. Idéalement, ces fragments sont ajoutés en premier qui ont une plus grande chance de montrer des interactions comme la liaison hydrogène car ils sont de nature directionnelle et sont responsables de la spécificité du ligand. De plus, les liaisons hydrogène conduisent à une prédiction plus précise de la géométrie. Une fois qu'un fragment particulier est ajouté, les poses avec le moins d'énergies sont considérées pour la prochaine itération [52].



**Figure III.5.** La méthode de construction incrémentale. (A) Le ligand est divisé en plusieurs fragments; (B) Le fragment d'ancrage est amarré au site de liaison de la cible moléculaire ; (C) Le fragment suivant est amarré après le fragment d'ancrage ; (D et E) Les autres fragments sont ancrés séquentiellement pour construire le ligand entier dans sa conformation de liaison. Les résidus dans le site actif sont représentés en bâton. Les liaisons hydrogène sont indiquées par des lignes pointillées [42].

# **CHAPITRE IV :**

## **MATERIEL ET METHODES**



## CHAPITER IV. MATERIEL ET METHODES

L'objectif principal de notre travail est de découvrir de nouveaux inhibiteurs de la PLpro en faisant appel à une étude théorique (*in silico*) qui permet de prédire l'affinité, les types d'interaction ainsi que les propriétés physicochimiques des molécules que retiendra l'opération de criblage virtuel de la chimiothèque PubChem.

Afin de réaliser cette étude, nous avons utilisées des programmes installées dans deux microordinateurs qui présentent les performances suivantes :

- **HP** avec une mémoire de 4 GO, et un processeur 2.20 GHZ sous le système d'exploitation Windows 7 ; 64 bits version 2009.
- **Condor** avec une mémoire de 2 GO, et un processeur 1.80 GHZ sous le système d'exploitation Windows 7 ; 32 bits version 2009.

### I. Tests de fiabilité de programme Surflex-dock

Avant d'aborder l'étude de l'inhibition de la PLpro par diverses molécules, nous avons tenté d'évaluer la fiabilité de programme Surflex-dock utilisé. Pour cela, nous avons fait appel à deux testes différents :

- L'écart quadratique moyen ou le RMSD ;
- Analyse visuelle.

#### I.1.Le RMSD (*Root Mean Square Deviation*)

Correspond à la moyenne de la déviation de chacun des atomes par rapport à ceux de la molécule d'origine. Le mieux possible signifie que la valeur du RMSD entre la pose du ligand calculée par le logiciel et la conformation dans le complexe expérimental est la plus petite possible. Le positionnement, c'est-à-dire l'identification correcte du site de liaison sur la protéine, l'orientation et la conformation du ligand influent sur la valeur du RMSD[54, 55].

Le ratio admis est une différence maximum de deux angströms au-delà de laquelle la prédiction est considérée comme non pertinente. La norme actuelle pour évaluer les performances d'un programme de docking est de faire un test à partir de plusieurs centaines de complexes protéines-ligands cristallisés [54, 55].

Un programme de docking moléculaire a été testé, Surflex-dock (v 1.3, 2005). Ce test a été réalisé sur 150 complexes protéine-ligand disponibles au niveau de la PDB et leurs RMSD déterminés. La prédiction est acceptable si la valeur du RMSD ne dépasse pas 2Å. La liste des

complexes protéines-ligands étudiés sera retrouvée dans l'annexe1. Les calculs de RMSD ont été réalisés par Surflex-dock grâce à la commande suivante :

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

Si le RMSD dépasse 2Å on peut utiliser la commande suivante :

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

## I.2. Analyses visuelles

C'est le deuxième test utilisé pour juger la performance du programme Surflex-dock, montre que les modèles des ligands simulés par ce programme sont correctement placés dans le site actif de la cible enzymatique. Cela témoigne davantage de la fiabilité du logiciel utilisé pour prédire les différentes interactions enzyme-inhibiteur.

## II. Criblage virtuel

Au cours de cette recherche, nous avons calculé *in silico* par criblage virtuel basé sur la structure de la cible enzymatique l'affinité d'une collection de 555 composés téléchargés à partir de la chimiothèque PubChem vis-à-vis du site actif de la PLpro. Les inhibiteurs retenus suite au criblage sont ensuite soumis à une analyse visuelle pour visualiser les différentes liaisons qu'elles impliquent avec le site actif de cette enzyme.

### II.1. Préparation de la cible enzymatique

Notre cible thérapeutique traitée est la PLpro. Avant son utilisation par Surflex-dock, le complexe de cette enzyme 7JN2 doit subir quelques modifications nécessaires. L'enzyme étant formée d'une seule chaîne, nous avons éliminé les molécules d'eau et les différents ligands pour permettre une utilisation plus aisée de l'enzyme avec un site actif libre lors du docking moléculaire. Le programme Surflex-dock fonctionne mieux avec le format .mol2, donc les molécules sont protonées et transformées dans le format .mol2 par le programme Open Babel.

## II.2. Préparation des ligands

La structure du ligand est téléchargée directement à partir de la PDB dans le format .sdf ; ensuite les différents ligands (inhibiteurs de référence ou autres) sont convertis au format .mol2 pour les adapter au docking par Surflex-dock par l'intermédiaire de programme Open Babel.

## III. Banques de données

### III.1. PDB (*Protein Data Bank*)

La PDB est une banque mondiale de données de structures tridimensionnelles 3D de macromolécules biologiques, essentiellement des protéines et des acides nucléiques. Ces structures sont principalement déterminées par les deux méthodes la cristallographie à rayons X et la RMN. Les structuralistes du monde entier sont tenus d'y déposer leurs données expérimentales, de cette façon elles appartiennent au domaine public. La consultation des structures est gratuite et peut se faire via l'internet. L'adresse électronique de cette banque est <https://www.rcsb.org/>[56].

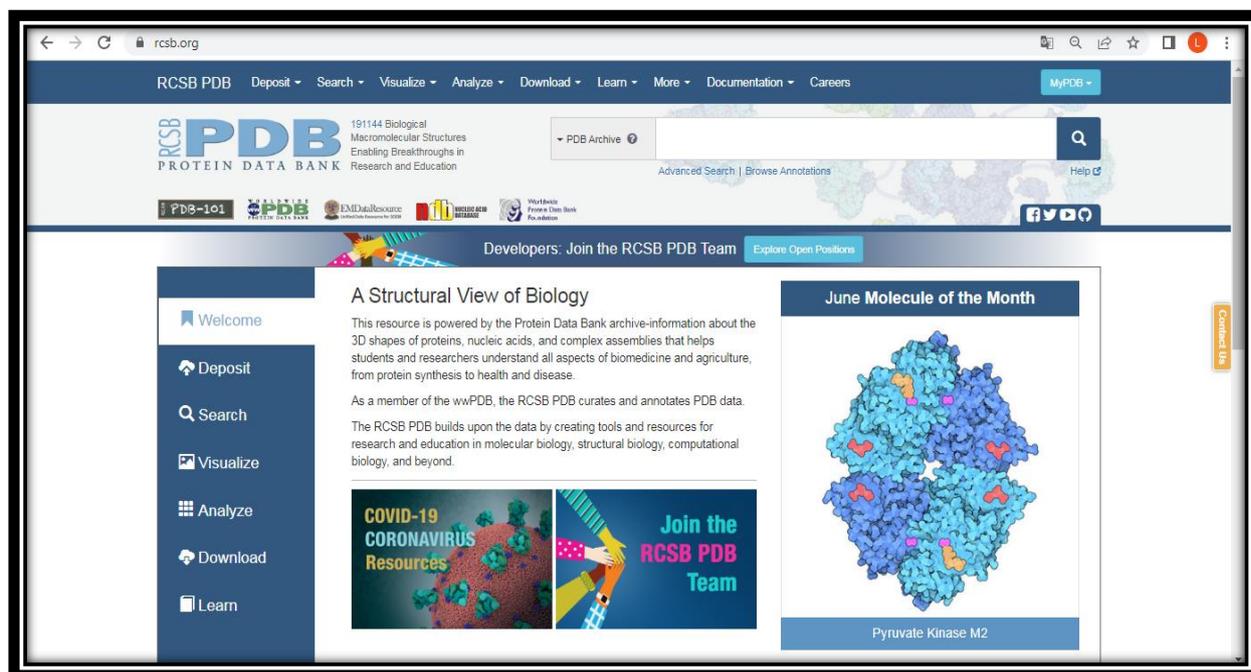


Figure IV.1. Page internet de la PDB.

### III.2. PubChem

C'est une banque de données américaine de molécules chimiques gérée par *le National Center for Biotechnology Information (NCBI)*, branche de la Bibliothèque nationale de médecine des États-Unis sous l'autorité de la *National Institute of Health (NIH)*. La

chimiothèque PubChem englobe plusieurs millions de composés mis en ligne, pour chaque substance une grande quantité de données de divers ordres chimique, biochimique, pharmacologique, production, toxicologie. De plus, tous ces informations concernant les différents composés chimiques sont gratuitement accessibles via le site Web (<https://pubchem.ncbi.nlm.nih.gov>) [58,65]. Cette banque a été utilisée pour obtenir des similaires structuraux de notre inhibiteur de référence.

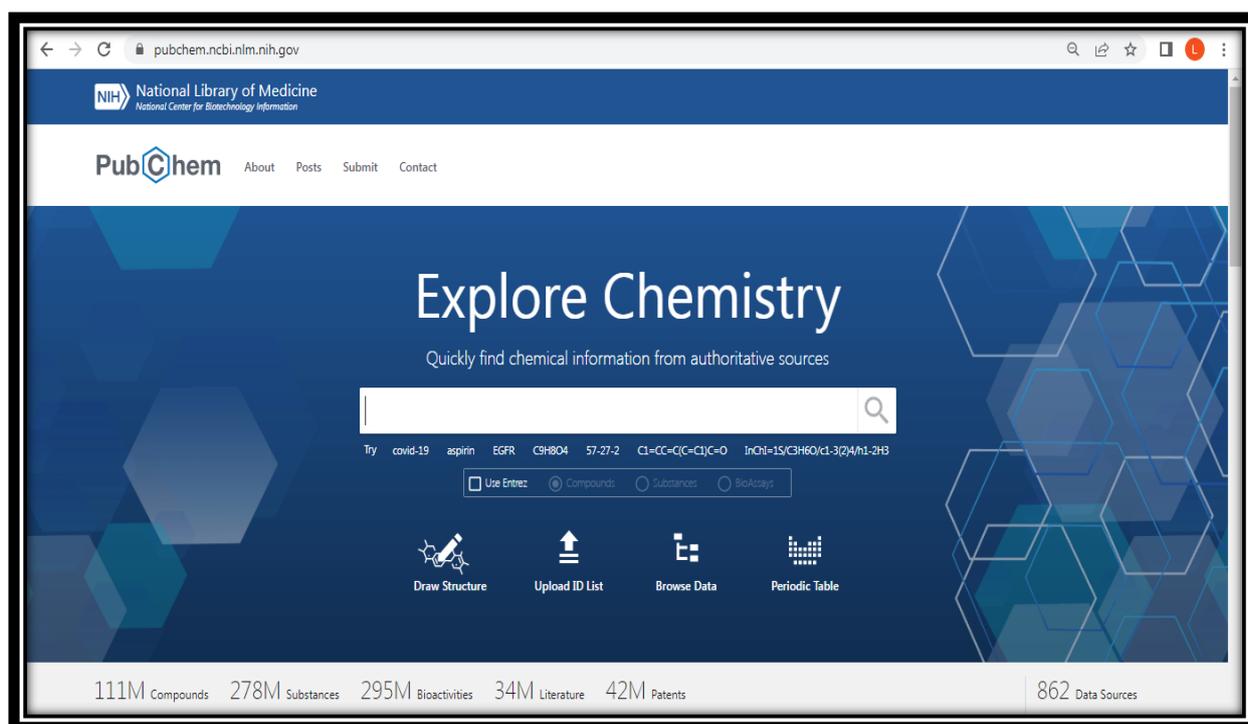


Figure IV. 2. Page internet de la chimiothèque PubChem.

## IV. Programmes utilisés

### IV.1. Surfex-dock (1.3 .2005)

Le logiciel Surfex-dock est un algorithme d'amarrage moléculaire flexible entièrement automatique qui combine la notation fonction du système d'amarrage Hammerhead avec un moteur de recherche qui s'appuie sur une méthode de similarité moléculaire comme moyen de générer rapidement des poses putatives appropriées pour fragments moléculaires [60, 59].

Le docking moléculaire par ce logiciel est réalisé en trois étapes [47] :

- Choix de la manière d'identifier le site actif soit à partir d'un ligand soit à partir du récepteur, et construire d'un *protomol* qui va servir de cible aux différents ligands ;
- Docking d'un ou de plusieurs ligands ;
- Traitement des résultats.

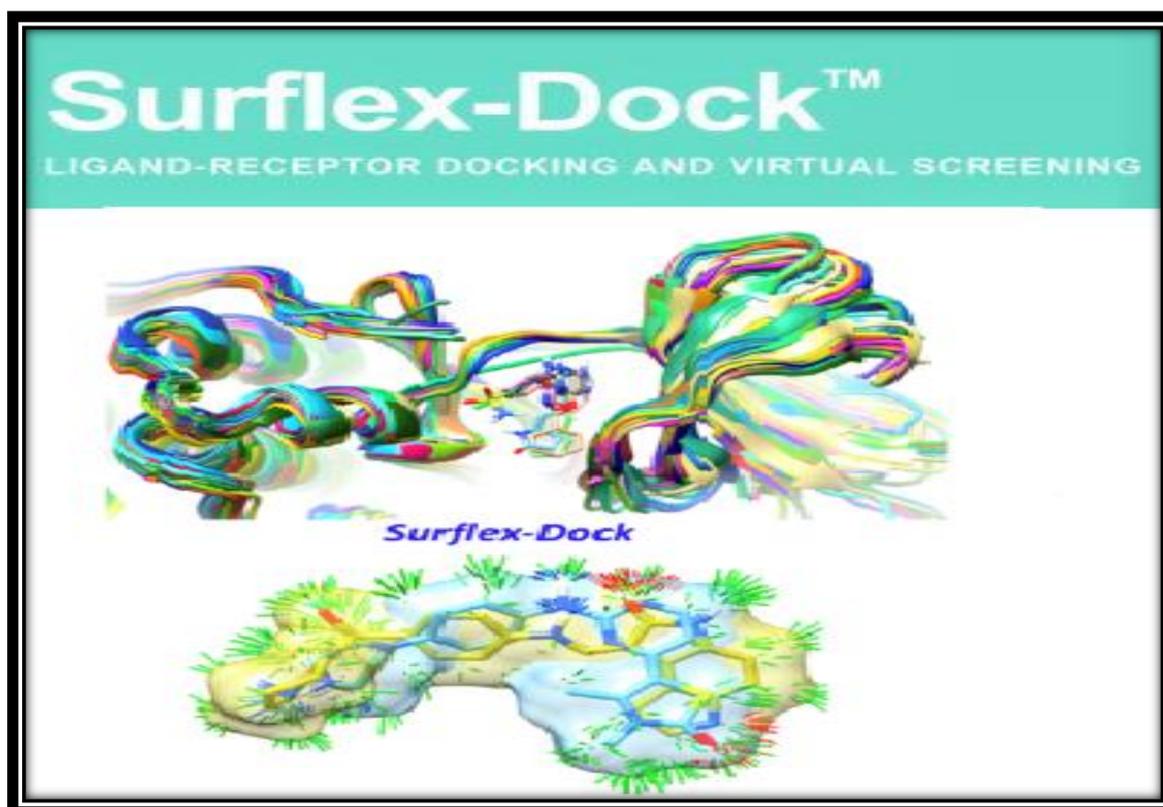


Figure IV.3. Page internet de programme Surflex-dock [61].

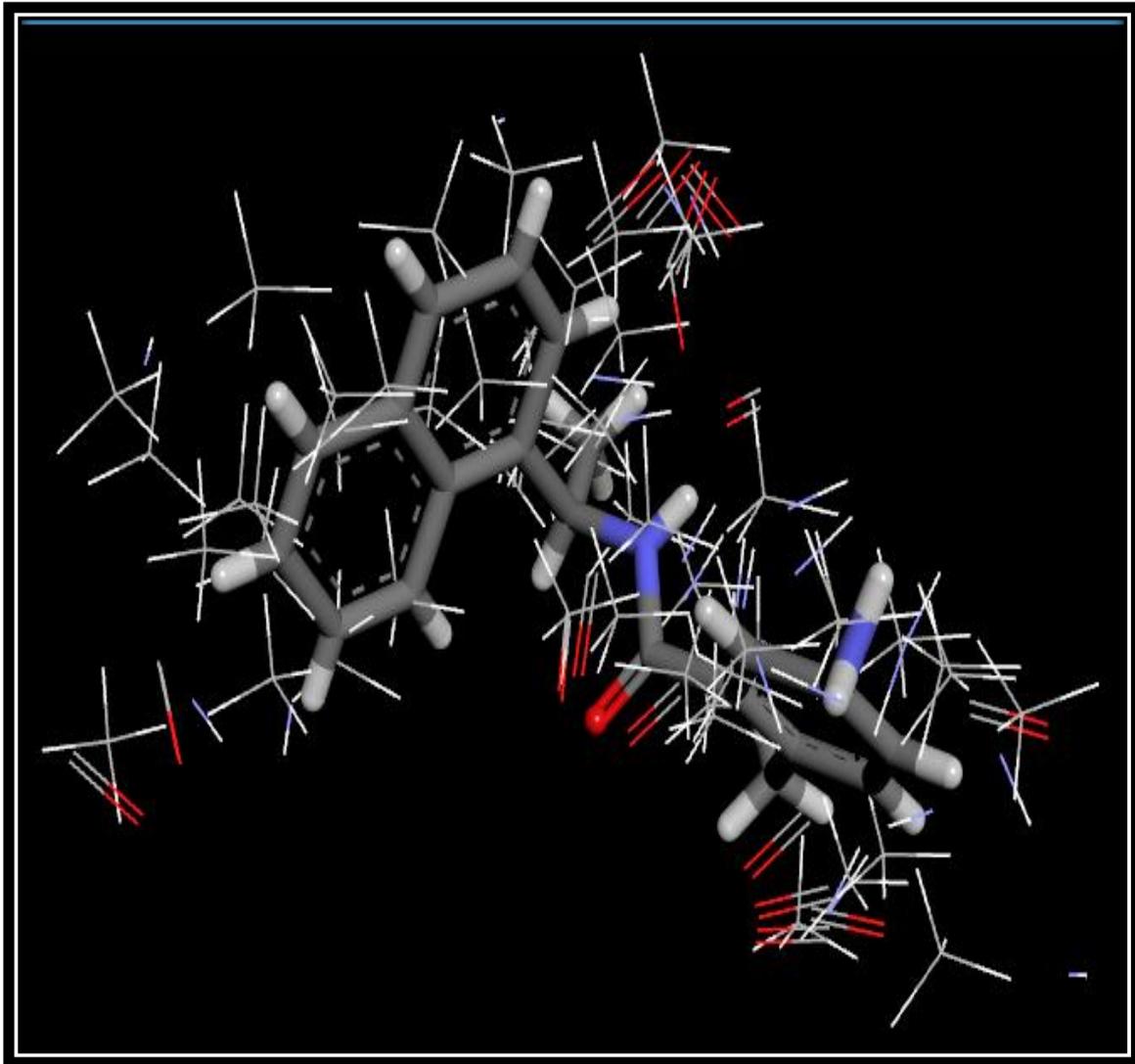
### Etapes de docking par Surflex-dock

Nous avons réalisé un docking moléculaire par le logiciel Surflex-dock selon le protocole suivant :

- **Génération de *Protomol*** : au cours de docking, nous avons utilisé une pseudo-molécule comme la cible à laquelle se rallie le ligand dans le site actif de la cible enzymatique. Ce dernier utilise trois types de fragments ( $\text{CH}_4$ ,  $\text{C}=\text{O}$  et  $\text{NH}$ ) qu'il place dans le site actif sous différentes positions et qu'il évalue au fur et à mesure. Les fragments les mieux scorés sont aussitôt assemblés pour former un *protomol* qui couvre l'intégralité du site actif [54, 58]. Pour la création de *protomol* nous avons utilisé la commande :

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

Avant le lancement des calculs, il est conseillé de visualiser d'abord le *protomol* avec l'inhibiteur de référence pour s'assurer que son volume et son extension sont ajustés au site actif de l'enzyme étudiée.



**Figure IV.4.** Génération du protomol pour la PLpro autour de l'inhibiteur TTT.

- **Docking d'une seule molécule** : l'arrimage moléculaire exige un ligand, un *protomol* et une cible enzymatique. Pour le docking d'une seule molécule nous avons utilisé la commande [55, 59] :

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

Le résultat de docking par surflex-dock donne par défaut, les dix meilleures «conformés» en  $M^{-1}$  au format .mol2. Chaque conformation possède trois scores :

- Le premier «final-0 » représente l'affinité ;

- Le second correspond au degré de pénétration impropre du ligand dans la protéine appelé «*crash score*», plus ce score est proche de 0 plus l'interaction est favorable ;
- Le troisième appelé «*polar score*» correspond au niveau de contribution des interactions polaires.

```

.....
Polishing: .....
Times (total 8): search 0 align1 2 merge+chain 6 polish 0
Nconfs (final 61): 3 12 2172 2762
[ret_conf 1.34] [ret_conf 4.85] [ret_conf 4.89] [ret_conf 3.40] [ret_conf 4.48]
[ret_conf 4.16] [ret_conf 1.34] [ret_conf 1.37] [ret_conf 4.83] [ret_conf 4.87]
[ret_conf 4.61] [ret_conf 4.20] [ret_conf 4.78] [ret_conf 3.65] [ret_conf 4.20]
[ret_conf 4.68] [ret_conf 3.65] [ret_conf 4.87] [ret_conf 4.85] [ret_conf 4.87]
[ret_conf 1.31] [ret_conf 4.68] [ret_conf 4.30] [ret_conf 4.98] [ret_conf 4.00]
[ret_conf 4.71] [ret_conf 4.68] [ret_conf 4.13] [ret_conf 4.61] [ret_conf 4.93]
[ret_conf 4.91] [ret_conf 5.07] [ret_conf 4.00] [ret_conf 5.07] [ret_conf 4.22]
[ret_conf 3.92] [ret_conf 4.68] [ret_conf 3.88] [ret_conf 4.61] [ret_conf 4.91]
[ret_conf 3.92] [ret_conf 2.47] [ret_conf 1.43] [ret_conf 4.00] [ret_conf 4.78]
[ret_conf 2.07] [ret_conf 4.22] [ret_conf 3.88] [ret_conf 1.93] [ret_conf 4.68]
final-0: 5.07 <crash -2.52 (self -1.05)>
final-1: 5.07 <crash -2.52 (self -1.05)>
final-2: 5.06 <crash -2.61 (self -1.12)>
final-3: 5.05 <crash -2.61 (self -1.16)>
final-4: 5.05 <crash -2.61 (self -1.16)>
final-5: 5.05 <crash -2.61 (self -1.16)>
final-6: 5.04 <crash -2.60 (self -1.16)>
final-7: 5.04 <crash -2.60 (self -1.16)>
final-8: 5.04 <crash -2.60 (self -1.16)>
final-9: 5.04 <crash -2.57 (self -1.11)>

```

Figure VI.5.Exemple d'un résultat de docking moléculaire donné par Surflex-dock.

## IV.2. Open Babel

C'est un Système utilisé pour la conversion des fichiers d'un format à un autre format. La conversion concerne tous les types de formats de molécules telle que : pdb, sdf, smiles, mol2...etc. L'objectif de ce programme est d'offrir aux utilisateurs une base pour rechercher, convertir, analyser ou stocker des données issues de la modélisation moléculaire, de la chimie, des matériaux à l'état solide et de la biochimie. Open Babel est obtenu à partir de cette adresse électronique : <http://openbabel.org> [56].

## IV.3. Discovery studio visualizer version 2021

Discovery Studio est un logiciel complet pour l'analyse et la modélisation de structures moléculaires, de séquences et d'autres données pertinentes pour les chercheurs en sciences de la vie. Le produit comprend des fonctionnalités d'affichage et de modification des données ainsi que des outils permettant d'effectuer une analyse de base des données [62].

Le programme Discovery studio visualizer est un visualiseur gratuit qui peut être utilisé pour ouvrir des données générées par d'autres logiciels de la gamme de produits Discovery Studio ainsi que d'autres logiciels tel que Surflex-dock. Il est conçu pour offrir un environnement interactif pour la visualisation et l'édition de structures moléculaires, de séquences, de données de réflexion de rayons X, de scripts et d'autres données. Il fournit également un ensemble complet de visualiseurs pour afficher des tracés et d'autres représentations graphiques de données. L'application fonctionne sous Windows et Linux et est un environnement de bureau entièrement intégré qui permet d'accéder aux fonctionnalités standard du système d'exploitation telles que le système de fichiers, le presse-papiers et les services d'impression. Cette application démarre en cliquant sur l'icône sur le bureau [62].

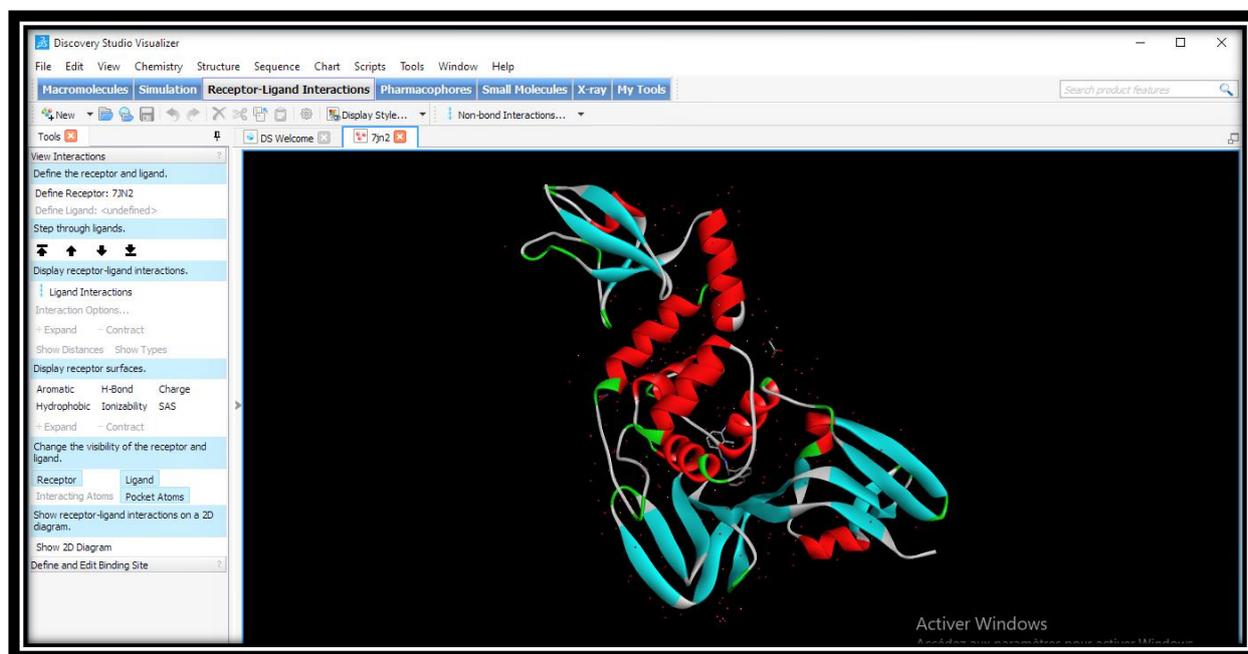


Figure IV.6. Interface graphique de logiciel Discovery studio visualizer.

## V. Filtrage ADME-Tox

Chaque médicament éventuel doit de se conformer à plusieurs critères de base, tels son faible coût de production, être soluble, stable, brevetable, mais doit aussi se conformer à des barèmes associés à ses propriétés pharmacologiques d'absorption, de distribution, de métabolisme, d'excrétion et de toxicité. La liste des aspects à inclure est très longue mais on admet communément que les molécules validant au moins deux conditions de la règle de *Lipinski* sont des candidats potentiels [63]. Ces règles sont :

- Le poids moléculaire du composé ne doit pas être supérieur à 500g/mol ;
- Le coefficient de partition (logP) doit être  $\leq$  à 5 ;

- Le nombre de donneurs de liaisons hydrogène doit être  $\leq$  à 5 ;

De plus, deux critères sont introduits par *Veber* [64] complémentaire à la règle de *lipinski* concernant les propriétés physico-chimiques des médicaments qui sont :

- La surface polaire (PSA, polar surface area) du composé doit être inférieure à  $140\text{\AA}^2$  ;
  - Le nombre de liaisons de rotation « rotatable bonds » doit être inférieur à 15.
- **Le coefficient de partage (logP ou clogP) :** est égal au logarithme du rapport des concentrations de la substance étudiée dans l'octanol et dans l'eau.  $\text{LogP} = \text{Log}(\text{Coct}/\text{Ceau})$ . Cette valeur permet d'appréhender le caractère hydrophile ou hydrophobe (lipophile) d'une molécule. En effet, si LogP est positif et très élevé, cela exprime le fait que la molécule considérée est bien plus soluble dans l'octanol que dans l'eau, ce qui reflète son caractère lipophile, et inversement. Diverses études statistiques ont mis en évidence les valeurs optimales de logP pour qu'un composé soit absorbé par le corps humain :
    - Pénétration au cerveau : 2,0 ;
    - Absorption orale : 1,8 ;
    - Absorption sub-linguale : 5 ;
    - Absorption percutanée : 2,6.

En effet, en complément de ces règles, d'autres critères sont pris en compte dans la sélection de molécules potentiellement candidates [65] :

- Nombre d'halogènes  $\leq 7$  ;
- Chaînes alkyles  $\leq$   $-(\text{CH}_2)_6\text{-CH}_3$  ;
- Nombre de cycles  $\leq 6$  ;
- Pas de grands cycles de plus de 7 membres ;
- Au moins un atome d'azote ou d'oxygène dans la molécule.

# CHAPITRE V :

## RESULTATS ET DISCUSSIONS



## CHAPITRE V. RESULTATS ET DISCUSSIONS

### I. Tests de fiabilité du programme de docking Surfex-dock

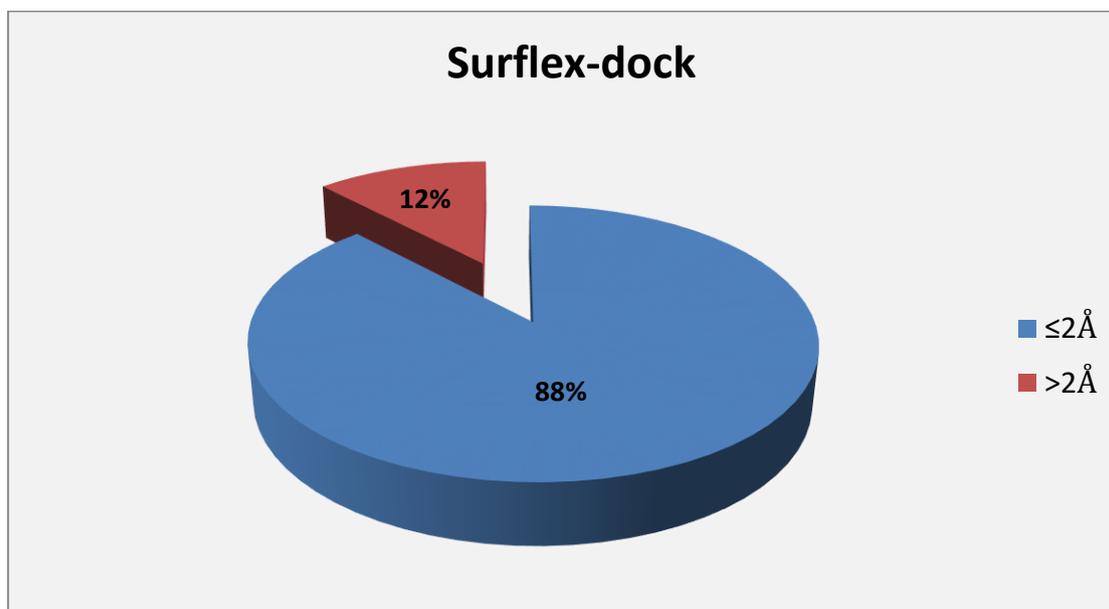
Afin de réaliser le protocole de docking moléculaire développé dans la partie « méthode » de chapitre IV, on a jugé utile d'évaluer tout d'abord la performance du programme Surfex-dock utilisé dans cette étude en faisant appel aux tests suivants :

- Ecart quadratique moyen ou le RMSD (*RootMean Square Deviation*)
- Analyse visuelle

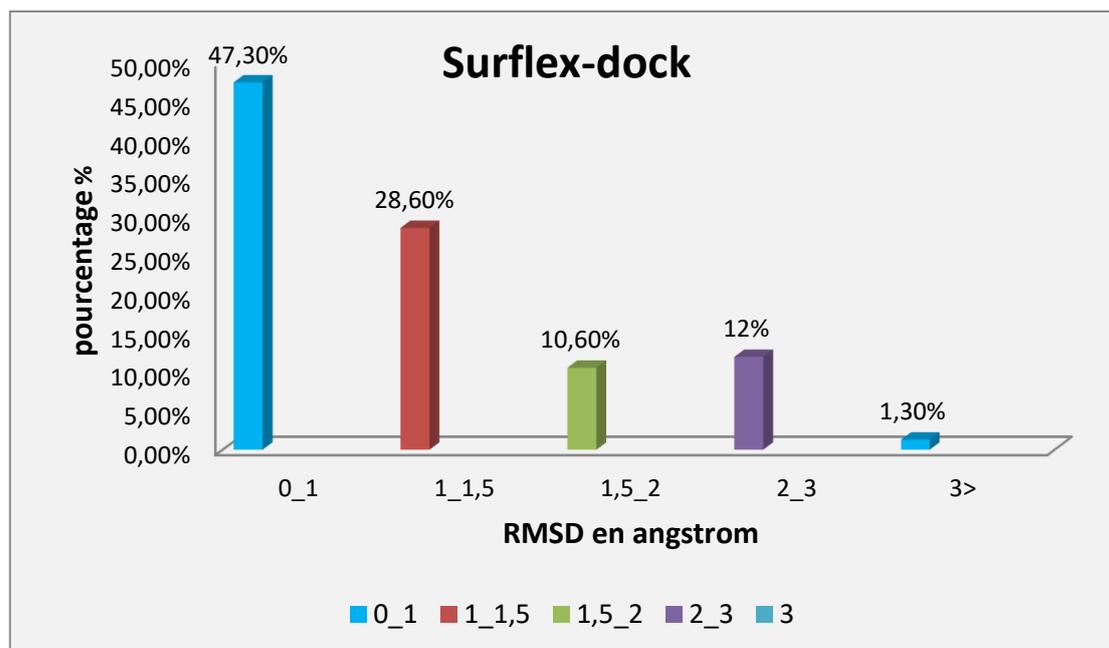
#### I.1. Test RMSD

L'objectif principal de ce test consiste à déterminer la bonne position du ligand prédite ou calculée par le logiciel de docking avec la référence co-cristallisée. Un programme de docking moléculaire est qualifié de performant si les valeurs de RMSD sont inférieures ou égales à une certaine valeur (généralement 2,0Å) [66-68].

Le test de fiabilité du programme Surfex-dock par le RMSD, a été réalisé en utilisant 150 complexes protéine-ligand que nous avons téléchargés de la PDB et regroupés dans le tableau 1 des annexes.



Graphique V.1. Résultat global du RMSD en %.



**Graphe V.2.** Résultats en % obtenus par Surflex-dock à divers intervalles de RMSD (Å).

Dans le graphe V.1, les résultats sont donnés en pourcent (%), à deux intervalles de RMSD représentés par deux couleurs. Il apparaît clairement sur ce graphe que le logiciel Surflex-dock est suffisamment performant parce que sur 150 complexes testés, nous avons trouvé 132 complexes ont une valeur de RMSD inférieure ou égales à 2 Å et sont considérés comme résultats positifs. Ce qui donne à un pourcentage de 88 % et que seulement 12 % des résultats sont supérieurs à 2 Å.

L'annexe montre que le RMSD minimum dans l'ensemble testé est de 0,06Å alors que le maximum est de 8,85 Å. D'après Ce graphe, dans le graphe V.2 la majorité des résultats positifs se situe dans l'intervalle 0 – 1. Conformément aux travaux de vieth M. *et al* (1998) [68].

Notre résultat confirme les résultats obtenus par Chikhi A, (2007) où le programme Surflex-dock avec deux autres logiciels (FlexX et Arguslab) ont été testé, il a montré que 88.4% des valeurs de RMSD calculé par le programme Surflex-dock sont inférieures ou égales à 2 Å [47].

Le même programme, utilisé par Chikhi A., Bensegueni A, (2008) présente 63% de résultats positifs à partir de 300 complexes testés [67]. Nous remarquons aussi que nos résultats sont identiques à celles obtenus par le travail de Teniou S, (2012) et Boucherit H, (2012) où la

performance de programme Surflex-dock a été évalué. Il a montré une performance pour reproduire les tests expérimentaux avec 85.8% des valeurs de RMSD sont inférieures ou égales à 2 Å pour le premier et 66% pour le second [72, 55].

De plus, ce test a été conforté par une analyse visuelle afin de confirmer davantage nos résultats.

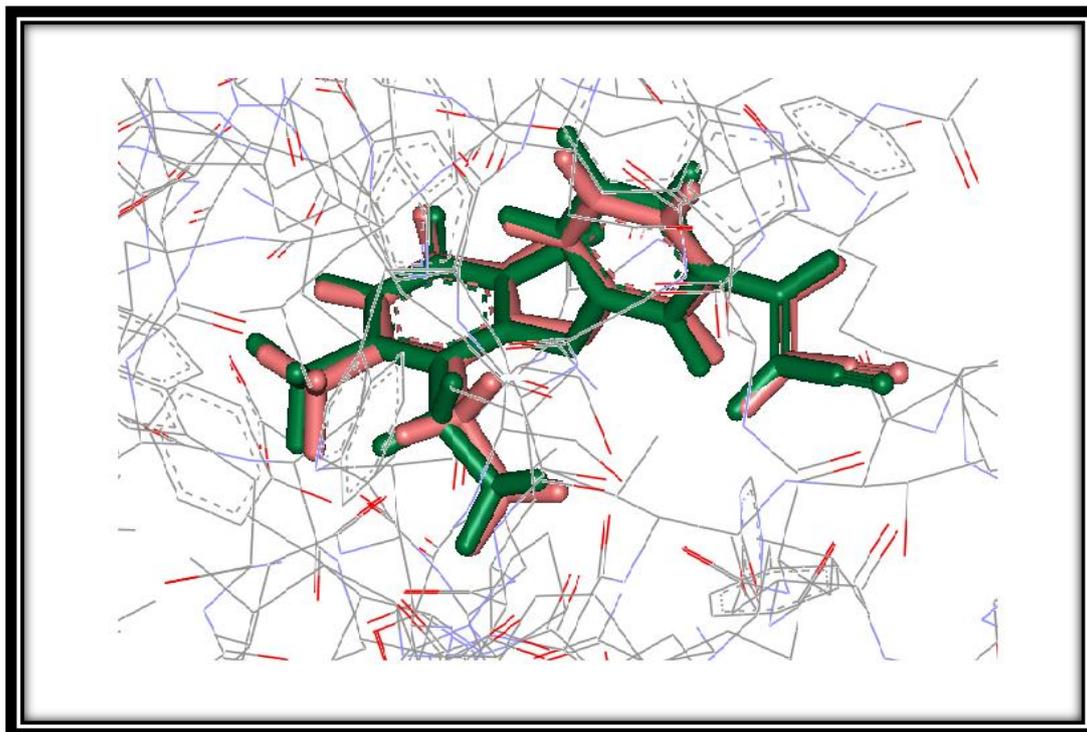
## I.2. Analyses Visuelles

L'analyse visuelle est une étape essentielle et complémentaire pour juger la performance du logiciel Surflex-dock, et confirmer les résultats du test RMSD. Aussi, elle permet une vérification de la superposition du ligand expérimental par rapport à celui docké.

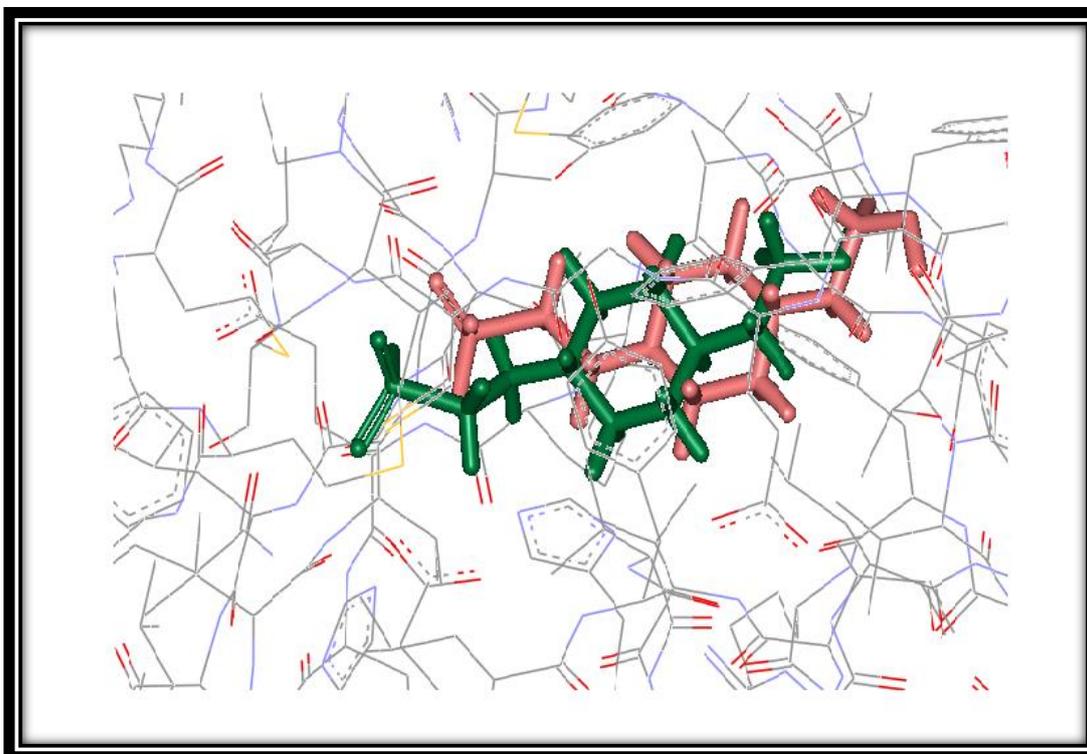
Les résultats du docking de trois complexes, le 2B5J (le meilleur), le 1VEY (l'intermédiaire) et le 1ICJ (le plus mauvais), sont montrés dans les figures V.1, V.2 et V.3 suivantes. La conformation des ligands générés par le docking moléculaire (vert) est comparée avec la pose expérimentale du ligand (rose), est montrée dans le tableau V.1 ci-dessous.

**Tableau V.1.** Valeurs de RMSD des 3 complexes étudiés.

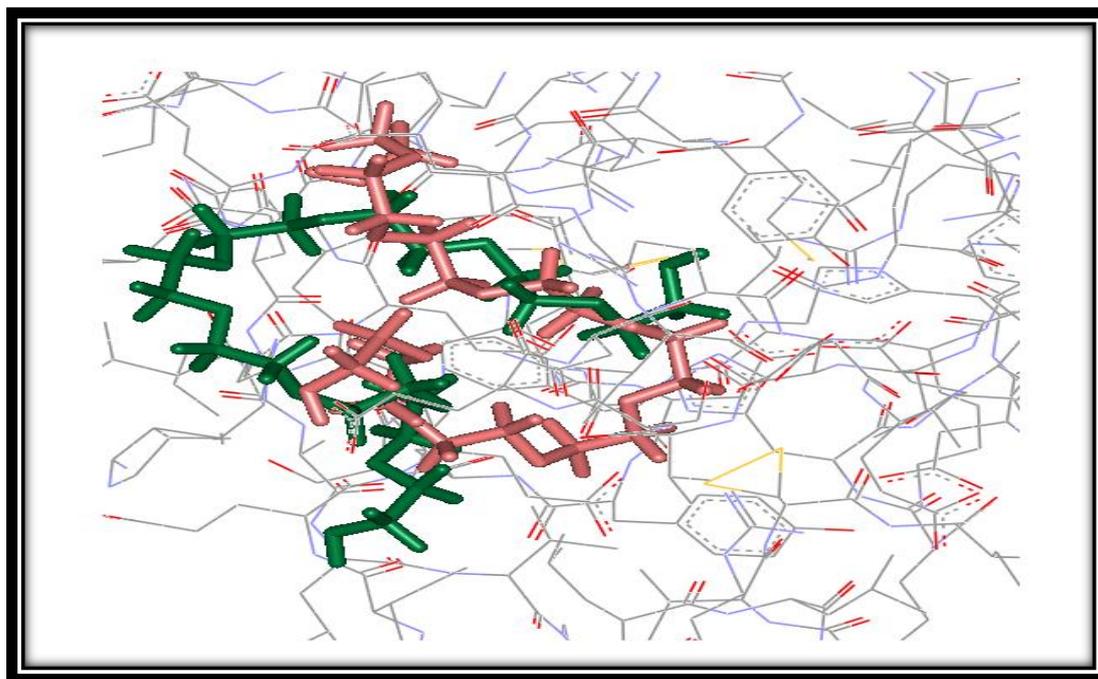
Code PDB de complexe Protéine-ligand	Code de ligand	RMSD (Å)
2B5J	3AC	0.22
1VEY	EPE	2.55
1ICJ	2PE	8.95



**Figure V.1.** Superposition du ligand donné par rayon-X (coloré en rose) et du finale-0 calculé par Surflex-dock (coloré en vert) du complexe 2B5J.



**Figure V.2.** Superposition du ligand donné par rayon-X (coloré en rose) et du finale-0 calculé par Surflex-dock (coloré en vert) du complexe 1VEY.



**Figure V.3.** Superposition du ligand donné par rayon-X (coloré en rose) et du finale-0 calculé par Surflex-dock (coloré en vert) du complexe 1ICJ.

Dans le premier cas, le complexe 2B5J avec un RMSD de 0.28 Å, la superposition est presque parfaite (figure V.1). Dans le deuxième cas, avec un RMSD de 2.55 Å, elle l'est moins (figure V.2), alors que pour le troisième complexe, avec un RMSD de 8.95 Å (figure V.3), il n'y a aucune superposition remarquable, c'est à dire le ligand générée est éloigné de ligand de référence.

### I.3. Conclusion

Les tests RMSD et l'analyse visuelle nous permet de conclure, que le logiciel Surflex-dock peut être utilisé pour prédire *in silico* les interactions enzyme-inhibiteur. Donc, nous pouvons l'utiliser pour cribler virtuellement la chimiothèque PubChem dans le but d'identifier des nouveaux inhibiteurs de la cible enzymatique PL protéase.

## II. Etude des interactions intervenant dans l'inhibition de la PL protéase par diverses molécules provenant de la PDB

Pour étudier le mode d'interaction de différents inhibiteurs de la PL protéase (PLpro) qui se trouve dans la PDB avec le site actif de notre cible enzymatique, nous avons utilisé le programme Surflex-dock, pour effectuer le docking moléculaire de ces molécules.

### II.1. Choix du complexe 7JN2

La structure de l'enzyme PLpro, objet de notre étude, provient de la PDB qui représente la principale source des complexes protéine-ligand sous forme de complexes cristallographiques 3D. Parmi les nombreux complexes de notre cible enzymatique. Nous avons choisi celui portant le code 7JN2 car elle constitue un compromis entre une bonne résolution (1.93 Å) et la présence d'un inhibiteur dans le site actif. Les caractéristiques de cette enzyme sont regroupées dans le tableau V.2 ci-dessous.

**Tableau V.2.** Principales caractéristiques du code 7JN2.

Code	Résolution (Å)	Facteur R	Classification	Nombre de chaîne	Nombre d'AA par chaîne	Nombre d'atomes par chaîne
7JN2	1.93 Å	0.209	EC : 3.4.22	1(A)	315	2502

### II.2. Site actif de la 7JN2

Le site actif des enzymes correspond à l'ensemble des acides aminés qui rentrent en contact avec le substrat à transformer. Il s'agit généralement d'une petite zone de l'enzyme. Il comporte deux sous parties : le site de fixation et le site catalytique [69]. Le site actif de la PLpro est composé de 13 acides aminés orientés de la manière suivante :

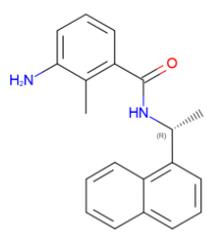
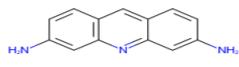
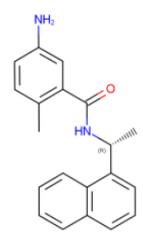
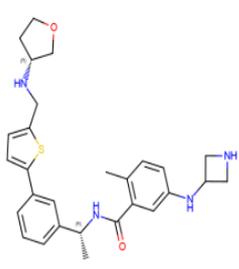
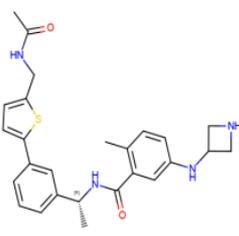
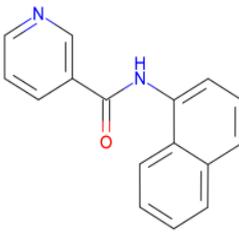
PRO247\_PRO248\_LYS157\_ASP164\_GLU167\_TYR264\_ASN267\_TYR273\_CYS270\_THR301\_ASP302\_TYR268\_GLN269.

### II.3. Interaction 7JN2-inhibiteurs

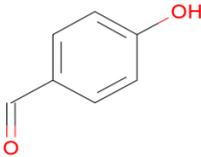
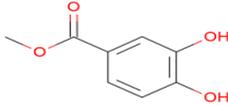
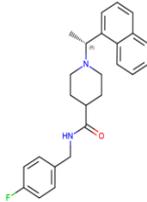
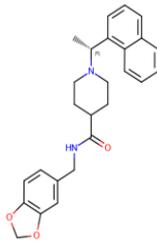
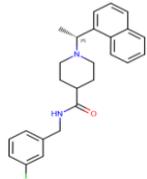
Dans le but d'identifier de nouveaux inhibiteurs de la PLpro. Nous avons réalisé une étude théorique dite *in silico* qui consiste à réaliser d'abord un criblage virtuel basé sur la structure de la cible enzymatique des inhibiteurs qu'on retrouve au niveau de la PDB par le programme de docking moléculaire Surflex-dock et de proposer le meilleur inhibiteur de cette enzyme.

Le docking de 18 molécules téléchargées à partir de la banque PDB avec le format *mol2* portant des codes différents est effectué sur la structure cristallographique 7JN2. Les structures et les résultats du docking par Surflex-dock de ces inhibiteurs sont représentés dans le tableau V.3 suivant.

Tableau V.3. Les valeurs d’Affinité des complexes 7JN2-inhibiteur étudiés.

Composé	Affinité (M <sup>-1</sup> )	Nom chimique	Structure
<b>Y41</b>	5.07	3-amino-2-méthyl-N-[(1R)-1-(naphthalen-1-yl)éthyl]benzamide	
<b>PRL</b>	3.56	PROFLAVIN	
<b>TTT</b>	<b>9.59</b>	5-amino-2-méthyl-N-[(1R)-1-naphthalen-1-ylethyl]benzamide	
<b>Y97</b>	7.48	5-(azetidino-3-ylamino)-2-méthyl-~ {N}-[(1~{R})-1-[3-[5-[[[(3~{R})-oxolan-3-yl]amino]méthyl] thiophen-2-yl] phényl] éthyl] benzamide	
<b>Y61</b>	6.00	N-[(1R)-1-(3-{5-[(acétylamino) méthyl] thiophen-2-yl} phényl) éthyl]-5-[(azetidino-3-yl) amino]-2-méthylbenzamide	
<b>L30</b>	5.03	N-(naphthalen-1-yl) pyridine-3-carboxamide	

<b>JWX</b>	7.27	(1R)-N-[(1H-indol-3-yl)methyl]-N-methyl-1-(naphthalen-1-yl)ethan-1-amine	
<b>XR8</b>	4.87	5-[(azetidin-3-yl) amino]-2-methyl-N-[(1R)-1-(3-{5-[(pyrrolidin-1-yl) methyl] thiophen-2-yl} phenyl) ethyl] benzamide	
<b>VBV</b>	7.09	5-(acryloylamino)-2-methyl-N-[(1R)-1-(naphthalen-1-yl)ethyl]benzamide	
<b>JW9</b>	<b>8.29</b>	4-({methyl[(1R)-1-(naphthalen-1-yl)ethyl]amino} methyl)pheno 1	
<b>Y96</b>	5.92	5-[(E)-(hydroxyimino)methyl]-2-methyl-N-[(1R)-1-(naphthalen-1-yl)ethyl]benzamide	
<b>XT7</b>	6.00	5-[(azetidin-3-yl) amino]-N-[(1R)-1-{3-[5-({[(1S, 3R)-3-hydroxycyclopentyl] amino} methyl) thiophen-2-yl] phenyl} ethyl]-2-methylbenzamide	

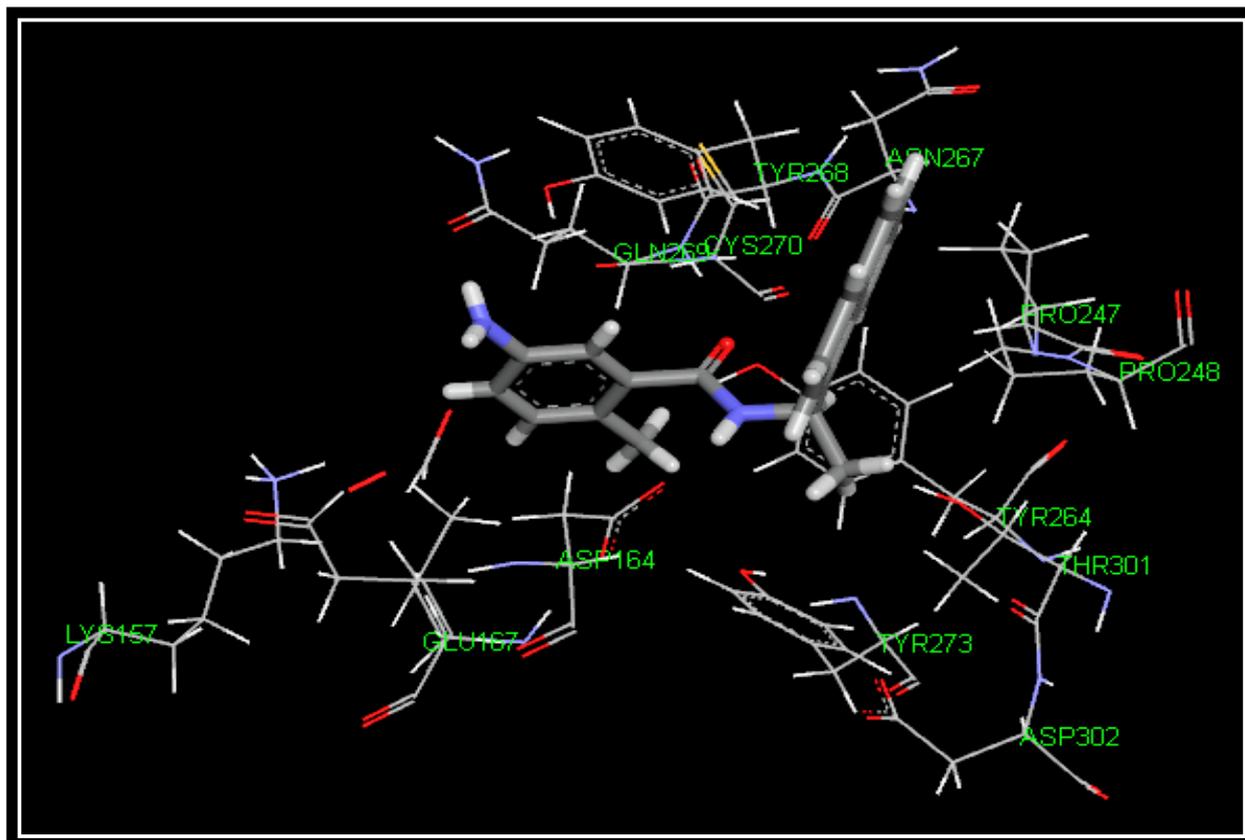
<b>HBA</b>	3.31	P- HYDROXYBENZALDEHY DE	
<b>HE9</b>	4.79	methyl 3,4-bis (oxidanyl) benzoate	
<b>A7L</b>	4.99	N-(2,5- dihydroxybenzylidene)- thiosemicarbazone	
<b>P85</b>	3.40	N-[(4-fluorophenyl)methyl]- 1-[(1R)-1-naphthalen-1- ylethyl]piperidine-4- carboxamide	
<b>GRM</b>	7.91	N-(1,3-benzodioxol-5- ylmethyl)-1-[(1R)-1- naphthalen-1-ylethyl] piperidine-4-carboxamide	
<b>S88</b>	7.38	N-(1,3-benzodioxol-5- ylmethyl)-1-[(1R)-1- naphthalen-1-ylethyl] piperidine-4-carboxamide	

Le tableau V.3. montre que parmi les 18 inhibiteurs dockés, le composé TTT dont le nom chimique 5-amino-2-méthyl-N-[(1R)-1-naphthalen-1-ylethyl]benzamide avec une Affinité égale à  $9.59 \text{ M}^{-1}$  et le composé JW9 dont le nom chimique 4-({méthyl[(1R)-1-(naphthalen-1-yl)éthyl]amino}méthyl)phénol avec un score de docking égale à  $8.29 \text{ M}^{-1}$ , les deux se présentent comme les inhibiteurs les plus puissants de la PLpro. Pour cela nous avons choisis ces composés comme modèle afin d'interpréter leurs différentes interactions établies entre l'enzyme et les deux inhibiteurs.

## II.4. Analyse visuelle des interactions « 7JN2-inhibiteur »

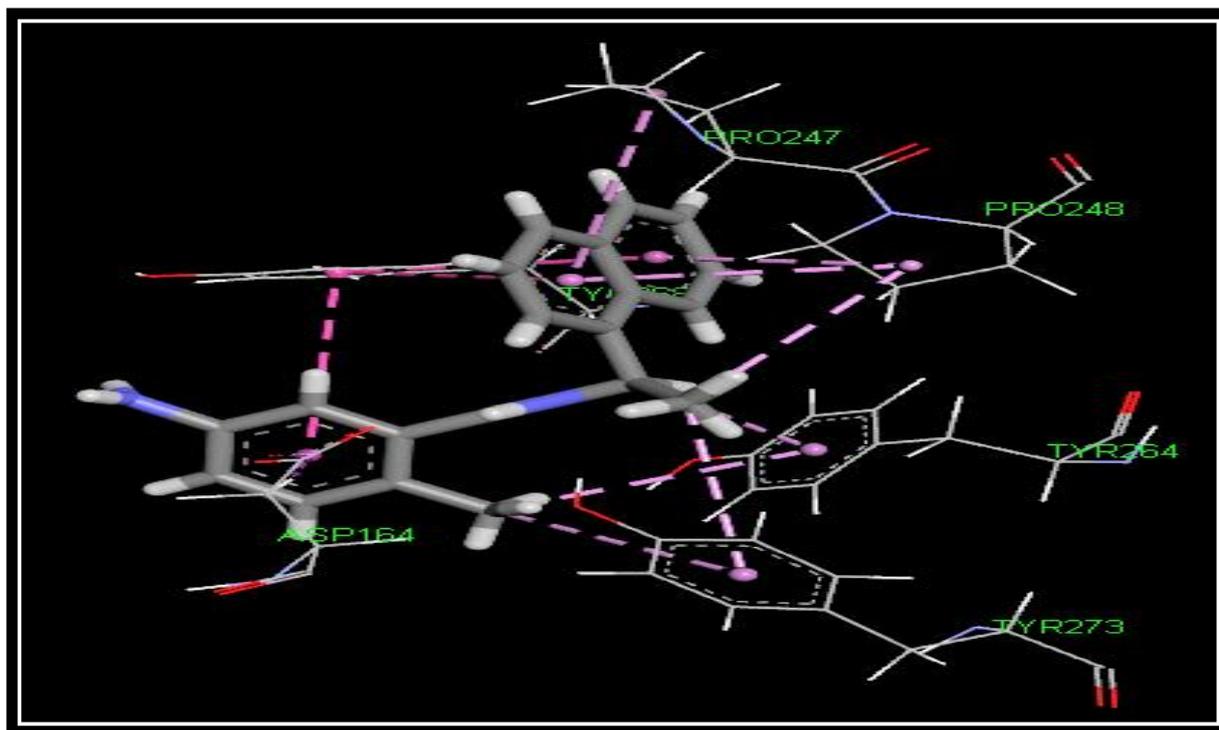
### II.4.1. Complexe 7JN2-TTT

L'analyse visuelle montre que l'inhibiteur TTT est bien placé dans le site actif de l'enzyme PLpro. Nous avons fait appel à l'interface graphique du programme Discovery Studio Visualizer pour visualiser et analyser le nombre et le type de liaison formant le complexe.



**Figure V. 4.** Analyse visuelle par discovery studio visualizer du mode d'interaction du ligand TTT avec le site actif de la 7JN2.

L'inhibiteur TTT établit plusieurs interactions hydrophobiques avec les résidus de site actif de notre cible enzymatique. Elles sont réalisées avec les acides aminés : PRO247\_PRO248 \_ASP164\_TYR273\_TYR264\_TYR268 (voir figure V.5).

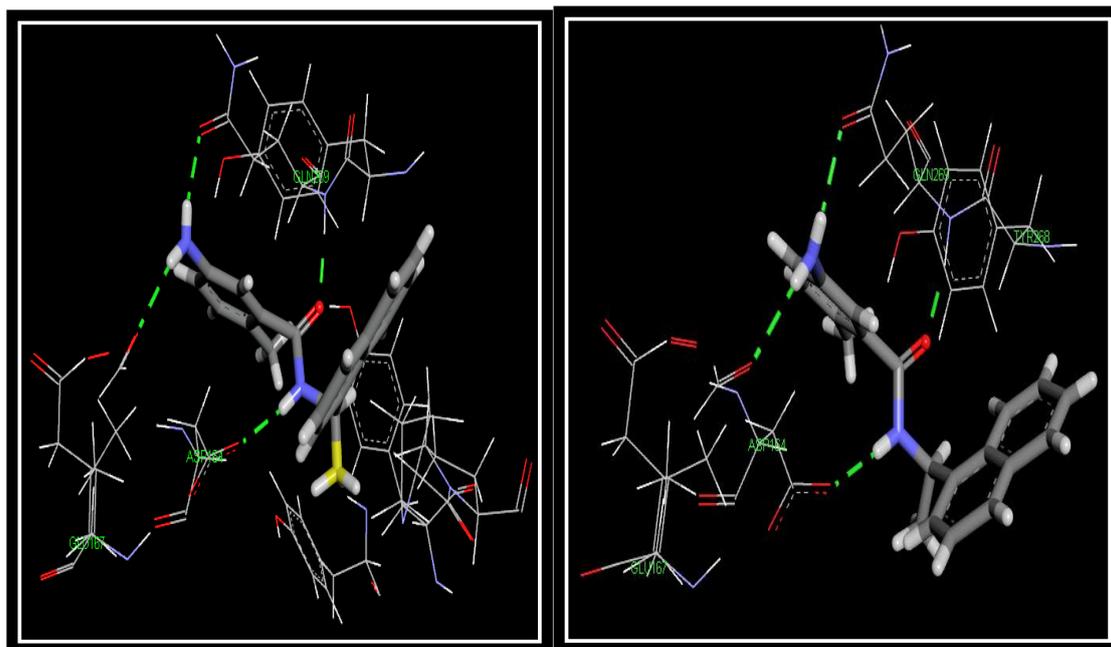


**Figure V.5.** Représentation des interactions hydrophobiques formées par le composé TTT et le site actif de la PLpro.

De plus, l'analyse visuelle du mode de positionnement de l'inhibiteur TTT dans le site actif de la PLpro montre la mise en place de plusieurs liaisons hydrogène. En particulier, les fonctions amine de ce composé, ils ont engagé dans quatre liaisons hydrogène :

- Trois ponts hydrogènes sont observés entre la fonction carbonyle des résidus Asp164, Glu269 et Glu167 et le groupement amine de l'inhibiteur TTT : (N-H.....O<sub>D2</sub>-Asp164), (N-H.....O<sub>E1</sub>-Gln269) et (N-H.....O<sub>E2</sub>-Glu167) avec des distances respectives de 2.03 Å, 2.19 Å et 2.26 Å;
- Un pont hydrogène est observé entre le carbonyle de TTT et groupement NH du résidu Gln269 (O.....H-N-Gln269 ; séparées par une distance de 1.86 Å).

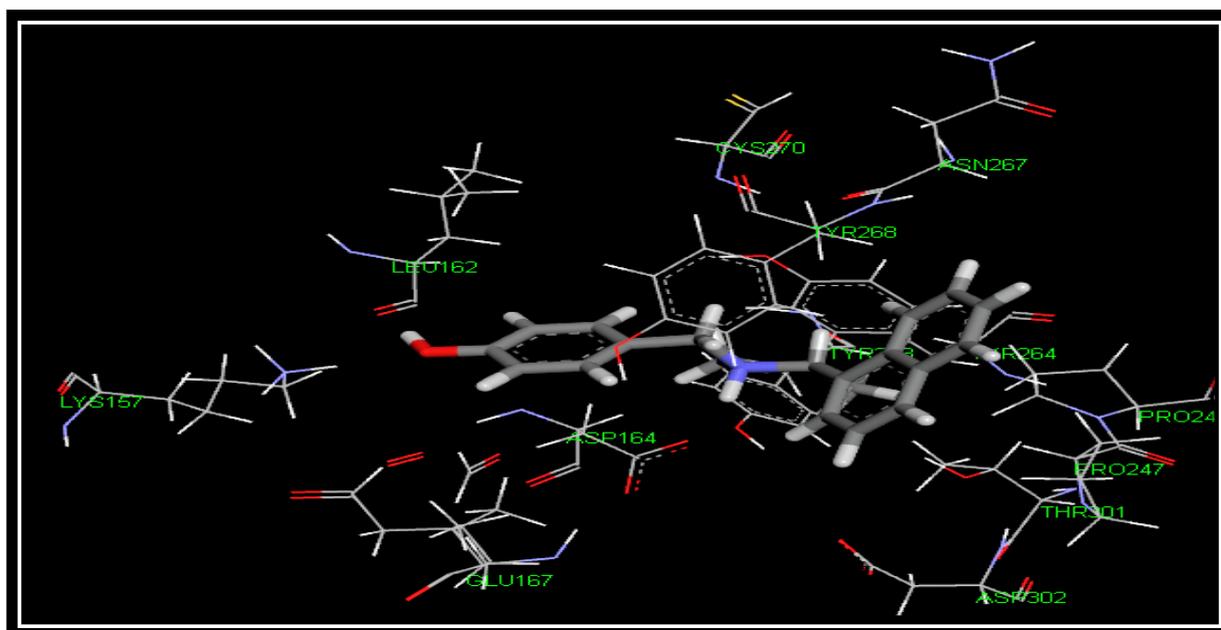
La figure V.6 ci-dessous montre le site actif de 7JN2 complexé avec l'inhibiteur TTT. Les acides aminés du site actif de notre enzyme sont représentés en « Line », l'inhibiteur est représenté en « Stick » et les distances entre les atomes susceptibles d'établir des liaisons hydrogènes sont représentées en lignes pointillées.



**Figure V.6.** Mode de liaison de l'inhibiteur TTT avec le site actif de la PLpro.

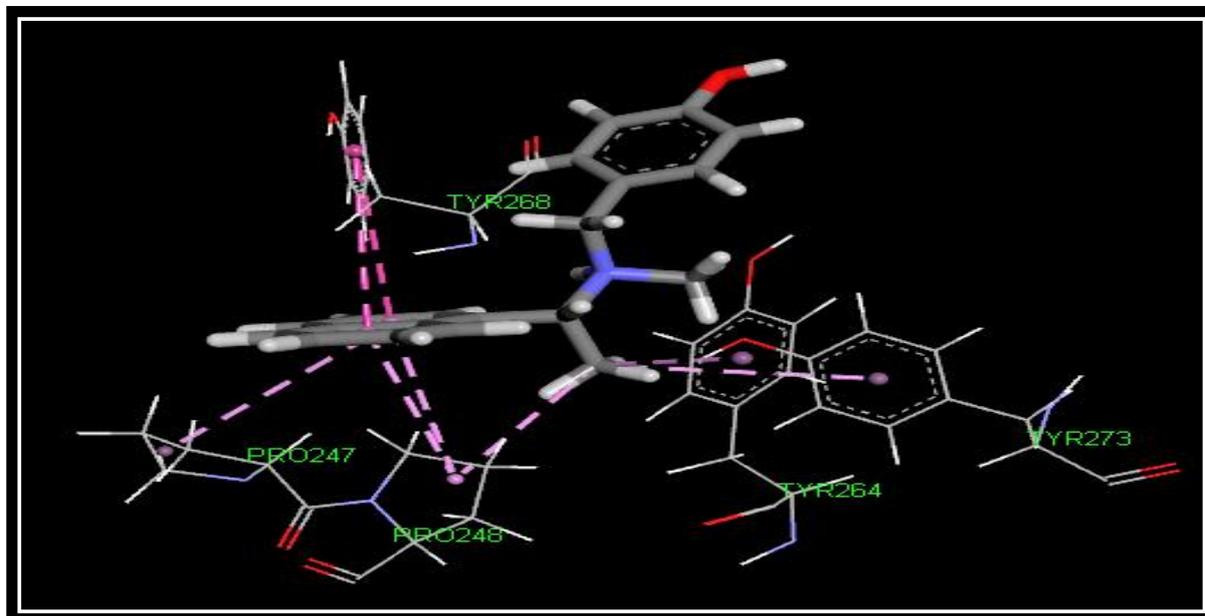
#### II.4.2. Complexe 7JN2-JW9

La visualisation des résultats du docking moléculaire montre que l'inhibiteur JW9 forme avec le site actif de la PLpro plusieurs interactions. De plus, cet inhibiteur est bien centré dans le site actif de notre cible enzymatique. Nous avons utilisé le programme discovery studio visualizer pour visualiser le nombre et le type de liaison former.



**Figure V.7.** Analyse visuelle par discovery studio visualizer de mode d'interaction du complexe 7JN2-JW9.

Le complexe 7JN2-JW9 est stabilisé par plusieurs interactions hydrophobiques (figure V.8). Elles sont réalisées avec les résidus d'acides aminés suivants : TYR268\_PRO247\_PRO248\_TYR264\_TYR273.

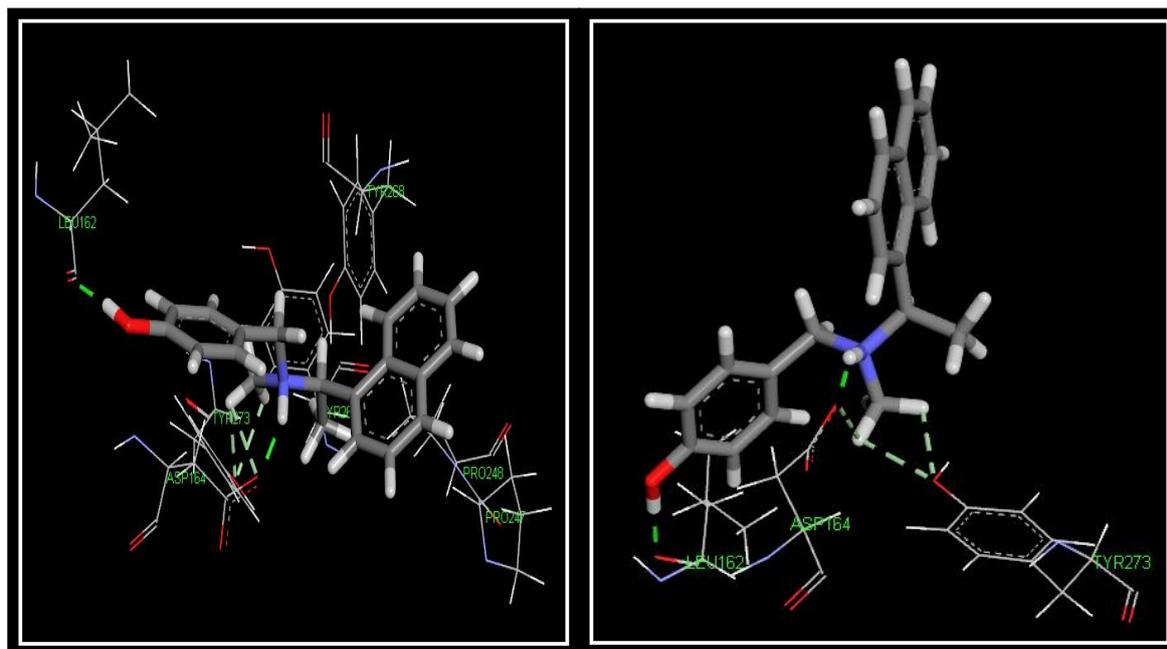


**Figure V.8.** Représentation des interactions hydrophobiques formées par le composé JW9.

De plus, ce complexe est stabilisé par la formation de cinq liaisons hydrogène.

- Liaison hydrogène entre la fonction hydroxyle de l'inhibiteur JW9 et l'oxygène du résidu Leu 162 (O-H.....O-Leu162 ; distance = 1.90Å) ;
- Le carboxyle de résidu Asp164 est établi une liaison hydrogène avec le carbone de l'inhibiteur JW9 (C-H.....O<sub>D2</sub>-Asp164 ; distance = 1.90Å) ;
- L'hydroxyle du résidu Tyr 273 est établi deux liaisons hydrogène avec le carbone de l'inhibiteur JW9 (C-H.....H-O-Tyr273; distance =2.51 Å) et(C-H.....H-O-Tyr273; distance =2.55 Å) ;
- Le groupement amine du composé JW9 forme un pont hydrogène avec la fonction carboxyle du résidu Asp164 (N-H.....O<sub>D2</sub>-Asp164 ; distance= 1.84Å).

Les liaisons de type hydrogène observées lors de l'interaction de cet inhibiteur avec le site actif de l'enzyme PLpro sont représentées dans la figure V. 9 ci-dessous.



**Figure V.9.** Mode de liaison de l'inhibiteur JW9 avec le site actif de la 7JN2.

### III. Proposition de nouveaux inhibiteurs de la PLpro

#### III.1. Criblage virtuel basé sur la structure de l'enzyme PLpro

Le CV est la stratégie *in silico* la plus utilisée pour l'identification de touches « hits » dans le cadre de la recherche de nouveaux médicaments. Le CV est une approche visant à simuler et à prédire l'affinité d'un très grand nombre de ligands pour le site actif d'une cible thérapeutique donnée, ce qui est assez facile à mettre en place, plus rapide, pour un coût bien moindre que les criblages expérimentaux [70 , 4].

Afin d'améliorer l'activité inhibitrice des composés TTT et JW9 nous avons utilisé la méthode du criblage virtuel basé sur la structure de la cible par docking moléculaire d'une collection des petites molécules chimiques.

Dans le premier criblage, nous avons réalisé un docking moléculaire par le logiciel Surflex-dock d'une collection de 394 similaire de l'inhibiteur TTT tirés de la chimiothèque PubChem et présentant une homologie structurale de 92 % à notre ligand de référence. Alor que, dans le second, nous avons réalisé un docking d'une collection de 161 similaires de l'inhibiteur JW9 avec 90 % de similarité. Ces composés chimiques sont téléchargés à partir de la banque de donnée PubChem. Finalement, nous avons comparé le score de docking de ces inhibiteurs (Affinité) à celui des inhibiteurs de référence (TTT et JW9) et proposé les meilleurs

vis-à-vis de notre cible enzymatique PLpro. Les résultats de simulation par Surflex-dock sont représentés dans le tableau V. 4 et tableau V.5 suivants.

**Tableau V.4.** Résultats de docking des similaires de composé TTT.

Code de ligand	Affinité (M <sup>-1</sup> )	Code de ligand	Affinité (M <sup>-1</sup> )	Code de ligand	Affinité (M <sup>-1</sup> )
<b>CID 24941262</b>	8.25	CID119758397	5.23	CID120418047	7.07
<b>CID101021718</b>	5.44	CID119787291	2.61	CID120418131	5.05
<b>CID104291714</b>	5.65	CID120417100	4.99	CID120418160	7.18
<b>CID104291750</b>	5.85	CID120417155	3.43	CID120418259	4.71
<b>CID104291775</b>	4.82	CID 120417174	5.11	CID120418455	6.21
<b>CID120418501</b>	5.96	CID 120417755	6.09	CID120418484	4.79
<b>CID119691301</b>	5.62	CID 120417772	4.93	CID 120418499	5.37
<b>CID8695502</b>	6.00	CID116649896	3.80	CID104291803	5.35
<b>CID120418636</b>	3.12	CID119547539	5.67	CID104292227	5.30
<b>CID120418833</b>	5.11	CID119691301	5.62	CID104292228	4.91
<b>CID120419253</b>	4.84	CID119691322	6.96	CID104292397	4.99
<b>CID120419347</b>	2.27	CID119727368	5.13	CID104292489	5.57
<b>CID120419597</b>	5.46	CID119748605	2.62	CID108541841	6.17
<b>CID120419807</b>	4.94	CID120618233	5.12	CID108553842	5.28
<b>CID120419925</b>	5.69	CID120618471	6.27	CID108554041	5.25
<b>CID120618200</b>	4.55	CID120621457	4.86	CID110440179	6.51
<b>CID120621473</b>	5.75	CID120621925	6.21	CID112762297	4.54
<b>CID120621933</b>	4.25	CID120630705	6.15	CID112767372	5.32
<b>CID120622860</b>	4.79	CID120631619	8.03	CID112811631	1.59
<b>CID120626798</b>	4.27	CID120632404	1.99	CID116649768	3.64
<b>CID120629868</b>	6.46	CID 120633281	2.27	CID116649872	3.82
<b>CID120630671</b>	7.49	CID 120633420	4.52	CID116649873	6.42
<b>CID 18069948</b>	6.46	CID 120633742	3.82	CID 120635958	4.66
<b>CID 18154080</b>	4.83	CID 120633878	5.31	CID 120636561	6.17
<b>CID 18292638</b>	3.88	CID 120634471	4.66	CID 120641216	6.41
<b>CID 20112464</b>	7.04	CID 120635439	4.70	CID 120641883	4.66
<b>CID 20121477</b>	5.80	CID 120635936	5.72	CID 120645188	2.03
<b>CID 24218958</b>	7.23	CID 29861785	2.25	CID 120645251	6.55

<b>CID 24411849</b>	7.52	<b>CID 30380379</b>	5.65	<b>CID 120645295</b>	5.18
<b>CID 24543517</b>	4.31	<b>CID 34636685</b>	6.96	<b>CID 121683954</b>	2.80
<b>CID 24636372</b>	9.40	<b>CID 38584641</b>	4.42	<b>CID 121683958</b>	4.65
<b>CID 25685722</b>	3.36	<b>CID 38755668</b>	7.95	<b>CID 121683972</b>	5.47
<b>CID 25685723</b>	4.45	<b>CID 40278262</b>	6.90	<b>CID 121825260</b>	5.37
<b>CID 25862174</b>	5.01	<b>CID 42906318</b>	3.42	<b>CID 124425003</b>	7.33
<b>CID 25862806</b>	5.97	<b>CID 43010534</b>	4.02	<b>CID 133733346</b>	6.48
<b>CID 26157522</b>	7.52	<b>CID43297530</b>	6.88	<b>CID 133827906</b>	8.31
<b>CID26160146</b>	6.41	<b>CID43297686</b>	8.40	<b>CID 133918959</b>	7.15
<b>CID27447549</b>	6.57	<b>CID43533069</b>	5.49	<b>CID 134003616</b>	6.24
<b>CID27447712</b>	7.15	<b>CID43533074</b>	5.61	<b>CID 144514255</b>	7.32
<b>CID28638832</b>	9.04	<b>CID43652012</b>	4.90	<b>CID 154727873</b>	4.48
<b>CID 28650458</b>	5.49	<b>CID 43652026</b>	5.33	<b>CID 154728196</b>	8.45
<b>CID 29861784</b>	3.87	<b>CID 43824499</b>	6.06	<b>CID 17456336</b>	5.84
<b>CID 43877777</b>	8.20	<b>CID 61207800</b>	4.77	<b>CID 17499980</b>	4.18
<b>CID 43887531</b>	8.17	<b>CID 61209358</b>	3.57	<b>CID 51184161</b>	4.34
<b>CID 43956033</b>	6.23	<b>CID 61289850</b>	4.62	<b>CID 51313962</b>	4.99
<b>CID 43956038</b>	5.84	<b>CID 61290224</b>	5.92	<b>CID 51339038</b>	5.04
<b>CID 45744583</b>	6.39	<b>CID 61290602</b>	6.89	<b>CID 53794251</b>	4.28
<b>CID 45744596</b>	5.16	<b>CID 61290793</b>	4.25	<b>CID 54833385</b>	7.75
<b>CID45744620</b>	6.82	<b>CID 61290973</b>	6.46	<b>CID 54835077</b>	8.93
<b>CID 45747438</b>	7.09	<b>CID61291555</b>	6.75	<b>CID54839918</b>	7.71
<b>CID45751961</b>	9.11	<b>CID61292468</b>	3.91	<b>CID54839948</b>	7.56
<b>CID45752068</b>	9.29	<b>CID61852360</b>	6.52	<b>CID55498930</b>	3.00
<b>CID45752096</b>	5.93	<b>CID61996694</b>	7.22	<b>CID55648822</b>	6.11
<b>CID45752116</b>	9.30	<b>CID62055528</b>	5.85	<b>CID55705757</b>	5.91
<b>CID45752117</b>	8.45	<b>CID62304077</b>	5.16	<b>CID55749195</b>	3.86
<b>CID46178562</b>	4.74	<b>CID62304078</b>	5.53	<b>CID60389686</b>	7.17
<b>CID46402392</b>	7.13	<b>CID62304243</b>	5.73	<b>CID60423818</b>	6.11
<b>CID46474318</b>	7.58	<b>CID 62304244</b>	5.58	<b>CID 60490794</b>	6.32
<b>CID 46434041</b>	7.65	<b>CID 62304579</b>	4.92	<b>CID 60535515</b>	6.53
<b>CID 46464519</b>	6.36	<b>CID 62304912</b>	5.50	<b>CID 60890874</b>	5.68
<b>CID 46465470</b>	6.96	<b>CID 62305086</b>	5.00	<b>CID 60891627</b>	6.35
<b>CID 51165469</b>	4.82	<b>CID 62306559</b>	4.89	<b>CID 61206881</b>	5.56

CID 78666436	5.74	CID 62313166	5.35	CID 82546440	4.18
CID 78776662	6.56	CID 62367730	4.02	CID 85820623	2.29
CID 78851555	5.89	CID 62499610	7.27	CID 86991469	6.38
CID 64734198	4.93	CID 62499794	8.06	CID 93675355	4.80
CID 71882862	4.44	CID 63011914	5.64	CID 93675356	7.32
CID 72119165	5.45	CID 73957885	3.61	CID 95733136	6.39
CID 97836488	5.70	CID 73969896	8.00	CID 95781960	5.37
CID 61209332	5.19	CID 97836487	5.89	CID 95781961	6.67
CID120625508	2.56	CID120631765	6.76	CID112839060	6.68
CID9862540	4.95	CID 7496009	6.35	CID143211681	2.43
CID9893483	7.33	CID7496034	4.55	CID143579827	6.98
CID24941262	8.25	CID7496035	4.78	CID146636717	6.12
CID44235655	4.71	CID7496036	4.38	CID146681430	8.91
CID 44235889	8.60	CID7496037	3.21	CID147545055	4.08
CID44235890	3.37	CID8326348	6.57	CID149879688	4.40
CID44236116	4.33	CID8326350	5.37	<b>CID149925965</b>	<b>9.52</b>
CID44828571	6.62	CID9882927	4.87	CID150358185	4.29
CID44828572	7.30	CID10214731	6.06	CID152343466	4.91
CID74369226	6.05	CID12583854	4.67	CID152379198	3.43
CID10914143	7.05	CID12583855	5.76	CID153583723	3.85
CID44236359	8.16	CID14678740	2.54	CID153835436	6.96
CID44828570	5.54	CID14839227	2.55	CID154815733	7.94
CID62367730	4.02	CID15542859	4.43	CID140524361	2.32
CID141096686	9.19	CID15951246	3.45	CID140524356	3.68
CID141096688	8.89	CID16889433	7.76	CID140524347	2.49
CID141991850	2.73	CID18175072	4.19	CID53794251	4.28
CID142214185	5.40	CID1831864	5.36	CID67105822	8.26
CID142214262	4.10	CID20115144	6.72	CID67599568	1.88
CID142988960	2.27	CID20223580	5.05	CID67600497	3.57
CID57209236	6.79	CID54182076	6.39	CID20223581	5.08
CID58497026	5.36	CID54196403	4.77	CID21917043	4.84
CID59534001	3.05	CID56663963	5.06	CID44208401	6.61
CID59627837	4.53	CID56675046	6.40	CID53973213	3.94
CID68562262	6.03	CID8695511	7.20	CID17499980	4.18

<b>CID68562418</b>	6.52	CID8695512	5.67	CID18069948	6.46
<b>CID68777337</b>	8.35	CID9412891	6.76	CID18154080	4.83
<b>CID68778083</b>	6.04	CID12583853	5.79	CID18292638	3.88
<b>CID68779807</b>	8.53	CID14872007	4.27	CID20112464	7.04
<b>CID68781200</b>	8.19	CID14872042	5.73	CID20121477	5.80
<b>CID68781927</b>	5.68	CID14872051	4.96	CID24218958	7.23
<b>CID68783352</b>	7.32	CID16105852	6.53	CID24411849	7.52
<b>CID68909104</b>	6.29	CID16567514	7.60	CID24543517	4.13
<b>CID68972497</b>	3.96	CID16967519	7.24	CID123465082	1.58
<b>CID70501501</b>	6.19	CID16781016	5.65	CID130372862	7.22
<b>CID70579833</b>	3.19	CID16785457	5.30	CID132227819	5.03
<b>CID70635022</b>	4.53	CID16889438	4.68	CID140524336	2.35
<b>CID70953340</b>	6.69	CID16889863	1.86	CID577677	6.57
<b>CID71441040</b>	5.19	CID16889867	6.27	CID959307	5.94
<b>CID72343105</b>	4.42	CID17057986	5.45	CID2127255	5.11
<b>CID74952323</b>	6.65	CID17170548	6.75	CID2236625	7.78
<b>CID74960507</b>	5.41	CID17194553	5.73	CID2236626	8.54
<b>CID77246650</b>	7.30	CID17195189	7.85	CID2980106	8.55
<b>CID90819712</b>	6.33	CID17196938	7.77	CID3050142	4.77
<b>CID91141962</b>	7.84	CID17198826	8.31	CID3426439	6.82
<b>CID91192123</b>	4.04	CID17201250	7.30	CID7496006	5.98
<b>CID91207411</b>	7.04	CID17218763	7.13	CID7496007	3.91
<b>CID91329062</b>	6.91	CID17222655	6.08	CID7496008	6.46
<b>CID91333324</b>	1.27	CID17230690	6.29	CID59627841	4.59
<b>CID91447187</b>	7.67	CID17231811	4.61	CID59627843	6.25
<b>CID91592481</b>	7.28	CID17234314	7.23	CID66908862	4.51
<b>CID118908271</b>	3.40	CID17392275	6.50	CID66908941	3.14
<b>CID118980727</b>	1.84	CID17445258	7.21	CID66991606	9.15
<b>CID123437211</b>	4.92	CID17456336	5.84	CID67105350	7.66
<b>CID67600598</b>	3.75	CID8695504	7.03	CID67105599	8.64
<b>CID68117836</b>	2.62	CID8695509	6.19	CID8695489	4.94
<b>CID68251952</b>	3.47	CID8695510	6.79	CID8695490	4.21
<b>CID68437409</b>	7.39	CID8695503	7.47	CID8695491	5.12
<b>CID68558714</b>	6.59	CID68562045	5.31	CID8695492	5.95

CID68560053	5.09	CID68562050	6.54	CID8695501	7.89
CID68562260	6.29				

Tableau V.5. Résultats de docking des similaires de composé JW9.

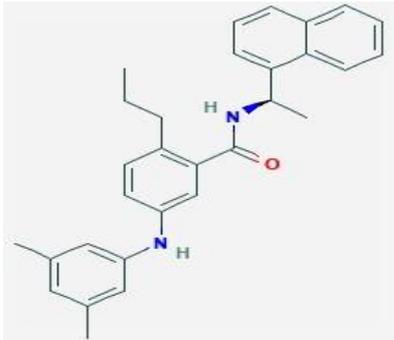
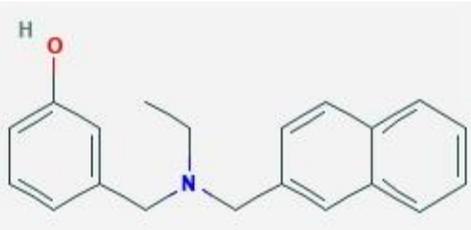
Code de ligand	Affinité (M <sup>-1</sup> )	Code de ligand	Affinité (M <sup>-1</sup> )	Code de ligand	Affinité (M <sup>-1</sup> )
CID101218928	-1.15	CID12031089	1.81	CID45276680	1.61
CID101218929	0.34	CID12035089	4.03	CID45276681	2.50
CID101218931	-0.51	CID12035090	2.86	CID45276682	1.15
CID101490397	3.08	CID12133041	3.18	CID45276839	2.28
CID101486814	0.38	CID1241349	1.44	CID45278355	-0.06
CID101218926	-1.16	CID1241350	4.72	CID45479132	6.43
CID101734007	2.97	CID131909802	3.76	CID46196604	-0.92
CID101734008	0.97	<b>CID142139076</b>	<b>8.20</b>	CID50959857	3.41
CID101734009	1.86	CID143084704	3.55	CID53235452	1.97
CID101734010	-0.29	CID15309628	6.29	CID53242721	0.35
CID101734011	1.25	CID15519385	3.18	CID53244140	1.48
CID101734013	1.44	CID17785784	3.53	CID53244141	-2.01
CID101734014	0.30	CID18125089	7.49	CID53769996	6.37
CID101842298	-1.67	CID18271171	2.42	CID56707345	4.49
CID10187274	-2.57	CID19881608	1.53	CID60284834	7.19
CID102351146	1.95	CID20780880	-3.34	CID60284870	4.03
CID102493799	5.46	CID21230994	2.13	CID60285177	5.42
CID11077538	4.46	CID21826591	1.29	CID62369091	6.76
CID11270037	-0.26	CID23500862	0.23	CID70680674	2.20
CID11312367	3.58	CID23541217	-3.22	CID73893173	5.58
CID11335695	3.22	CID23622267	3.90	CID75526221	5.43
CID11360175	0.93	CID24077722	2.18	CID79001474	2.64
CID11428353	1.14	CID2845587	3.58	CID80094923	6.78
CID11430012	-0.39	CID28665942	7.60	CID80094925	4.45
CID115280542	3.34	CID28665946	5.32	CID8221228	3.81
CID115280545	2.68	CID3029535	2.63	CID82957233	2.64
CID115280641	1.65	CID39963651	5.97	CID82957541	3.68
CID115280812	3.27	CID39963653	6.65	CID82977116	4.66

<b>CID115281972</b>	4.45	CID43224316	8.11	CID82977332	4.47
<b>CID115572166</b>	6.51	CID44536475	3.93	CID82977613	2.94
<b>CID95205597</b>	3.50	CID95205596	4.49	CID92136292	1.97
<b>CID97783595</b>	5.43	CID11199642	-0.17	CID92168518	1.70
<b>CID97783596</b>	2.41	CID22415918	5.54	CID22415918	5.54
<b>CID101218932</b>	-2.29	CID11809720	0.30	CID45276679	0.27
<b>CID60574</b>	0.46	CID129056708	7.44	CID58666455	-2.47
<b>CID234822</b>	4.80	CID129056718	5.36	CID59178738	2.50
<b>CID577695</b>	5.38	CID129056722	5.81	CID31236537	6.14
<b>CID11383025</b>	0.49	CID139076690	1.51	CID67427834	4.73
<b>CID11771125</b>	1.92	CID139084449	2.43	CID67926225	6.44
<b>CID12912982</b>	6.16	CID139758150	7.43	CID68818084	0.34
<b>CID12912986</b>	5.24	CID141526926	4.35	CID70404862	1.87
<b>CID53216175</b>	2.06	CID142099028	4.89	CID70405464	1.06
<b>CID53739606</b>	4.17	CID142814354	2.74	CID704055227	-1.31
<b>CID129876386</b>	2.58	CID142924736	1.50	CID70435245	1.22
<b>CID257443</b>	5.83	CID54040501	3.70	CID72825567	-0.25
<b>CID11011300</b>	1.43	CID57453710	-3.01	CID72957717	3.15
<b>CID11033165</b>	3.86	CID85432648	-0.22	CID127028826	2.92
<b>CID11223823</b>	-2.91	CID85592519	5.07	CID127028828	1.39
<b>CID18505277</b>	6.79	CID89189350	1.76	CID12702829	2.32
<b>CID143242758</b>	0.24	CID95117936	3.79	CID129056702	8.04
<b>CID145300129</b>	6.99	CID95117937	2.06	CID129056705	5.77
<b>CID146506292</b>	2.93	CID85946536	2.43	CID90849040	1.51
<b>CID151106148</b>	4.62	CID88936026	-3.27	CID101218925	-0.25
<b>CID153887472</b>	3.76	CID89189350	1.76		

### III.2. Molécules sélectionnées comme nouveaux inhibiteurs

Le criblage réalisé fait ressortir deux composés présentés dans le tableau V.6 suivant et proposé comme nouveaux inhibiteurs théoriquement puissants envers notre cible enzymatique PLpro.

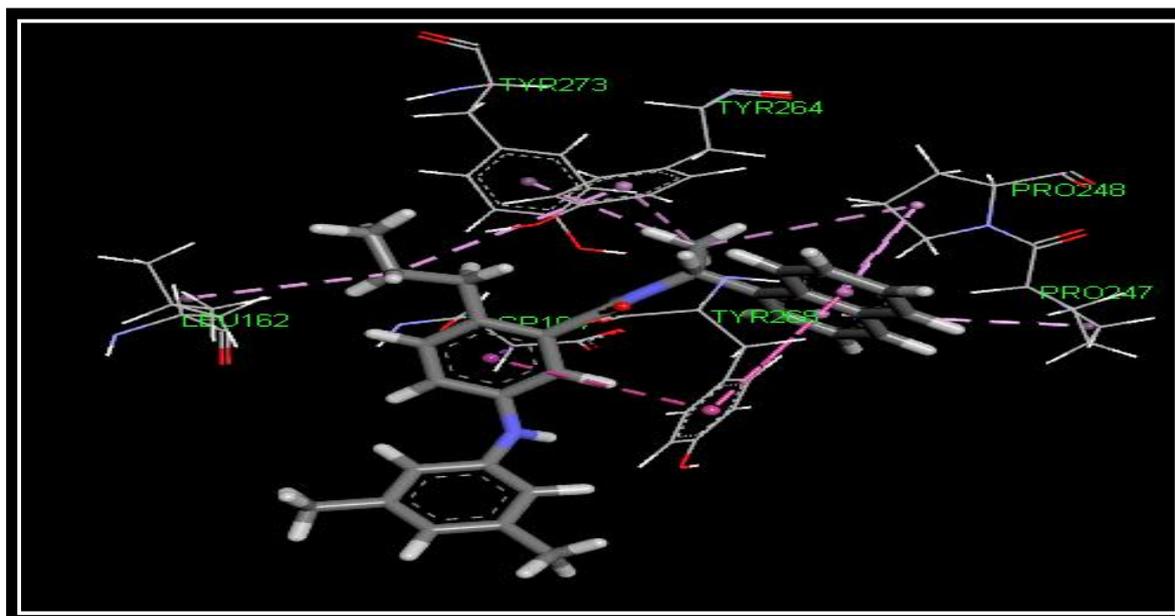
**Tableau V.6.** Données chimiques et structurales des composés sélectionnés comme candidats.

Code de l'inhibiteur	Structure	Nom chimique	Affinité (M <sup>-1</sup> )
CID 149925965		: 5-(3,5-dimethylanilino)-N-[(1R)-1-naphthalen-1-ylethyl]-2-propylbenzamide	9.52
CID 142139076		3-[[Ethyl(naphthalen-2-ylmethyl)amino]methyl]phenol	8.20

### III.3. Prédiction du mode d'interaction des meilleurs inhibiteurs

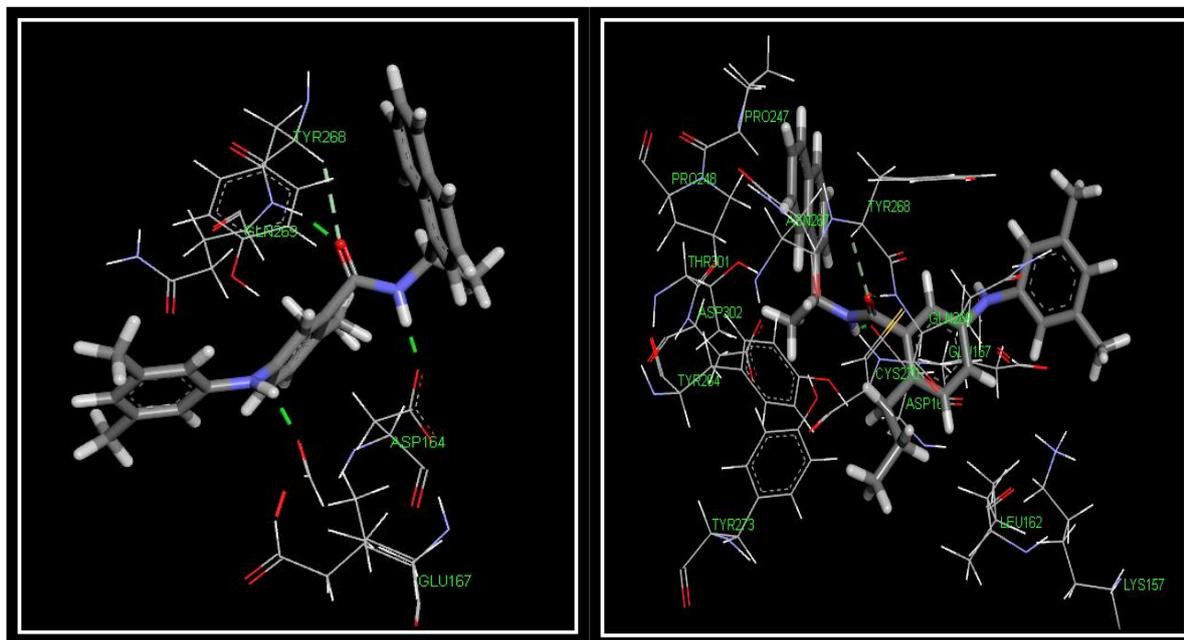
#### III.3.1. Interaction de complexe 7JN2- CID 149925965

La visualisation des résultats de criblage par docking moléculaire montre que l'inhibiteur CID149925965 forme avec le site actif de la 7JN2 plusieurs interactions hydrophobiques. Elles sont réalisées avec les résidus des acides aminés suivants : LEU162, TYR273, TYR264, PRO248, PRO247, ASP164 et TYR268. Ces interactions sont représentées dans la figure V.10 ci-dessous.



**Figure V.10.** Représentation des interactions hydrophobiques formées par l'inhibiteur CID149925965 avec la 7JN2.

De plus, le complexe 7JN2-CID149925965 est stabilisé par la formation de quatre liaisons hydrogène. La figure V.11 montre le site actif de la 7JN2 complexée avec cet inhibiteur.



**Figure V.11.** Représentation des interactions hydrogène formées par le similaire CID149925965 dans le site actif de 7JN2.

- Deux liaisons hydrogène sont réalisées entre le groupement amine de ce similaire et la fonction carbonyle des résidus Asp164 et Glu267 ;
- Un pont hydrogène est observé entre le carbonyle de l'inhibiteur et le groupement NH du résidu Gln269 ;
- Un pont hydrogène est formé entre le carbonyle de similaire CID149925965 et le carbone du résidu Tyr268.

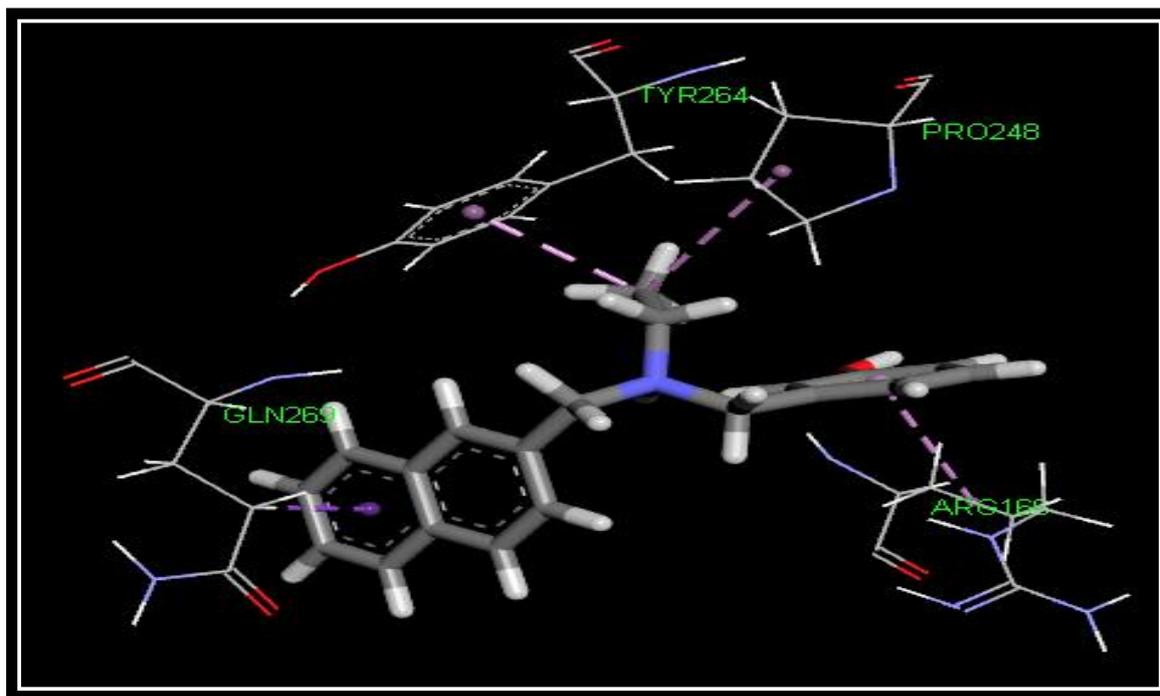
Dans le tableau V.7 ci-dessous nous avons résumé ces différents ponts.

**Tableau V.7.** Les quatre ponts hydrogène formés par l'inhibiteur CID149925965.

<b>Liaisons Hydrogène</b>	<b>Atomes du ligand CID149925965</b>	<b>Atomes et résidus du site actif</b>	<b>Distance (Å)</b>
<b>1</b>	N-H	O <sub>D2</sub> -Asp164	1.92
<b>2</b>	N-H	O <sub>E2</sub> -Glu267	1.81
<b>3</b>	O	H-N-Gln269	1.64
<b>4</b>	O	H-C <sub>A</sub> -Tyr268	2.68

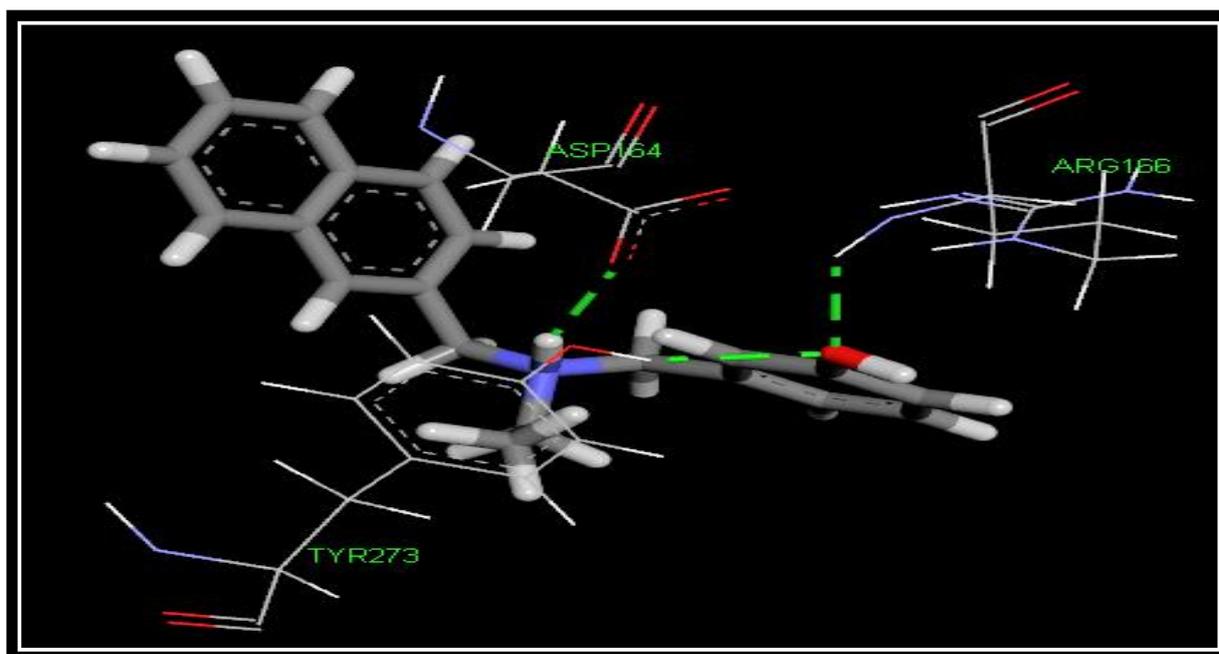
### III.3.2. Interaction de complexe 7JN2-CID 142139076

L'analyse visuelle montre que le similaire CID142139076 forme avec le site actif de la PLpro plusieurs interactions hydrophobiques avec les acides aminés ARG166, PRO248, TYR264 et GLN269. Ces interactions sont représentées dans la figure V.12 ci-dessous.



**Figure V.12.** Représentation des interactions hydrophobiques formées par le composé CID142139076 avec la 7JN2.

Dans le cas de l'inhibiteur CID142139076, le complexe enzyme-inhibiteur est stabilisé par la formation de trois liaisons hydrogène. Les liaisons de type hydrogène observées lors de l'interaction de cet inhibiteur avec le site actif de notre cible sont représentées dans la figure V.13 ci-dessous.



**Figure V.13.** Mode de liaison de l'inhibiteur CID142139076 avec le site actif de la 7JN2.

- Un pont hydrogène est formé entre la fonction carboxyle du résidu Asp164 et le groupement amine de l'inhibiteur CID142139076 ;
- Un pont est formé entre la fonction carboxyle de l'inhibiteur et le groupement amine du résidu Arg166 ;
- Un pont hydrogène est établi entre l'hydroxyle du résidu Tyr273 et la fonction carboxyle de l'inhibiteur CID142139076.

Dans le tableau V.8 ci-dessous nous avons résumé ces différents ponts.

**Tableau V.8.** Les trois ponts hydrogène formés par l'inhibiteur CID142139076.

<b>Liaisons Hydrogène</b>	<b>Atomes du ligand CID142139076</b>	<b>Atomes et résidus du site actif</b>	<b>Distance (Å)</b>
<b>1</b>	N-H	O <sub>D2</sub> -Asp164	2.19
<b>2</b>	O	H-N-Arg166	2.22
<b>3</b>	O	H-O-Tyr273	2.27

## VI. Filtrage ADME/Tox de composés proposés

Nous avons appliqué la règle de 5 *Lipinski*. Cette règle décrit les propriétés physico-chimiques d'un composé testé [71, 63]. Donc, il est nécessaire d'évaluer les paramètres permettant leur validation comme antibiotiques (tableau V.9). Elle permet de calculer les propriétés moléculaires importantes des meilleurs composés prédits (logP, le nombre de donneurs et accepteurs de liaison hydrogène...etc.) ; directement sur la page internet de la chimiothèque PubChem.

**Tableau V.9.** Critères de la règle de *Lipinski* pour les différents inhibiteurs étudiés.

N	Composés	PM	nOH,NH	nO,N	ClogP	nrotb
1	TTT	304.4	4	2	2	3
2	JW9	291.4	4.5	1	2	4
3	CID149925965	436.6	8	2	2	7
4	CID142139076	291.4	4.5	1	2	5

**PM** : poids moléculaire (g /mol),

**nOH,NH** : nombre de donneurs de liaisons H ,

**nO,N** : nombre d'accepteurs de liaisons H ,

**clogP** :logP ou coefficient de partition calculé,

**nrotb** : nombre de liaisons rotables.

Les résultats du tableau V.9 montrent que les quatre composés étudiés s'inscrivent parfaitement dans la marge des critères imposés par la règle de Lipinski. On constate que, tous les inhibiteurs étudiés ont un poids moléculaire inférieurs à 500g/mol et un LogP inférieur à 5 ; ce qui indique un meilleur caractère lipophile. Donc, ces composés proposés sont capable de présenter une activité biologique sans avoir des problèmes d'absorption par voie oral. De ce fait, nous pouvons proposer ces composés comme étant des nouveaux inhibiteurs plus puissants de l'enzyme PLpro.

Cependant, d'autres critères sont pris en compte dans la sélection de molécules potentiellement candidates (tableau V.10).

Tableau V.10. Autres critères.

N	Composés	Nombre d'halogènes	Chaînes alkyles	Nombre de cycles	Nombre d'atome d'oxygène	Nombre d'atome d'azote
1	TTT	0	0	3	1	2
2	JW9	0	0	3	1	1
3	CID149925965	0	0	4	1	2
4	CID142139076	0	0	3	1	1

Les résultats du tableau V.10 montrent que les 4 molécules étudiées s'inscrivent parfaitement dans la marge de ces critères.

Les résultats d'affinités et structuraux concernant le docking des structures étudiées (TTT, JW9, CID149925965 et CID142139076) dans le site actif de la PLpro ont permis d'obtenir les informations souhaitées concernant le mode d'interaction spécifique de ces inhibiteurs.

# CONCLUSION ET PERSPECTIVE



### CONCLUSION ET PERSPECTIVE

Notre travail de master avait pour objectif d'acquérir des connaissances dans les simulations informatiques, notamment le criblage virtuel basé sur la structure de la cible par le logiciel Surflex-dock, afin de découvrir *in silico* de nouvelles molécules agissant comme inhibiteurs très puissants de la PL protéases (PLpro). C'est une cible thérapeutique potentielle intéressante pour développer de nouveaux traitements pour le SARS COV-2.

Au préalable, nous avons testé la fiabilité du programme Surflex-dock par deux tests, pour confirmer que ce logiciel est adapté à notre étude *in silico*.

- Le premier test consiste à calculer le RMSD : sur 150 complexes protéine-ligand pris arbitrairement de la PDB, 88% ont présenté un RMSD parfaitement dans les normes c'est-à-dire des valeurs inférieures ou égales à 2 Å contre 12% de valeurs supérieures à 2 Å.
- Le deuxième test est une analyse visuelle effectuée par le logiciel discovery studio visualizer. C'est une étape essentielle pour confirmer les résultats de premier test. Si les valeurs de RMSD étaient inférieures à 2 Å, nous avons observé une bonne superposition des deux conformations. Par contre, si les valeurs du RMSD étaient supérieures à 2 Å, la superposition devient mauvaise.

Ces tests nous fournissent la preuve que le programme Surflex-dock est efficace et peut être utilisé de manière fiable pour simuler les interactions PLpro-Inhibiteur, objet de notre recherche.

Dans la deuxième partie de ce travail nous avons pu élucider le mécanisme d'interaction entre la PLpro et des inhibiteurs pris de la PDB. Les résultats ont montré que les composés TTT et JW9 avaient le meilleur effet inhibiteur sur l'enzyme PLpro.

Dans la troisième partie, nous avons tenté de découvrir de nouveaux inhibiteurs plus affins et plus puissants de la PLpro. Nous avons réalisés un criblage virtuel de 555 molécules issues de la base de données PubChem. Cette étude met en évidence les composés suivants : CID149925965 et CID142139076 comme inhibiteurs potentiels pour notre cible.

Enfin, il est important de noter que les quatre molécules testées : TTT, JW9, CID149925965 et CID142139076 sont conformes aux critères imposés par la règle de *Lipinski*,

## CONCLUSION ET PERSPECTIVE

---

ils sont donc capable de présenter une activité biologique sans avoir des problèmes d'absorption par voie oral.

En termes de perspectives d'avenir, nous proposons à :

- Utiliser d'autres programmes de criblage virtuel, parmi les plus récents et les plus performants tel que FlexX, Gold et Glide, pour tester tous les inhibiteurs de la PLpro étudiés jusqu'à présent afin de proposer les meilleurs ;
- La recherche de molécules similaires non toxiques et plus actives possibles sur la cible PLpro ;
- Acheter les meilleurs inhibiteurs choisis à partir des fournisseurs ;
- Tester *in vitro et/ou in vivo* l'activité inhibitrice des inhibiteurs choisis sur l'enzyme PLpro et calculer l'activité biologique  $IC_{50}$  et le KI de ces inhibiteurs.

# BIBLIOGRAPHIE



## REFERENCES BIBLIOGRAPHIQUES

- [1] Sallard E., Halloy J., Casane D., Van Helden J., Decroly É. 2020. Retrouver les origines du SARS-CoV-2 dans les phylogénies de coronavirus. *Journal of médecine sciences*. **36** :783-796.
- [2] La maladie à coronavirus COVID-19. [En ligne].<https://www.academie-sciences.fr/fr/Rapports-ouvrages-avis-et-recommandations-de-l-Academie/covid-19-les-medicaments.html> .Consulté le [25/06/2022].
- [3] SARS-CoV-2- Protéase de type papaïne (PLPro)-Antigènes (Protéines et Peptides). [En ligne].<https://www.clinisciences.com/achat/cat-sars-cov-2protease-de-typepapaine-5137.html>. Consulté le [22/05/2022].
- [4] Mokrani E.H., Bensegueni A., Chikhi A., Teniou S., Merabti B., Boucherit H. 2012. Mise en évidence par criblage virtuel d'un nouvel inhibiteur xanthinique de la DPP4 pour le traitement du diabète de type 2. *Les technologies de laboratoire*. **7** : 41-48.
- [5] Hediye-Bag D. Définition d'un cas COVID ambulatoire par le biais d'une étude narrative et d'une étude typologique d'une cohorte de 1 500 patients ambulatoire. Thèse de Doctorat en Médecine. Université paris SACLAY. France. 2021 .133p .
- [6] Coronavirus et Covid-19 · Inserm, La science pour la santé. [En ligne]. <https://www.inserm.fr/dossier/coronavirus-sars-cov-et-mers-cov/>. Consulté le [07 /03 /2022].
- [7] Maladie Covid-19 (nouveau coronavirus) | Institut Pasteur. [En ligne]. <https://www.pasteur.fr/fr/centre-medical/fiches-maladies/maladie-covid-19-nouveau-coronavirus>. Consulté le [21/03/2022].
- [8] Maladie covid-19 : définition, durée, évolution, traitements. [En ligne]. <https://sante.journaldesfemmes.fr/fiches-maladies/2619933-maladie-covid-19-coronavirus-signification-definition-nom-incubation-traitement-depistage-duree-evolution/>. Consulté le [09/03/2022].
- [9] Covid : les pays riches "bloquent" les plans de vaccination des pays en développement. [En ligne] .<https://www.bbc.com/afrique/monde-56475128>. Consulté le [12/03/ 2022].

[10] Pillaiyar T., Wendt L.L., Manickam M., Easwaran M. 2021. The recent outbreaks of human coronaviruses: A medicinal chemistry perspective. *Medicinal research review*. **41** :72-135.

[11] Kamel S., Raynor A., Zozor S., Lacape G., Brunel V., Nivet V.A., Collin D.Ch., Peoc'h K., Cohen A., Ben Lassoued A., Chévrier M., Alemann M., Lessinger J.M., Bérard A.M., Sapin V., Beauvieux M.Ch., LEVY P. 2021. Atteintes myocardiques au cours de la maladie à coronavirus 19 (Covid-19) : principaux mécanismes physiopathologiques et utilité clinique des biomarqueurs cardiaques. *Annales de biologie clinique*. **79** : 213-4.

[12] Structure du virus COVID-19; Anatomie du SRAS-CoV-2; Modèle de virus pandémique; Syndrome respiratoire aigu sévère Diagramme Coronavirus 2; Schéma du virus Coronavirus Disease 2019; Organisation Novel Coronavirus. [En ligne]. <https://www.shutterstock.com/fr/image-illustration/covid19-virus-structure-sarscov2-anatomy-pandemic-1854731590>. Consulté le [10/03/2022].

[13] Protect Yourself. [En ligne]. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html>. Consulté le [12/03/2022].

[14] Drugs Available Under Emergency Use Authorization. [En ligne]. <https://uofmhealthwest.org/about/covid-19-updates/treatments>. Consulté le [12/03/2022].

[15] Fan L., Jiang S., Yang X., wang Z., Yang Ch. 2020. COVID-19 Drug Treatment in China. *Springer Nature Switzerland AG*. **6**:146-154.

[16] Marchal A. Anti-infectieux et SARS-CoV-2 : étude comparative des pratiques de prescription en médecine générale lors des deux vagues successives de COVID-19 en 2020 en Alsace. Thèse de Doctorat en Médecine Université de Strasbourg faculté de Médecine de Strasbourg, France. 2021. 107p.

[17] Quelle est la différence entre les tests de dépistage COVID-19 ? Comment savoir lequel privilégier ? [En ligne]. <https://www.normandie.ars.sante.fr/covid-19-les-differents-types-de-tests>. Consulté le [16/03/2022].

[18] Bahl sh., Javaid M., Bagha A.k., Singh R. p., Haleem A., Vaishya R., Suman R. 2020. Biosensors applications in fighting COVID-19 pandemic. *official publication of appollo hospitals group*. **17** :221-223.

- [19] Bonny V., Maillard A., Mousseaux C., Plaçais L., Richier Q. 2020. COVID-19 : physiopathologie d'une maladie à plusieurs visages. *Revue Médecine Interne*. **6**:375-89.
- [20] Steinkühler C. 2018. Viral Proteases. *Encyclopedia of Molecular Pharmacology*. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-38918-7\\_146](https://doi.org/10.1007/978-3-540-38918-7_146).
- [21] Turk D., Guncar G., Podobnik M., Turk B. 2009. Revised Definition of Substrate Binding Sites of Papain-Like Cysteine Proteases. *Journal of Walter de Gruyter*. **379**:137-147.
- [22] Hilgenfeld R., Kusov Y., Lei J. 2018. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Journal of Antiviral Research*. **149**: 58-74.
- [23] Osipiuk J., Azizi S.A., Dvorkin S., Endres M., Jedrzejczak R., Jones K.A., Kang S., Kathayat R. S., Kim Y.C.H., Lisnyak V.G., Maki S.L., Nicolaescu V., Taylor C.A., Tesar C.H., Zhang Y.A., Zhou Z., Randall G., Michalska K., Snyder S.A., Dickinson B.C., Joachimiak A. 2021. Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Journal of Nature Communications*. **12** (743):1-9.
- [24] Liang J. J., Pitsillou E., Ververis K., Guallar V., Hung A., Karagiannis T.C. 2021. Investigation of small molecule inhibitors of the SARS-CoV-2 papain-like protease by all-atom microsecond modelling, PELE Monte Carlo simulations, and in vitro activity inhibition. *Journal of Chemical Physics Letters*. **788**:1-9.
- [25] Razali R., Asis H., Budiman C. 2021. Structure-Function Characteristics of SARS-CoV-2 Proteases and Their Potential Inhibitors from Microbial Sources. *Journal of Microorganisms*. **9**(12):1-25.
- [26] SARS CoV 2 Structural proteins number of copies per virion have been estimated. [En ligne]. <https://twitter.com/ismaelmingarro/status/1246545114547314689>. Consulté le [21/05/2022].
- [27] Les inhibiteurs de la protéase du SRAS-CoV-2 sont-ils des antiviraux efficaces dans le traitement du COVID-19 ? [En ligne]. <https://ma-clinique.fr/les-inhibiteurs-de-la-protéase-du-sras-cov-2-sont-ils-des-antiviraux-efficaces-dans-le-traitement-du-covid-19>. Consulté le [19/05/2022].

[28] Hognon C., Marazzi M., García-Iriepa C.2022. Atomistic-Level Description of the Covalent Inhibition of SARS-CoV-2 Papain-like Protease. *International Journal of Molecular Science*.**23** (10):1-14.

[29] Delre P., Caporuscio F., Saviano M., Felice M., Angiatordi G. 2020. Repurposing Known Drugs as Covalent and Non-covalent Inhibitors of the SARS-CoV-2 Papain-Like Protease. *Journal of Molecular Studies of COVID-19 Chemistry*.**8**: 1-17.

[30] Gao X., Qin B., Chen P., Zhu K., Hou P., Wojdyla J.A., Wang M., Cui S.H.2021. Crystal structure of SARS-CoV-2 papain-like protease. *Journal of Acta Pharmaceutica Sinica B*.**11**: 237-245.

[31] Freitas B., Durie I., Murray J., Longo J., Miller H., Crich D., Jeff Hogan R., Tripp R., Pegan S.2022. Characterization and Non covalent Inhibition of the Deubiquitinase and deISGylase Activity of SARS-CoV-2 Papain-Like Protease. *ACS Infectious Diseases*.**6** (8):2099-2109.

[32] GRL0617  $\geq$  98% (HPLC). [Enligne]. [https://www.googleadservices.com/pagead/aclk?sa=L&ai=DChcSEwi65YznpZn4AhX9j2gJHUxMAPUYABAAGgJ3Zg&ae=2&ohost=www.google.com&cid=CAESauD2QzjyNn\\_PRg1YmF9iVUENBznEYbc\\_iYiHmiVtTKt3EwcwMaMENr\\_qwuwuT5r2CZuo\\_IOch0piIshIW6O6AIF98K-4y59M9jmNmzQBlo9tEF0udvRMYddYoYY5Qh2zJGKf8UyAtczWvw&sig=AOD64\\_2mjiA9YoYIegobTW21thWT2dv7Ew&q&adurl&ved=2ahUKEwibITnpZn4AhVbiv0HHRpMDPsQ0Qx6BAgDEAE](https://www.googleadservices.com/pagead/aclk?sa=L&ai=DChcSEwi65YznpZn4AhX9j2gJHUxMAPUYABAAGgJ3Zg&ae=2&ohost=www.google.com&cid=CAESauD2QzjyNn_PRg1YmF9iVUENBznEYbc_iYiHmiVtTKt3EwcwMaMENr_qwuwuT5r2CZuo_IOch0piIshIW6O6AIF98K-4y59M9jmNmzQBlo9tEF0udvRMYddYoYY5Qh2zJGKf8UyAtczWvw&sig=AOD64_2mjiA9YoYIegobTW21thWT2dv7Ew&q&adurl&ved=2ahUKEwibITnpZn4AhVbiv0HHRpMDPsQ0Qx6BAgDEAE). Consulté le [25/05/2022].

[33] Ma C., Wang J.2022. Validation and Invalidation of SARS-CoV-2 Papain-like Protease Inhibitors. *ACS Pharmacology & Translational Science*. **5**:102-109.

[34] Kumar S., Purohit D., Pandey P., Neeta A. 2017. Molecular docking and its application towards modern drug discovery. *Journal of pharmacy and pharmaceutical sciences*.**6**:691-696.

[35] Muegge I., Oloff S. 2007. Advances in virtual screening. *Journal of national library of medicine*.**3**:405-411.

[36] Louacheni F. 2013. Développement d'un portail web pour le criblage virtuel sur la grille de calcul. Institut de la Francophonie pour l'Informatique.

- [37] Rognan D. 2014. Méthodes de criblage in silico de chimiothèques : Méthodes de criblage virtuel. *L'expertise technique et scientifique de référence*. <https://doi.org/10.51257/a-v1-pha1020> .
- [38] Tripathi A., Misra K. 2017. Molecular Docking: A Structure Based Drug Designing Approach. *Journal of JSM Chemistry*. **5**:1-5.
- [39] Arrault A. Stratégies de docking-scoring assistées par analyse de données. Application au criblage virtuel des cibles thérapeutiques COX-2 et PPAR gamma. Thèse de Doctorat en Modélisation Moléculaire. Université d'Orléans. France. 2007. 191p.
- [40] Merzoug A. Recherche de nouvelles molécules bioactives par criblage virtuel et applications à l'inhibition de la peptide déformylase. Thèse de Doctorat en Biochimie. Université Mentouri Constantine1. Algérie. 2018. 193p.
- [41] Boucherit H. Recherche in silico de nouveaux composés bioactifs et applications à l'inhibition de la méthionine aminopeptidase. Thèse de Doctorat en Biochimie. Université Mentouri Constantine1. Algérie. 2020. 210p.
- [42] Ferreira L. G., dos Santos R. N., Oliva G and Andricopulo A. D. 2015. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules*. **20**: 13384-13421.
- [43] Hernández-Santoyo A., Tenorio-Barajas A.Y., Altuzar V., Vivanco-Cid H., Mendoza-Barrer C. 2013. Protein-Protein and Protein-Ligand Docking. *Journal of intechOpen*. 1-21. DOI: 10.5772/56376.
- [44] Bechelane Maia E.H., Assis L.C., De Oliveira T.A., da Silva A.M., Taranto A.G. 2020. Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Journal of frontiers in chemistry*. **8**: 343.
- [45] 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 en Bio-informatique. Paris : École Abbé Grégoire. France. 2014. 317p.
- [46] Liao C., Sitzmann M., Angelo Pugliese A., Nicklaus M. C. 2013. Software and resources for computational medicinal chemistry. *Future Medicinal Chemistry* : **3**(8): 1057–1085.

- [47] Chikhi A. calculs et modelisations des interactions peptide deformylase substances antibactériennes à l'aide de techniques de "docking" (arrimage) moléculaire .Thèse de Doctorat en Microbiologie. Université Mentouri Constantine 1. Algérie. 2007. 126p.
- [48] General Introduction Virtual Screening strategy. [En ligne], <https://u2roopasingh.wordpress.com/general-introduction/> . Consulté le [05/05/2022].
- [49] Beguiret O., Belgat N. Conception in silico de nouveaux inhibiteurs de la rénine par criblage virtuel. Mémoire de Magistère en Biologie. Université Mentouri Constantine 1. Algérie. 2016. 65p.
- [50] Alickovic L., Kellogg G. E. Accounting for side-chain flexibility in protein-ligand docking: 3D Interaction Homology as an approach of quantifying side-chain flexibility of Tyrosine. <https://slideplayer.com/slide/16185649/> . Consulté le [05/05/ 2022].
- [51] Zhang M., White R.A., Wang L., Goldman R., Kavraki L., Hassett B. 2005. Improving conformational searches by geometric screening .*Journal of Bioinformatics*. **21**: 624-630.
- [52] Sethi A., Joshi Kh., Sasikala K., Alvala M. 2019. Molecular Docking in Modern Drug Discovery: Principles and Recent Applications. *Drug Discovery and Development – New advances*. 1-22.
- [53] Morris G.M., Lim-Wilby M. 2008. Molecular Docking. *Methods Molecular Biology*. [https://doi.org/10.1007/978-1-59745-177-2\\_19](https://doi.org/10.1007/978-1-59745-177-2_19) .
- [54] Kramer B., Rarey M., Lengauer T. 1999. Evaluation of the FlexX Incremental Construction Algorithm for Protein–Ligand Docking. *Proteins: Structure, Function, and Genetics*. **37** : 228-241.
- [55] Boucherit H. 2012. Étude théorique des interactions intervenant dans l'inhibition de la méthionine aminopeptidase de *Mycobacterium tuberculosis* par diverses molécules. Mémoire de magistère en Biochimie. Université Mentouri Constantine 1. Algérie. 110p.
- [56] Hioual K. S. 2012. Etude comparative de l'efficacité de deux programmes de docking et application à l'inhibition de la neuraminidase. Mémoire de magistère en Biochimie. Université Mentouri Constantine 1. Algérie. 124p.

- [57] Centre national de la recherche scientifique. PubChem. [En ligne]. <https://pubchem.ncbi.nlm.nih.gov/>. Consulté le [06/06/2022].
- [58] Définition- PubChem. [En ligne].<https://dictionnaire.sensagent.leparisien.fr/PubChem/ru-ru/> . Consulté le [06/06/2022].
- [59] BioPharmics LLC. 2005. Surflex Manual: Docking and Similarity (version 1.3).
- [60] Ajay N. J.2003. Surflex: Fully Automatic Flexible Molecular Docking Using a Molecular Similarity-Based Search Engine>>. *UCSF Cancer Research Institute and Comprehensive Cancer Center*.**46**:499-511.
- [61] BioPharmics LLC. Computational Tools for Drug Discovery.[En ligne]. <https://www.biopharmics.com/>. Consulte le [06/06/2022].
- [62] Introduction to the Discovery Studio Visualizer.[En ligne]. [http://www.adrianomartinelli.it/Fondamenti/Tutorial\\_0.pdf](http://www.adrianomartinelli.it/Fondamenti/Tutorial_0.pdf). Consulté le [8/6/2022].
- [63] Lipinski C.A., Lombardo F., Dominy B.W et feeney P J.2001.Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*.**46**:3-26.
- [64] Veber D.F., Johnson S.R. , Cheg H. Y., Smith B. R., W et kopple K. D.2002.Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *Journal of medicinal chemistry*.**45**:2615-2623.
- [65] Wang Y., Xiao J., O. Suzek T., Zhang J., Wang J., H. Bryant S.2009.PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Research*.**37**:623-633.
- [66] Verdonk M.L., Cole J.C., Hartshorn M.J., Murray C.W., Taylor R.D. 2003.Improved protein–ligand docking using GOLD.*Proteins*.**52**: 609-623.
- [67] Chikhi A., Bensegueni A. 2008. freeware. *Journal of Computer Science & Systems Biology*.**1**:081-086.
- [68] Vieth M., Hirst J.D., Kolinski A., Brooks C.L. 1998. Assessing energy functions for flexible docking. *Journal of Computational Chemistry*. **19**:1612-1622.

[69] Les enzymes. [En ligne].[https://biotech.spip.ac-rouen.fr/IMG/pdf/Les\\_enzymes.pdf](https://biotech.spip.ac-rouen.fr/IMG/pdf/Les_enzymes.pdf) .  
[Consulté le 05/05/2022].

[70] Walters W. P., Stahl M. T., Murcko M. A. 1998. Virtual screening - an overview. *Drug Discovery Today*.**3**(4): 160-178.

[71] Lipinski C A., Lombardo F., Dominy B W., Feeney P J. 1998.Experimental and energy functions for flexible docking. *Journal of Computational Chemistry*. **19**: 1612-22.

[72] Teniou S. Conception in silico de nouveaux inhibiteurs de la topoisomérase I de cellules cancéreuses. Mémoire de magistère en Biochimie. Université Mentouri Constantine 1. Algérie. 2012.103p.

# ANNEXES



**Annexe 1. Liste de 150 complexes utilisés pour tester la fiabilité de programme Surflex-dock avec le RMSD.**

<b>Numéro</b>	<b>Code</b>	<b>RMSD Å</b>
1	1A4G	0.36
2	1A4Q	0.44
3	1AQN	1.00
4	1C1C	0.54
5	1W8N	1.04
6	1LRU	1.24
7	1FKO	0.61
8	1BQN	1.53
9	1NSC	1.06
10	1VEV	1.10
11	1X70	1.20
12	1NNB	1.00
13	1F8B	1.07
14	1WCQ	0.63
15	1VCJ	0.45
16	1MT3	1.76
17	1G2A	1.35
18	1CX2	1.71
19	1E8V	0.98
20	1MWE	0.35
21	1EET	0.47
22	1V3D	1.33
23	1YW9	1.08
24	1L7H	0.83
25	1IVD	1.78
26	1WS1	1.46
27	1EUS	0.59
28	1B9T	0.80
29	1FK9	0.34
30	1EP4	0.59
31	1SLI	1.13
32	1HNV	0.95
33	1YW7	1.09
34	1C23	0.59
35	1IX1	0.90
36	1R5H	1.52
37	1R58	0.41
38	1L7F	1.18
39	1S1W	0.75
40	1C27	0.70
41	1V3C	1.13
42	1LOK	0.68
43	1ICJ	<b>8.95</b>
44	1CP6	0.96

45	1NSN	0.85
46	1EUU	0.94
47	1TKT	1.97
48	1DTT	0.47
49	1HNI	1.17
50	1BSZ	1.13
51	1XNZ	0.16
52	1G27	1.54
53	1YVM	0.06
54	1TV6	2.54
55	1XRY	2.53
56	1VEY	<b>2.55</b>
57	1FT7	2.49
58	1YW8	2.94
59	1S17	2.17
60	1R5G	2.15
61	1B9S	1.27
62	1KQ0	0.46
63	2BAN	1.24
64	2QWN	0.70
65	2RIP	0.59
66	2HTU	0.39
67	2BUB	1.09
68	2SIM	1.44
69	2HPT	1.13
70	2WON	0.48
71	2QTB	0.55
72	2VVZ	1.34
73	2RF2	0.50
74	2AI7	1.91
75	2OPH	0.41
76	2QT9	1.39
77	2VG7	0.98
78	2ZOQ	1.02
79	2B5J	<b>0.22</b>
80	2VW1	0.55
81	2DQM	1.23
82	2B9V	1.38
83	2FJP	0.33
84	2ZD1	1.10
85	2GG5	0.26
86	2QWB	1.59
87	2Q96	0.07
88	2ADU	2.71
89	2OS3	2.50
90	2P98	0.89
91	2X52	2.02
92	2BER	1.54

93	2ZE2	1.91
94	2RKI	1.41
95	2AIE	2.36
96	2EVO	1.89
97	2IIV	2.42
98	2CCV	0.89
99	2HT7	1.71
100	3D4L	0.37
101	3IU9	1.35
102	3E3U	0.58
103	3DLG	0.72
104	3DOL	0.95
105	3KWF	0.60
106	3GCQ	1.11
107	3AH1	0.86
108	3GCU	0.71
109	3EBH	0.95
110	3C6U	0.75
111	3G5K	0.82
112	3NEW	0.47
113	3HV9	0.70
114	3GCP	1.23
115	3B7E	1.18
116	3DOK	1.03
117	3GC8	0.74
118	3ZHZ	2.30
119	3FMM	0.44
120	3PTB	1.18
121	3OBJ	0.99
122	3NNU	0.46
123	3L8X	1.14
124	3HRB	0.79
125	3GC7	0.63
126	2M2W	1.25
127	3K6L	1.24
128	3IU9	1.35
129	3B7U	1.24
130	3TAV	0.82
131	3CHR	2.14
132	3FGD	2.13
133	3FHR	2.41
134	3PKA	1.73
135	3D27	0.72
136	3MBL	1.82
137	3LQL	0.91
138	3ONC	0.34
139	3DM2	0.36
140	3EYK	1.15

<b>141</b>	3IU7	2.62
<b>142</b>	3CMD	0.13
<b>143</b>	3IU8	0.22
<b>144</b>	3HL7	1.56
<b>145</b>	3MH3	1.17
<b>146</b>	3KED	3.02
<b>147</b>	3FFI	0.14
<b>148</b>	3SLI	0.37
<b>149</b>	3CKZ	0.34
<b>150</b>	4OLC	0.73

---

**Thème : Identification de nouveaux inhibiteurs de la PLprotéase en tant qu'une stratégie thérapeutique pour traiter le COVID-19 : criblage *in silico* et filtrage ADME-Tox**

---

**Résumé**

Le COVID-19 est une maladie respiratoire potentiellement mortelle causée par le virus SARS-COV-2. Il s'est propagé rapidement, d'abord à travers la Chine, puis à l'étranger, provoquant une épidémie mondiale.

Notre travail s'inscrit dans le cadre de rechercher *in silico* de nouveaux traitement du coronavirus en inhibant l'enzyme PL protéases (PLpro)avec de nouveaux inhibiteurs. Nous avons effectué un criblage virtuel basé sur la structure de la cible enzymatique PLpro de SARS-COV-2. Cette approche permet de simuler les interactions entre une protéine et des milliers de molécules dans le but de découvrir des composés ayant une activité inhibitrice plus élevée contre une cible thérapeutique.

Le programme Surflex-dock a été utilisé pour découvrir de nouveaux inhibiteurs plus puissants de la PLpro. Avec un pourcentage important de valeurs RMSD inférieures ou égales à 2 Å (88%), on peut dire que ce programme fonctionne bien. On peut donc s'en servir pour se plonger dans le mécanisme d'inhibition de PLpro sans trop de risque d'erreur. En effet, une collection de 555 molécules de la chimiothèque PubChem a été testée contre le site actif de la PLpro. Le criblage virtuel par Surflex-dock ressortir les composés CID149925965 et CID142139076 comme meilleurs inhibiteurs potentiellement plus actif envers la PLpro avec des affinités égales à  $9.52M^{-1}$  et  $8.20 M^{-1}$ , respectivement.

Enfin, l'application de la règle de *Lipinski* nous renseigné de manière positive sur les propriétés pharmacocinétiques (ADME-Tox) de ces deux inhibiteurs proposés.

---

**Mots clés :** SARS-COV-2, PL protéases, criblage virtuel, RMSD, Surflex-dock, règle de *Lipinski*

---

**Devant le jury :**

**Présidente :** MERZOUG Amina

M.C.A Centre.Univ.A.Boussouf - Mila

**Promotrice :** BOUCHERIT Hanane

M.C.B Centre.Univ.A.Boussouf - Mila

**Examinatrice:** LALAOUI Meryem

M.A.BCentre.Univ.A.Boussouf - Mila