



الجمهورية الجزائرية الديمقراطية الشعبية  
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE  
وزارة التعليم العالي و البحث العلمي  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA  
RECHERCHE SCIENTIFIQUE



Université Frères Mentouri Constantine  
Faculté des Sciences de la Nature et de la Vie

جامعة الاخوة منتوري قسنطينة  
كلية عاوم الطبيعة و الحياة

قسم الكيمياء الحيوية و البيولوجيا الخلوية و الجزيئية

**Département de Biochimie et Biologie Cellulaire et Moléculaire**

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

**Domaine : Sciences de la Nature et de la Vie**

**Filière : Sciences biologiques**

**Spécialité : Biochimie appliquée**

Intitulé :

Conception *in silico* de nouveaux inhibiteurs de l'Anaplastic lymphoma kinase (ALK) une cible thérapeutique pour le traitement du cancer du poumon non à petites cellules (CPNPC).

**Présenté et soutenu par :** BERLAT Ahmed Samy

**Le :** 14/07/2021

BOUANAKA Ahmed

**Jury d'évaluation :**

**Président du jury :** Mme GUENDOUZE.A (Maître de conférences B, UFM - Constantine 1)

**Examineurs :** Mme TENIOU.S (Maître assistante, UFM - Constantine 1)

**Encadreur :** Mme DEMMAK.R-G (Maître de conférences B, USB - Constantine 3)

*Année universitaire*

*2020 – 2021*

# *Remerciements*

*Nous remercions tout d'abord le tout puissant qui, par sa grâce nous a permis d'arriver au bout de nos efforts en nous donnant la santé et la volonté d'entamer et de terminer ce mémoire.*

*Ce travail est le fruit de la combinaison d'efforts de plusieurs personnes, on souhaite adresser nos remerciements les plus sincères aux personnes qui nous ont apporté leur aide et qui ont contribué de près ou de loin à l'élaboration de ce mémoire.*

*Tout d'abord ce travail ne serait pas aussi riche et n'aurait pas pu voir le jour sans l'aide et l'encadrement de **Mme DEMMAK,R.G.** Maître de conférence «B» à l'Université Salah Boubnider Constantine 3, On la remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité et surtout ses conseils qui nous ont été d'une aide inestimable.*

*Nous remercions **Mme GUENDOUZE,A.** Maître de conférence «B» à l'Université des Frères Mentouri Constantine 1 , pour l'honneur qu'elle nous fait en acceptant de présider ce jury.*

*On tiens à exprimer nos vifs remerciements envers notre examinatrice **Mme Teniou,S.** Maître assistante «A» à l'Université des Frères Mentouri Constantine 1 d'avoir accepté d'évaluer notre modeste travail et d'être parmi le jury.*

*Nos sincères remerciements vont au Chef de notre Spécialité **Mr BENSEGUENI,A.** Professeur à l'Université des Frères Mentouri Constantine 1. Ainsi qu'à tous les enseignants de la Spécialité Biochimie Appliquée du Département de Biochimie et Biologie Cellulaire et Moléculaire, Université des Frères Mentouri Constantine 1.*

# ***Dédicaces***

## ***À mes chers parents***

*Aucune dédicace ne saurait exprimer mon respect, mon amour  
éternel et ma*

*Considération pour les sacrifices que vous avez consentis pour mon  
instruction*

*Merci pour tout le soutien et l'amour que vous m'avez apporté tout au  
long de mes études.*

## ***À mes chères sœurs***

*Comme preuve de mes amours fraternels, de mes profondes affections  
et de mes gratitude, je vous souhaite une vie pleine de bonheur et de  
succès.*

*À toute ma famille pour leur soutien tout au long de mon parcours  
universitaire.*

*À mon binôme SAMY, en souvenir de tous les moments passés  
ensemble*

*À toute ma promotion Biochimie Appliquée 2021, et surtout mes  
ami(e)s KENZA, RAMZI, SAMI, ALI, ABDELHAK, ABDELMALEK,  
je n'oublierai jamais les moments passés avec vous.*

***Ahmed***

# *Dédicaces*

*Je remercie en premier lieu dieu le tout puissant de m'avoir donné la force, la volonté, et la patience durant toutes mes années d'étude.*

## *À ma très chère mère*

*Quoi que je fasse ou quoi que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance et tes prières me guident, ta présence à mes côtés a toujours été ma source de force pour affronter les différents obstacles.*

## *À mon très cher père*

*Tu as toujours été à mes côtés pour me soutenir, me diriger et m'encourager. Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de toi. Que ce travail traduit ma gratitude et mon affection.*

## *À mon très cher frère, ma source de joie et de motivation.*

*À mon binôme AHMED, ce fut un plaisir de travailler avec toi, un grand MERCI à vous.*

*À mon ami CHELLALI RAMZI, merci pour tes conseils et ton aide, je te souhaite le meilleur pour l'avenir.*

*À toute ma promotion Biochimie Appliquée 2021, et surtout mes ami(e)s KENZA, SAMI, ALI, ABDELHAK, ABDELMALEK, je n'oublierai jamais les moments passés avec vous.*

*À toutes les personnes ayant participé de près ou de loin à la réalisation de ce projet.*

***Samy***

# TABLE DES MATIERES

<b>Introduction .....</b>	<b>1</b>
---------------------------	----------

## **Partie Bibliographique**

### CHAPITRE 1 : LE CANCER DU POUMON

1. Généralités .....	2
2. Cancer du poumon .....	3
2.1. Classification des tumeurs broncho-pulmonaires .....	5
3. Épidémiologie.....	5
3.1. Statistiques mondiales .....	5
3.2. Statistiques nationales .....	6
4. Les facteurs de risques .....	7
5. Symptômes.....	8
5.1. Les signes généraux .....	8
5.2. Les signes indirects .....	8
6. Diagnostics .....	9
6.1. L'analyse d'imagerie .....	9
6.2. Confirmation du diagnostic .....	10
7. Traitements .....	10
7.1. La chirurgie .....	10
7.2. La chimiothérapie .....	11
7.3. La radiothérapie .....	11

## CHAPITRE 2 : ANAPLASTIC LYMPHOMA KINASE

1. Généralités .....	12
2. Structure .....	12
3. Site actif .....	14
4. Mécanisme d'action .....	14
5. L'alk et le cancer .....	15
6. Les inhibiteurs de l'alk .....	17
6.1. Inhibiteurs de première génération .....	17
6.2. Inhibiteurs de deuxième génération .....	18
6.3. Inhibiteurs de troisième génération .....	19

## CHAPITRE 3 : DOCKING MOLECULAIRE

1. Généralités .....	20
2. Définition de docking moléculaire .....	20
3. Principe de docking .....	21
4. Types de docking .....	21
5. Outils de docking moléculaire .....	22
5.1. Récepteur 3D (la cible) .....	22
5.2. Ligands 3D (Chimiothèques) .....	23
5.3. Programmes (Logiciels) .....	24

## **Partie Pratique**

### **CHAPITRE 4 : MATERIALS AND METHODS**

1. Softwares and tools used.....	27
1.1. Surfex .....	27
1.2. ChemDraw 3D.....	27
1.3. Open Babel .....	27
1.4. BIOVIA Discovery Studio .....	27
1.5. Excel Microsoft 2016 .....	27
2. PDB (Protein Data Bank).....	27
3. Evaluation of Programs.....	27
3.1. The RMSD (root mean square deviation) .....	27
3.2. Correlation coefficient (R) .....	28
4. Preparation of molecules for molecular docking .....	28
5. Calculation of interactions “enzyme- similar”.....	29
5.1. PubChem .....	29
6. Lipinski rule .....	29

### **CHAPITRE 5 : RESULTS AND DISCUSSION**

1. Reliability of the programs used.....	30
1.1. The RMSD.....	30
1.2. Visual analysis.....	31
1.3. The correlation coefficient .....	33

2. Study of the interactions involved in the inhibition of the human ALK .....	33
2.1. Interaction 5IUI-45Q .....	33
3. Proposed new inhibitors of human ALK.....	34
3.1. Inhibition of the human ALK by the similar of compound 45Q.....	34
4. Drug-likeness Prediction .....	37
Conclusion et perspective.....	39

## **Références bibliographiques**

## **Annexes**

## **Résumés**



## LISTE DES ABREVIATIONS

**3D** : Trois dimensions.

**Å** : Angström.

**ADME** : Absorption, Distribution, Métabolisme, Excrétion.

**ADN**: Acide désoxyribonucléique.

**ALK** : Anaplastic lymphoma kinase.

**ARN**: Acide ribonucléique.

**ATP** : L'adénosine triphosphate.

**CBNPC** : Cancers bronchiques non à petites cellules.

**CBPC** : Cancers bronchiques à petites cellules.

**EGFR**: Epithelial Growth Factor Receptor.

**EML4**: Echinoderm microtubule-associated protein-like 4.

**HAP** : Hydrocarbures aromatiques polycycliques.

**IC50** : Inhibitory Concentration 50.

**IRM** : l'imagerie par résonance magnétique.

**LDL-A**: Low-Density Lipoprotein class A.

**Log P** : Le coefficient de partition Eau/Octanol.

**MAM**: Meprin, A-5 protein, receptor protein tyrosine phosphatase mu.

**MAPK**: Mitogen-activated protein kinases.

**NCBI**: National Center for Biotechnology Information.

**NCI**: National Institutes of Health.

**NSCLC**: Non-small cell lung cancer.

**PDB:** Protein Data Bank.

**PI3K:** Phosphoinositide 3-kinase.

**PLC :** Phospholipase C.

**RMN :** Résonance magnétique nucléaire.

**RMSD:** Root Mean Square Deviation.

**RTK:** Récepteur Tyrosine Kinase.

**STAT:** Signal Transducers and Activators of Transcription.

**TDM :** Tomodensitométrie.

## LISTE DES FIGURES

<b>Figure 01</b> : La formation de métastases.....	3
<b>Figure 02</b> : Anatomie des poumons.....	4
<b>Figure 03</b> : Incidence et mortalité mondiale en 2020.....	6
<b>Figure 04</b> : Incidence et mortalité chez les hommes en 2020.....	6
<b>Figure 05</b> : Structure de l'ALK.....	13
<b>Figure 06</b> : Motifs structuraux caractéristiques de l'ALK.....	13
<b>Figure 07</b> : Site actif de l'ALK.....	14
<b>Figure 08</b> : Voies de signalisation de l'ALK.....	15
<b>Figure 09</b> : Représentation schématique de translocation entre EML4 et ALK.....	16
<b>Figure 10</b> : Les Voies de signalisation de EML4-ALK.....	17
<b>Figure 11</b> : la structure du Crizotinib.....	18
<b>Figure 12</b> : la structure de l'alectinib.....	19
<b>Figure 13</b> : la structure de la céritinib.....	19
<b>Figure 14</b> : la structure de Lorlatinib.....	19
<b>Figure 15</b> : Principe générale d'un programme de docking.....	21
<b>Figure 16</b> : Modèle clé-serrure.....	22
<b>Figure 17</b> : Interface d'accueil du PDB.....	23
<b>Figure 18</b> : Interface d'accueil du PubChem.....	24
<b>Figure 19</b> : Utilisation des différents logiciels de docking en 2010-2011.....	25
<b>Figure 20</b> : structure of the complex 5IUI.....	28
<b>Figure 21</b> : Results as a percentage of the RMSD test at two intervals.....	30

<b>Figure 22</b> : Linear correlation curve between Log IC <sub>50</sub> of the 30 inhibitors and their scores .....	<b>33</b>
<b>Figure 23</b> : Representation of interactions formed by 45Q with the active site of 5IUI..	<b>34</b>
<b>Figure 24</b> : Representation of interactions formed by the similar CID_ 70654694.....	<b>35</b>
<b>Figure 25</b> : Representation of interactions formed by the similar CID_70655002.....	<b>36</b>

## **LISTE DES TABLEAUX**

<b>Tableau 1</b> : Les programmes de Docking moléculaire.....	<b>25</b>
<b>Tableau 2</b> : The RMSD values of the 3 complexes studied and superpositions.....	<b>31</b>
<b>Tableau 3</b> : Docking results of the reference ligand and the two similar compounds with the highest affinity.....	<b>35</b>
<b>Tableau 4</b> : The predicted pharmacokinetic and toxicity proprieties of the most promising compounds.....	<b>37</b>



# ***Introduction***

## Introduction

La recherche dans l'industrie pharmaceutique est un processus qui prend du temps avec un risque élevé d'échec. Pour cela, le processus de développement des médicaments a été amélioré grâce à l'utilisation des techniques, qui sont efficaces lorsqu'elles sont combinées à des outils informatiques. Le Docking est une des méthodes de la modélisation moléculaire, devenu un outil indispensable dans le domaine de la conception de molécules médicamenteuses pour but à prédire les détails de l'interaction moléculaire entre une molécule et une protéine cible ayant un intérêt thérapeutique [1].

Le cancer du poumon représente un problème de santé publique majeur. En effet c'est le type de cancer le plus fréquent au monde chez l'homme comme chez la femme. De plus, le cancer du poumon est la cause principale de mort par cancer dans le monde [2].

L'identification de sous-ensembles moléculaires du cancer du poumon avec des mutations motrices ciblables a modifié le paysage du traitement. En 2007, la découverte d'altération moléculaire, activant les fusions du gène de la kinase du lymphome anaplasique (ALK), a été découverte dans un ensemble des patients atteints de cancer du poumon non à petites cellules (NSCLC). Une fusion d'ALK avec EML4 provoque l'activation constitutive du domaine ALK kinase et la signalisation oncogène subséquente, généralement via les voies MAPK, JAK-STAT et PI3K-AKT [3].

Dans ce contexte, nous nous sommes intéressés à l'usage de la modélisation moléculaire, en utilisant le « Docking moléculaire » pour étudier l'inhibition de l'enzyme *anaplastic lymphoma kinase (ALK)* et la prédiction des modes d'interactions possibles entre « enzyme–ligand ».

L'objectif du présent travail est de :

- Tester la fiabilité du programme de docking moléculaire "Surflex" utilisé dans cette étude.

Nous nous sommes servis du RMSD où critère de déviation de la structure du ligand calculée par le programme par rapport à sa conformation donnée par diffraction des rayons X. Le test a été réalisé sur 100 complexes pris de la Protéine Data Bank "PDB".

- Calculer les affinités des similaires de ligand de référence.
- Mettre en évidence les liaisons hydrogènes impliquées dans la formation des complexes « inhibiteurs-ALK ».
- Etudier *in silico* les propriétés biologiques ADME/ Tox de ces composés.

# *Chapitre 1 : Le cancer du poumon*

## **I.1.Généralités**

Le terme « cancer » est utilisé depuis l'antiquité pour désigner toute excroissance anormale comme des tumeurs ou des ulcérations. Depuis la découverte scientifique de cette maladie, ce terme est d'autant plus utilisé pour désigner les maladies oncologiques (« oncos » masse, grosseur)[4].

Le cancer est une maladie chronique, possiblement mortelle, qui apparaît lorsque des substances toxiques entraînent des mutations génétiques dans l'organisme, et qu'un processus inflammatoire les fait croître. Notre système immunitaire empêche normalement le développement du cancer en réparant l'ADN (constituant essentiel de la cellule vivante) ou en tuant les cellules endommagées. La maladie se manifeste lorsque l'équilibre entre des cellules cancéreuses « dormantes » dans l'organisme et les défenses naturelles du corps se rompt et le terrain devient alors propice [5].

Malgré le nombre immense des cellules de l'organisme adulte, leur prolifération physiologique reste un processus strictement contrôlé qui permet le renouvellement ou la répartition des tissus lésés. Parmi ces milliards de cellules une seule n'obéit plus à cette organisation et prolifère sans le moindre contrôle. Quand une prolifération cellulaire mono ou oligo clonale a tendance à persister, à s'accroître, à détruire les tissus avoisinants et à diffuser sous forme de métastases, on parle de cancer [6].

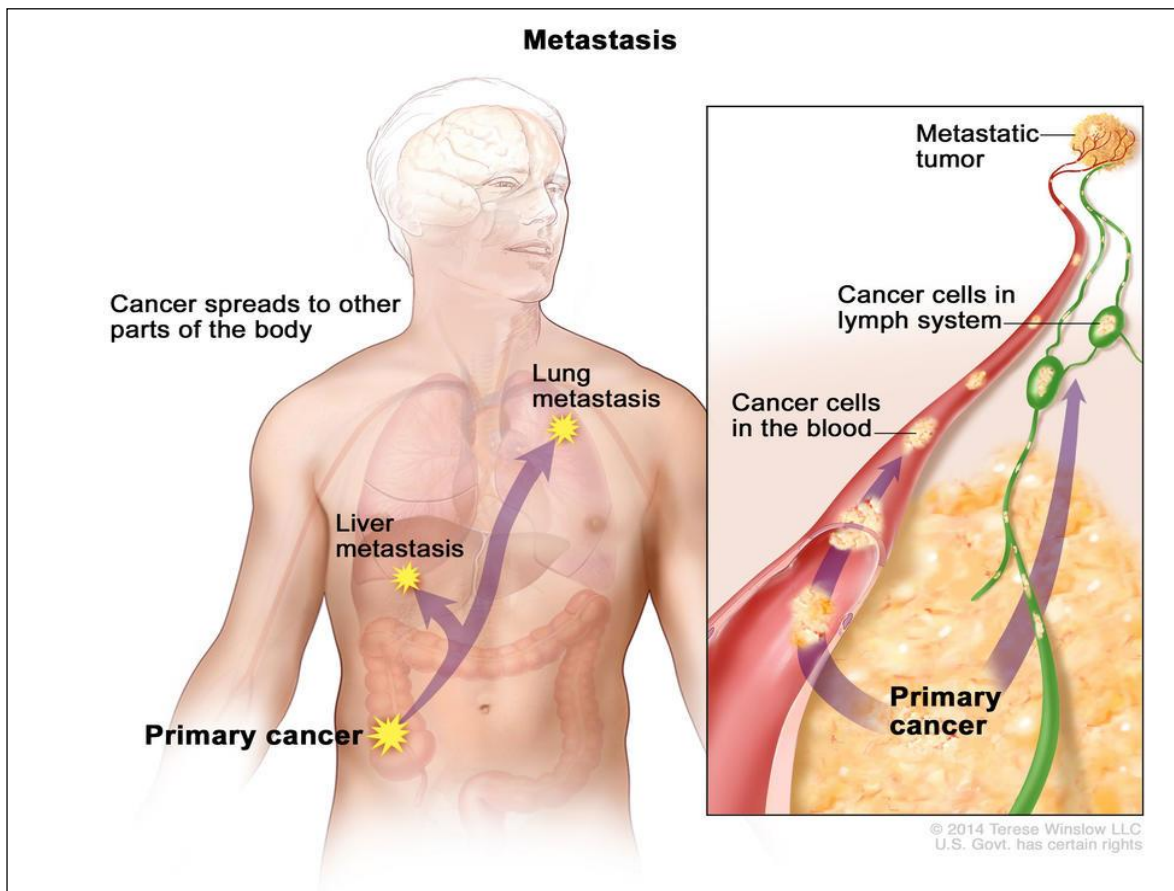
Le cancer se caractérise par une croissance cellulaire non régulée ainsi que par l'invasion et la propagation des cellules à partir de leur site d'origine, vers d'autres sites de l'organisme. Cela conduit à la distinction entre une tumeur bénigne et une tumeur maligne. Les tumeurs bénignes ne se propagent pas dans l'ensemble de l'organisme (c'est-à-dire qu'elles ne forment pas de métastases), bien que certaines d'entre elles puissent menacer le pronostic vital en raison de leur localisation (par exemple, une tumeur cérébrale bénigne qui peut être difficile à enlever). Les tumeurs malignes, en revanche, ne restent pas encapsulées, présentent des caractéristiques d'invasion et forment des métastases[7]. Un cancer qui s'est propagé de l'endroit où il a commencé à un autre endroit du corps est appelé cancer métastatique. Le processus par lequel les cellules cancéreuses se propagent à d'autres parties du corps est appelé métastase.

Le cancer métastatique a le même nom et le même type de cellules cancéreuses que le cancer initial, ou primaire. Par exemple, un cancer du sein qui se propage et forme une



tumeur métastatique dans le poumon est un cancer du sein métastatique, et non un cancer du poumon.

Au microscope, les cellules cancéreuses métastatiques ont généralement le même aspect que les cellules du cancer d'origine. De plus, les cellules cancéreuses métastatiques et les cellules du cancer d'origine ont généralement certaines caractéristiques moléculaires en commun, comme la présence de modifications chromosomiques spécifiques [8].



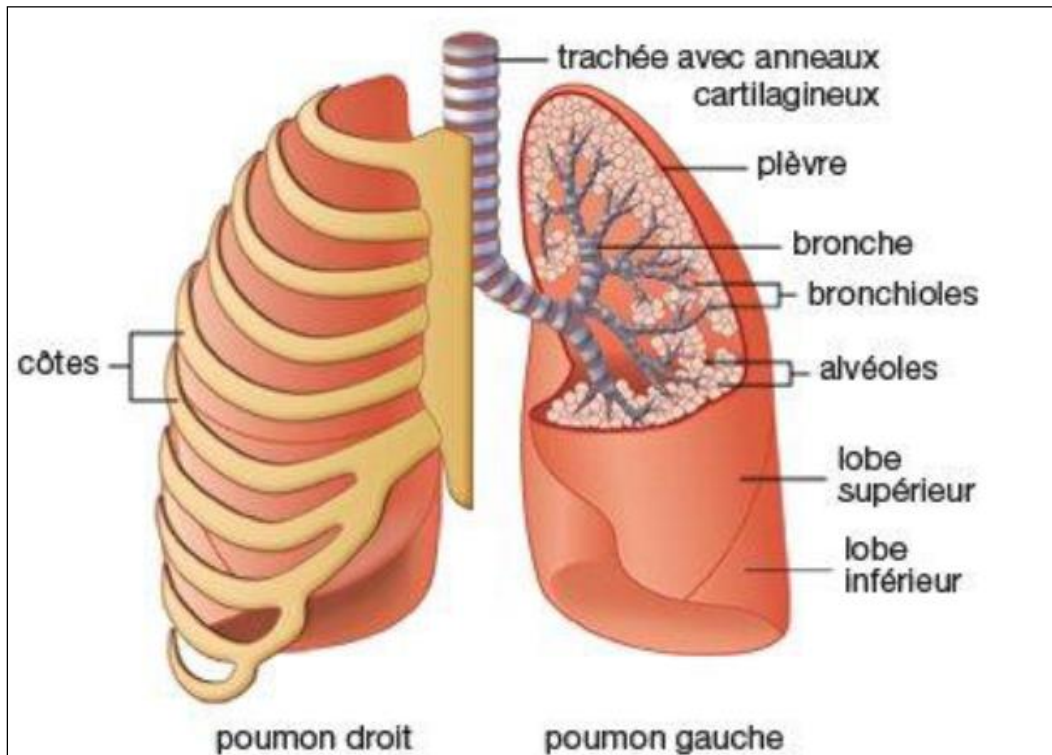
**Figure 01** :La formation de métastases [8]

## I.2. Cancer du poumon

Organe de la respiration, le poumon joue également un rôle important d'épuration et de protection de l'organisme vis-à-vis de l'environnement avec lequel il est en contact aérien permanent [9].

Les poumons sont situés dans le thorax, plus particulièrement au sein de la cage thoracique où ils occupent la majeure partie. Les deux poumons, droit et gauche, sont

séparés par le médiastin (figure 2)[10]. Ils sont composés de bronches et de conduits aériens plus petits, d'alvéoles, de tissu conjonctif, de vaisseaux sanguins, de vaisseaux lymphatiques et de nerfs, tous enchâssés dans une matrice de tissu conjonctif élastique. Chaque lobe est constitué d'un grand nombre de lobules [11].



**Figure 2** : Anatomie des poumons [10]

Le cancer du poumon est un processus en plusieurs étapes impliquant des altérations génétiques et épigénétiques, au cours duquel les lésions de l'ADN qui en résultent transforment les cellules épithéliales pulmonaires normales en cancer [12].

Sa tumorigenèse est liée à l'activation de protéines favorisant la croissance et à l'inactivation de gènes suppresseurs de tumeurs. L'activation d'oncogènes favorisant la croissance peut se produire par l'amplification de gènes ou d'autres altérations génétiques, y compris des mutations ponctuelles et des réarrangements structuraux conduisant à une signalisation incontrôlée par des voies oncogènes [13].

### **I.2.1 Classification des tumeurs broncho-pulmonaires**

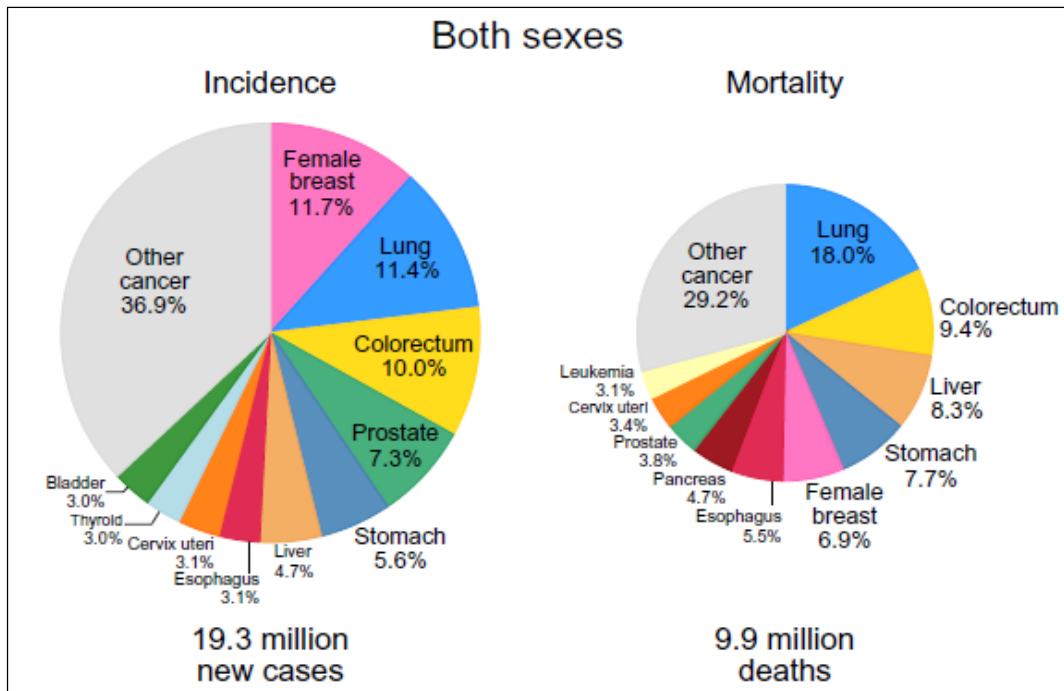
Sur le plan histologique, il s'agit d'un groupe hétérogène de tumeurs, parmi lesquelles on distingue deux types : les cancers bronchiques non à petites cellules (CBNPC) (85%) et les cancers bronchiques à petites cellules (CBPC) (15%). Dans le CBNPC, les caractéristiques histogénétiques et immunohistochimiques permettent de distinguer 3 types : l'adénocarcinome (40%), le carcinome épidermoïde (30%) et le carcinome à grandes cellules (10%) [14].

## **I.3. Épidémiologie**

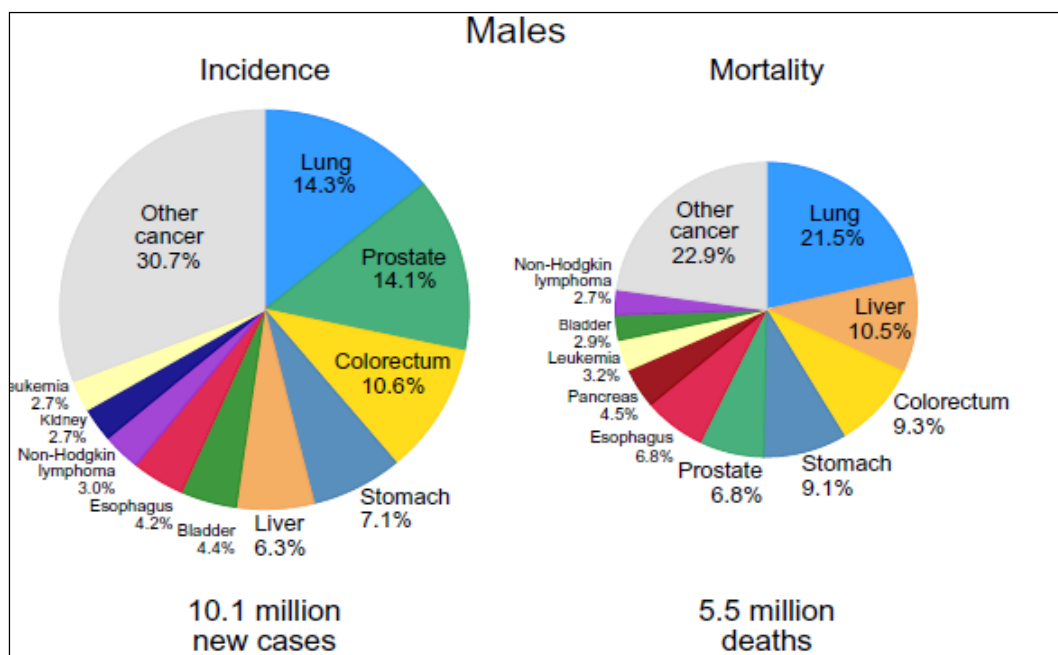
### **I.3.1 Statistiques mondiales**

Au niveau international, le cancer du poumon reste la principale cause de décès liés au cancer chez les hommes et les femmes [15].

Avec une estimation de 2,2 million de nouveaux cas de cancer et 1,79 million de décès, le cancer du poumon est le deuxième cancer le plus fréquemment diagnostiqué et la principale cause de décès par cancer en 2020, représentant environ un cancer diagnostiqué sur 10 (11,4 %) et un décès sur 5 (18,0 %) (Figure 3). Le cancer du poumon est la principale cause de morbidité et de mortalité par cancer chez les hommes (Figure 4), tandis que chez les femmes, il se classe au troisième rang pour l'incidence, après le cancer du sein et le cancer colorectal, et au deuxième rang pour la mortalité, après le cancer du sein. Les taux d'incidence et de mortalité sont environ deux fois plus élevés chez les hommes que chez les femmes [16].



**Figure 3** : Incidence et mortalité mondiale en 2020 [16]



**Figure 4** : Incidence et mortalité chez les hommes en 2020 [16]

### I.3.2 Statistiques nationales

Selon les statistiques de The Global Cancer Observatory (GLOBOCAN). Le cancer du poumon constitue en Algérie (2020), un véritable fléau et il continue de se propager d'une

manière inquiétante. En effet, on estime jusqu'à 4 774 le nombre de nouveaux cas enregistrés (8.2%) avec principalement 4 159 nouveaux cas enregistrés chez l'homme (15.2%), et à 4 333 le nombre de décès (13.2%) [17].

#### **I.4. Les facteurs de risques**

La plupart des causes du cancer bronchique sont environnementales, mais il existe des facteurs de susceptibilité et de variations individuelles aux cancérogènes pulmonaires.

Les principaux facteurs de risque sont le tabagisme actif, l'exposition au tabagisme passif, les expositions professionnelles, et ont des interactions entre eux car un même individu peut avoir l'association de plusieurs facteurs de risque.

##### **➤ Tabagisme actif**

Fumer du tabac, en particulier sous forme de cigarette, est la principale cause du cancer du poumon. La fumée du tabac contient de nombreuses substances chimiques nuisibles qui peuvent causer le cancer bronchique (carcinogènes). Le tabagisme affecte la santé et provoque des changements génétiques dans les cellules pulmonaires qui mènent au développement d'un cancer du poumon.

Le risque d'être atteint d'un cancer du poumon varie en fonction de la durée du tabagisme, de l'âge auquel la personne a commencé à fumer et du nombre de cigarettes fumées chaque jour [18]

##### **➤ Tabagisme passif**

La composition chimique de la fumée inhalée directement de la cigarette et celle de la fumée d'une cigarette se consommant sont similaires. Le tabagisme passif est aussi appelé fumée de tabac ambiante, ou inhalation involontaire de fumée.

Les personnes qui sont exposées à la fumée secondaire risquent davantage d'être atteintes d'un cancer du poumon. La fumée secondaire est un facteur de risque important du cancer du poumon chez les non-fumeurs [19].

##### **➤ Radon**

Le radon est un gaz incolore, inodore et sans goût qui provient de la désintégration naturelle de l'uranium présent dans le sol et la pierre. L'inhalation du radon risque d'endommager les cellules qui tapissent les poumons [20].

L'exposition au radon fait augmenter le risque de cancer du poumon. Le radon est la principale cause de cancer du poumon chez les non-fumeurs, et la deuxième plus importante cause de ce cancer chez les fumeurs, le risque d'apparition du cancer du poumon dépend de la quantité de radon à laquelle la personne a été exposée, de la durée de l'exposition et de la consommation de tabac. Le risque engendré par le radon est beaucoup plus élevé chez les fumeurs [21].

### ➤ La pollution de l'air ambiant

La pollution de l'air ambiant est impliquée dans divers problèmes de santé, dont le cancer, et en particulier dans le développement du cancer du poumon. Le risque de cancer du poumon augmente en fonction de l'exposition à la pollution atmosphérique. Les polluants présents dans l'air varient d'un endroit à l'autre selon les sources d'émissions locales.

Des chercheurs ont démontré que des composants individuels de la pollution de l'air extérieur causent le cancer. Ce sont entre autres les gaz d'échappement des moteurs au diesel, le benzène, la matière particulaire et certains hydrocarbures aromatiques polycycliques (HAP) [22].

## I.5. Les symptômes du cancer du poumon

Le cancer peut causer certains symptômes. Mais ces signes peuvent aussi être dus à d'autres maladies. Ils doivent donc être évalués par un médecin. Voici les principaux :

### I.5.1 Les signes généraux

- Des problèmes respiratoires, par exemple un essoufflement ou une respiration sifflante.
- Une toux qui s'aggrave ou persistante.
- Une douleur thoracique persistante plus vive lors de quintes de toux ou de respirations profondes.
- Crachat de sang.
- Perte involontaire de poids, perte de l'appétit et faiblesse généralisée.
- Infection aux poumons (bronchite ou pneumonies), fréquente ou continue [23].

### I.5.2 Les signes indirects

Regroupés sous le terme de « syndrome paranéoplasique », pouvant apparaître plusieurs années avant la tumeur du poumon, observé dans 10 à 20 % des cas :

- Déformation des dernières phalanges et des ongles des doigts (hippocratisme digital).
- Hypercalcémie.
- Troubles de la régulation de l'hydratation.
- Thromboses veineuses et phlébites [24].

## I.6. Les moyens de diagnostic

Le bilan diagnostique repose notamment sur :

### I.6.1 L'analyse d'imagerie :

Des techniques comme la radiographie, l'échographie, la tomodensitométrie TDM, l'imagerie par résonance magnétique IRM permettant de procéder à un examen approfondi des tissus, des organes et des os.

Les analyses d'imagerie sont également utilisés pour déterminer où le cancer se situe, En plus pour surveiller la réaction du cancer au traitement [25].

- **Radiographie thoracique** : Une radiographie thoracique est souvent le premier test effectué lorsque les symptômes semblent correspondre à ceux du cancer du poumon. Il s'agit d'une procédure rapide et indolore qui prend une image en utilisant une faible dose de rayons X [26].
- **Tomodensitométrie** : Utilise également des rayons X, mais il crée des images plus détaillées qu'une radiographie conventionnelle. La tomodensitométrie peuvent détecter des tumeurs plus petites, à un stade plus précoce, et déterminer si le cancer s'est propagé aux ganglions lymphatiques voisins ou à proximité des gros vaisseaux sanguins, ou encore à la cage thoracique [26].
- **Imagerie par résonance magnétique** : l'IRM utilise des champs magnétiques et des ondes radio pour produire une image 3D. L'IRM n'a pas de place dans le cancer du poumon, sauf cas particuliers (exemple : tumeur de l'apex pulmonaire), elle est plus sensible aux autres facteurs des mouvements cardiaques et respiratoires. En plus, la faible densité en protons du parenchyme pulmonaire et l'inhomogénéité du champ magnétique dans les poumons ont empêché le développement antérieur de l'IRM thoracique [27].

### **I.6.2 Confirmation du diagnostic :**

Le diagnostic est confirmé par l'examen anatomopathologique, préférentiellement par une histologie plutôt qu'une cytologie. Il est envisagé en premier lieu une bronchoscopie souple. Si la bronchoscopie souple est non contributive, il sera réalisé une ponction transpariétale si la taille et la localisation de la lésion le permettent.

L'envoi de matériel tumoral à la plateforme d'Oncologie Moléculaire est indiqué dans les cas de carcinome non à petites cellules non épidermoïde devant une tumeur localement avancée ou métastatique. Il y sera recherché les mutations du gène de l'EGFR ainsi que celles des biomarqueurs émergents : BRAF, HER2, PI3KCA et la translocation EML4-ALK [28].

## **I.7. Traitement**

L'identification de l'histologie du cancer du poumon ainsi son stade d'avancement peut faciliter la prescription d'un traitement médical personnalisé et plus efficace [29][30]. Une classification TNM est appliquée et permet de regrouper les tumeurs en fonction de la taille des nodules tumoraux (T), de l'envahissement des ganglions lymphatiques (N) situés plus ou moins loin de la tumeur ainsi que de la présence ou de l'absence de métastases à distance (M) [31].

Mais généralement, dans le cas des cancers bronchiques, la progression tumorale est définie en fonction des stades qui vont de I pour les moins agressifs sans envahissement des ganglions à IV pour les plus agressifs avec présence de métastases [32].

### **I.7.1 La chirurgie**

Les patients atteints du CBNPC de stade I, II et IIIA sont généralement opérés pour retirer la tumeur si celle-ci est résécable et si le patient peut tolérer la chirurgie. Les chirurgiens peuvent enlever un lobe ou une section du poumon contenant la tumeur. Pour déterminer si la tumeur est résécable, des études d'imagerie et des biopsies sont réalisées, ainsi qu'une évaluation des facteurs propres au patient pour déterminer l'opérabilité [33]. Généralement, la chirurgie correspond à une lobectomie, on retire tout un lobe pulmonaire, c'est l'opération la plus fréquente ou à une segmentectomie, on retire la seule partie où se trouve la tumeur ou à une pneumectomie, on retire la totalité du poumon ainsi que les ganglions correspondants [34].



### **I.7.2 La chimiothérapie**

Près de 70 % des patients atteints de cancer du poumon présentent une maladie localement avancée ou métastatique au moment du diagnostic. La chimiothérapie est bénéfique pour la palliation chez les patients atteints d'une maladie localement avancée ou métastatique. La chimiothérapie adjuvante est généralement indiquée pour les patients atteints de CBNPC de stades IIA à IIIA réséqués. Bien que la chimiothérapie soit appropriée pour de nombreux patients atteints de cancer du poumon, on a le sentiment que l'utilisation des agents chimiothérapeutiques traditionnels a atteint un plateau thérapeutique. Une meilleure compréhension de la biologie du cancer a révélé de nombreuses stratégies thérapeutiques potentielles, notamment le ciblage de l'EGFR et d'autres voies de transduction du signal et d'angiogenèse [35].

### **I.7.3 La radiothérapie**

La radiothérapie utilise des rayonnements ionisants pour détruire les cellules cancéreuses en les empêchant de se multiplier. Ces rayonnements (appelés aussi rayons ou radiations) sont dirigés avec précision, sous forme de faisceaux, sur la zone à traiter. Ils visent, à travers la peau, soit la tumeur et les ganglions lymphatiques voisins, soit le lit tumoral (c'est-à-dire la région où se trouvait la tumeur avant l'intervention chirurgicale), soit des métastases.

Il existe différentes techniques de radiothérapie : la radiothérapie conformationnelle en trois dimensions (3D) et la radiothérapie stéréotaxique. La plus utilisée pour traiter les cancers du poumon est la radiothérapie conformationnelle en trois dimensions (3D). Il s'agit d'une radiothérapie externe. Cette technique consiste à faire correspondre le plus précisément possible (autrement dit à conformer) le volume de la zone à traiter et le volume sur lequel vont être dirigés les rayons. Elle utilise des images en 3D de la zone à traiter et des organes avoisinants, obtenues par scanner, parfois associées à d'autres examens d'imagerie médicale (IRM, TEP) [36].

***Chapitre 2 : Anaplastic Lymphoma  
Kinase (ALK)***

## II.1. Généralités

L'anaplastic Lymphoma Kinase (ALK), est un récepteur transmembranaire à activité tyrosine kinase (RTK), qui appartient à la superfamille des récepteurs de l'insuline [37], qui a été clonée en 1997 [38]. Elle est retrouvée en majorité au niveau du cerveau et joue un rôle dans le développement et le fonctionnement du système nerveux central et périphérique chez l'adulte, et pourrait également être impliquée dans la formation de la synapse et la migration des cellules musculaires [39][40]. Le gène ALK humain est localisé sur le bras court du chromosome 2 (2p23) [41].

## II.2. Structure

Il s'agit d'une protéine monomérique transmembranaire de 1 620 acides aminés et de 200 kDa après modifications post-traductionnelles, codée par un ADNc de 6 226 paires de bases. Sa partie extracellulaire est composée de 1038 résidus d'acides aminés (acides aminés 1-1038), incluant un peptide signal N-terminal, suivi de deux domaines MAM (meprin, A-5 protein, receptor protein tyrosine phosphatase mu) qui pourrait jouer un rôle dans les interactions intercellulaires, et un domaine LDL-A (Low-Density Lipoprotein class A), suivie d'une région riche en glycine, faisant suite à un domaine transmembranaire qui relie la région extracellulaire à la région juxtamembranaire du domaine intracellulaire (figure 5) [40][42].

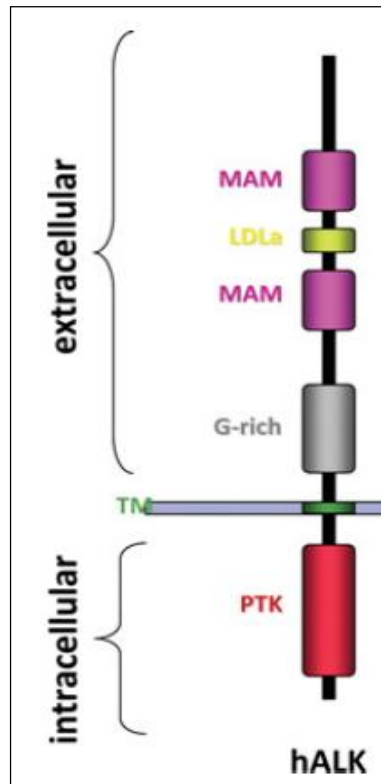
La région intracellulaire comporte le domaine à activité tyrosine kinase de l'ALK, cette région est composée d'un petit lobe amino-terminal et d'un lobe carboxyterminal. Le lobe N-terminal contient cinq feuillets  $\beta$  antiparallèles, une hélice  $\alpha$ C ( $\alpha$ C-regulatory helix) pouvant être dans une conformation active ou inactive, et un domaine de liaison à l'ATP riche en glycine (P-loop).

Le lobe C-terminal quant à lui est composé de cinq hélices  $\alpha$  et de deux feuillets  $\beta$ . Il comprend deux régions essentielles à l'activité de la kinase :

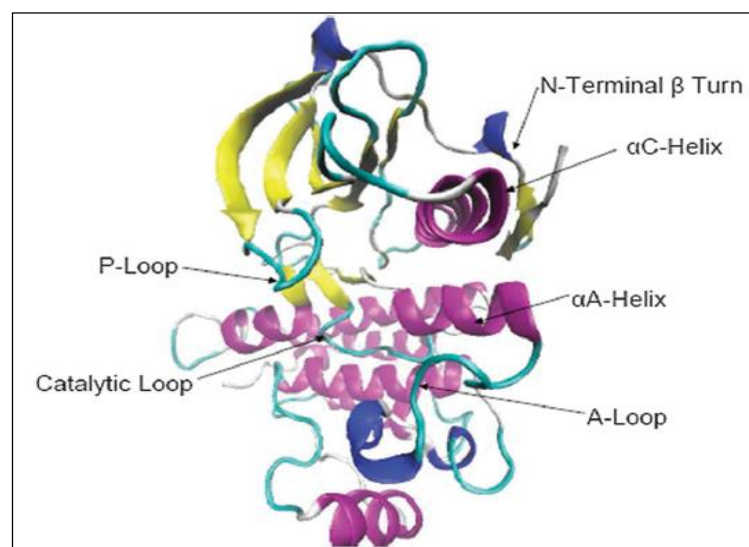
-La boucle d'activation (activating-loop, A-loop) qui commence par un motif Acide aspartique- Phénylalanine-Glycine (motif DFG) et qui a pour fonction de moduler l'efficacité de liaison au substrat ainsi que la vitesse de catalyse. Elle comprend un motif type des récepteurs à l'insuline (motif YxxxYY) comprenant trois tyrosines (Y1278, Y1282, et Y1823) qui représente les principaux sites d'autophosphorylation de la kinase.

La phosphorylation de la tyrosine Y1278 est nécessaire à l'auto-activation du domaine kinase de l'ALK [43].

-Le site catalytique de la kinase, composé de résidus permettant le transfert du phosphate sur le substrat(figure6)[44].



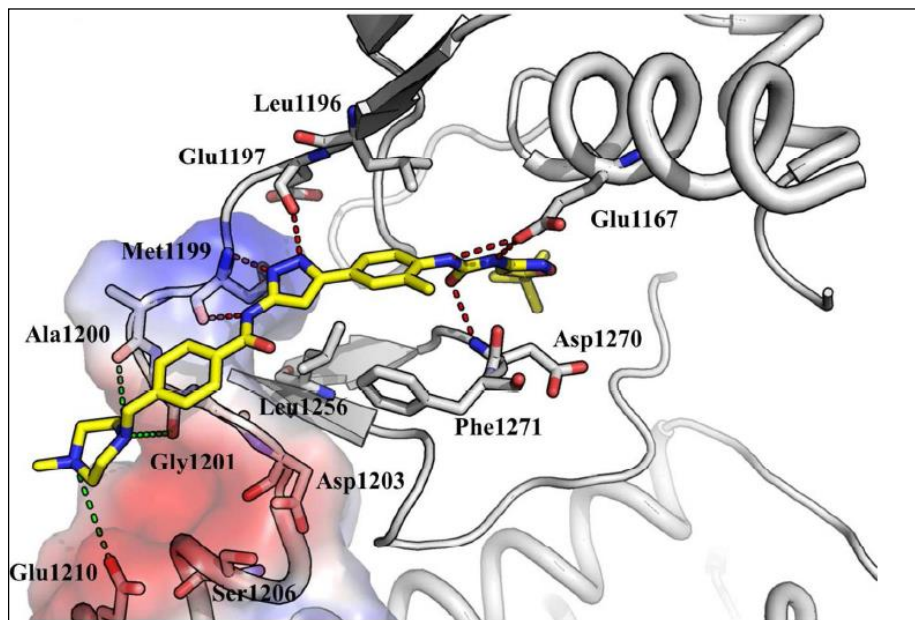
**Figure 5** : Structure de l'ALK [42].



**Figure 6** : Motifs structuraux caractéristiques de l'ALK [44].

### II.3. Site actif

Les interactions par liaison hydrogène ont été reconnues comme la force la plus importante liée à la stabilisation du ligand dans la structure du complexe. En outre, les interactions cation-pi sont également importantes pour la liaison des ligands et la fonction catalytique. Le crizotinib est capable de maintenir trois interactions de liaison hydrogène et une interaction cation-pi, et une interaction hydrophobe. Les résidus Met1199, Asp1203 et Glu1197 sont capables de former une liaison hydrogène avec le crizotinib, et le résidu Lys1150 a pu former des interactions de liaison hydrogène et une interaction cation-pi, en outre le résidu Leu1122 a pu faire des interactions hydrophobes. Il est nécessaire de mentionner l'importance du résidu Met1199 dans la stabilité du ligand, et du résidu Leu1122 qui est important pour l'inhibition de la kinase ALK (figure7) [45][46].

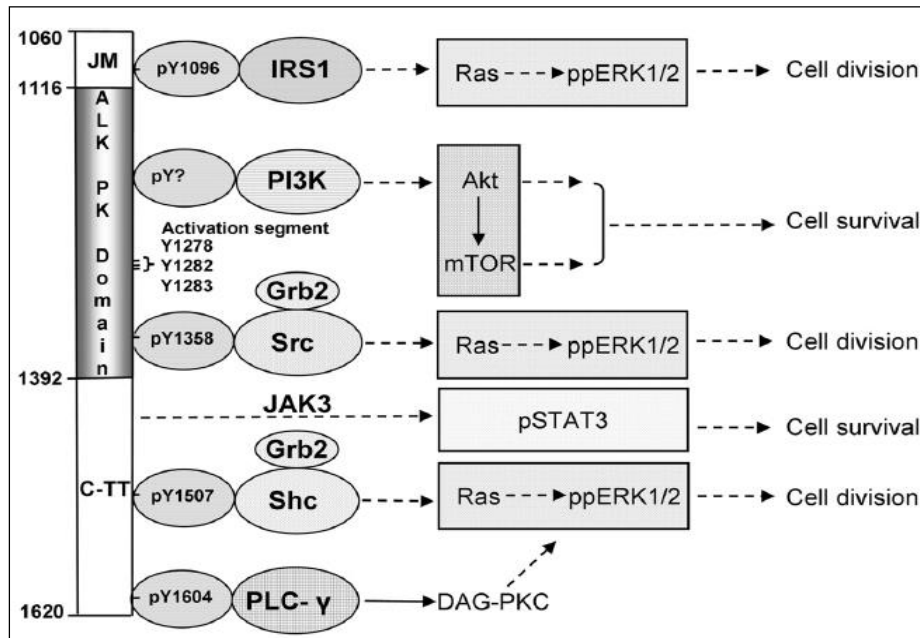


**Figure 7** : Site actif de l'ALK [47].

### II.4. Mécanisme d'action

La liaison du ligand au domaine extracellulaire active généralement les protéines tyrosine kinases du récepteur en induisant la dimérisation ou l'oligomérisation du récepteur. L'activation d'ALK induite par le ligand par trans-autophosphorylation va permettre la phosphorylation d'autres résidus situés dans la partie C-terminale de son domaine kinase. Certains de ces résidus peuvent alors servir de points d'ancrage, et permettent la liaison de protéines participant à la transduction du signal dépendant de l'ALK. Ces voies comprennent la voie Ras/Raf/MEK/ERK1/2, la voie JAK/STAT (Janus activated

kinase/signal transducer and activator of transcription), la voie PI3K (phosphatidylinositol 3-kinase) /Akt (PKB) et la voie PLC (phospholipase C). Les voies PLC- et Ras/ERK1/2 participent à la prolifération cellulaire, et les voies JAK/STAT et PI3K/Akt assurent la survie cellulaire (figure8) [48].



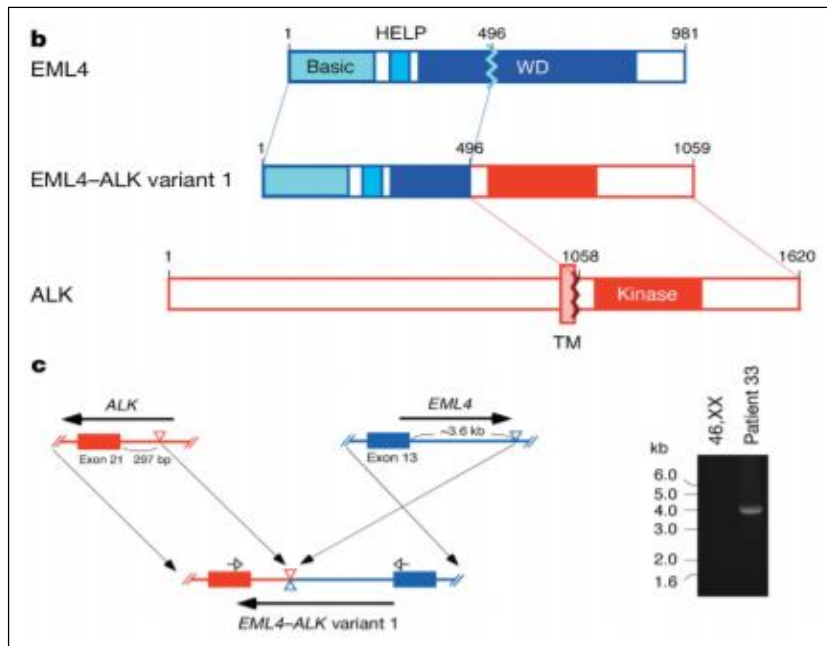
**Figure 8** : Voies de signalisation de l'ALK [48].

## II.5. Relation entre l'ALK et le cancer

L'activation des voies de signalisation oncogéniques par ALK dans les cellules cancéreuses peut se produire soit par l'activation des mutations ponctuelles ou l'amplification génique [46].

L'identification de la protéine de fusion EML4-ALK en 2007, par Soda et al [46]. Il a également été montré dans cette étude que la protéine de fusion EML4-ALK possédait des propriétés de transformation à la fois *in vitro* et *in vivo* [47].

La translocation se produit par une inversion paracentrique entre les gènes EML4 et ALK situés dans le bras court du chromosome 2. À ce jour, au moins 15 variantes EML4-ALK ont été identifiées, (Figure 9) [47]. Conduisant une activation de l'ALK et la phosphorylation des effecteurs en aval impliqués dans la cancérogène, qui provoque la multiplication non contrôlée des cellules cancéreuses [48].

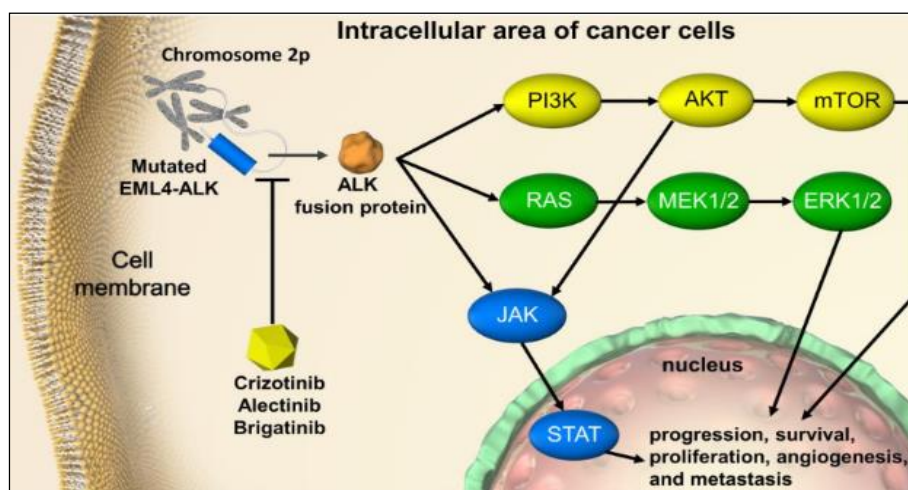


**Figure 9** : Représentation schématique de translocation entre EML4 et ALK [49].

La translocation génique la plus courante (EML4-ALK) qui améliore la phosphorylation de l'ALK, activant ainsi :

- Phosphatidylinositol-3 kinase-AKT (PI3K-AKT).
- Système d'activation de réseau (RAS).
- Le stimulateur de transcription (JAK / STAT) cascades des signaux (Figure 10).

Cela affecte non seulement les facteurs cancérogènes et la croissance tumorale, mais également les métastases dans d'autres organes, y compris les métastases du système nerveux central [49].



**Figure 10** : Les Voies de signalisation de EML4-ALK [49].

## II.6. les inhibiteurs d'ALK

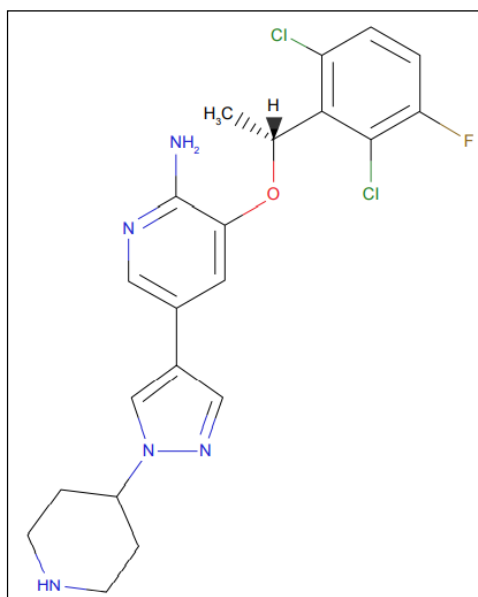
L'identification du remaniement EML4-ALK dans 5 % des CPNPC a permis le développement rapide d'inhibiteurs de ALK de premier génération et a modifier à la fois la prise en charge et le pronostic de ces patients.

Cependant, Dans les 12 à 16 mois qui suivent son introduction, il est constaté une résistance thérapeutique survenant soit par mutation ou amplification au niveau du gène ALK. De nouveaux inhibiteurs, d'ALK de 2<sup>ème</sup> et 3<sup>ème</sup> génération ont été développés [50].

### II.6.1 Les inhibiteurs de première génération

Le crizotinib, une petite molécule biodisponible par voie orale, est le premier ALK-TKI évalué et approuvé pour le traitement des patients atteints de CPNPC (Figure 11). Des études précliniques ont démontré que le crizotinib inhibe la prolifération cellulaire et favorise l'apoptose.

Le crizotinib inhibe également l'EMT et la ROS1, une kinase qui partage 77% des séquences d'acides aminés des sites de liaison de l'adénosine triphosphate (ATP) avec l'ALK [51].



**Figure 11** : la structure du Crizotinib[51]

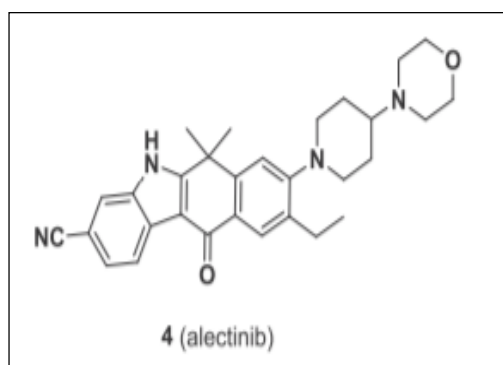


## II.6.2 Les inhibiteurs de deuxième génération

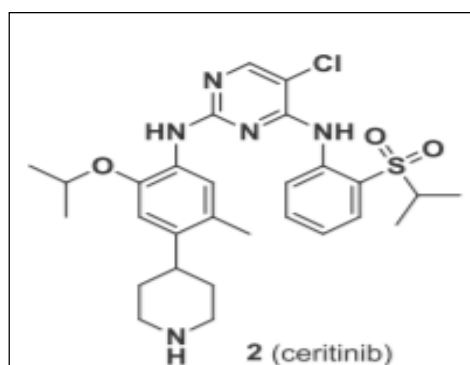
Contrairement au crizotinib, le ceritinib ne présente pas d'activité sur MET, ROS1. Son activité a été démontrée *in vitro* sur des modèles de cultures primaires issues de CBNPC résistants au crizotinib et porteurs des mutations L1196M ou G1269A [52].

Ces molécules ont montré leur efficacité sur les patients naïfs de traitement par inhibiteurs de ALK, avec notamment pour l'alectinib des taux de réponse de plus de 93 %. De même pour le céritinib (Figure 12,13), une activité anti-tumorale est clairement notée. Mais surtout, ces thérapeutiques restent efficaces chez des patients prétraités par un inhibiteur d'ALK avec des taux de réponse moyens de 55 %.

De plus, la pénétration de ces médicaments au niveau cérébral semble être meilleure que le crizotinib, et ainsi ils ont un effet anti-tumoral très intéressant sur les métastases cérébrales ; ceci sous-couvert d'un profil de toxicité relativement acceptable [50]



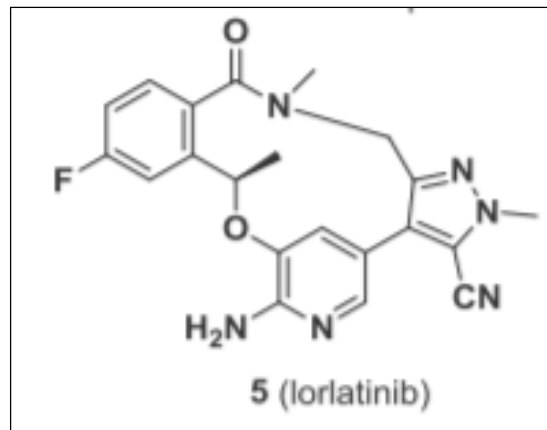
**Figure 12** : la structure de l'alectinib [56].



**Figure 13** : la structure de la céritinib[56].

## II.6.3 Les inhibiteurs de troisième génération

Lorlatinib est un inhibiteur d'ALK de troisième génération avec une efficacité et une sélectivité optimale, ainsi qu'une pénétration élevée vers le système nerveux central. Lorlatinib peut vaincre la résistance induite par plus de 16 mutations secondaires dans le domaine kinase d'ALK acquises lors de la progression aux ALK TKI de premières et deuxièmes générations. Le traitement par lorlatinib est donc efficace chez les patients préalablement traités par un ALK TKI de première ou deuxième génération, et est actuellement approuvé pour cette indication (Figure 14) [57].



**Figure 14** : la structure de Lorlatinib [56]

## ***Chapitre 3 : Docking Moléculaire***

## 1. Généralités

La modélisation moléculaire est un domaine scientifique multidisciplinaire ; elle exploite des lois de la chimie, de la physique et de la biologie dans des programmes informatiques spécifiques afin de calculer les affinités structurales et les propriétés des entités chimiques et biochimiques (Protéines, acides nucléiques, complexes moléculaires, solides, cristaux etc..). Le but est la compréhension ou la prédiction des phénomènes auxquels s'intéressent ces disciplines[58].

La conception d'un nouveau médicament suit un processus de recherche long et coûteux, principalement engagé par l'industrie pharmaceutique avec un support académique important.

le temps nécessaire à la découverte d'un nouveau médicament est estimé, en moyenne, entre 12 et 14 ans pour un coût total de 800 millions à 1 milliard de dollars[59].

Les méthodes de criblages *in silico* ont été développées pour diminuer le coût et le temps considérable nécessaires au processus de développement d'un nouveau médicament. En effet, Les méthodes de criblage virtuel sont désormais considérées comme étant une partie intégrante du processus en permettant de réduire le nombre considérable de molécules à tester expérimentalement [60].

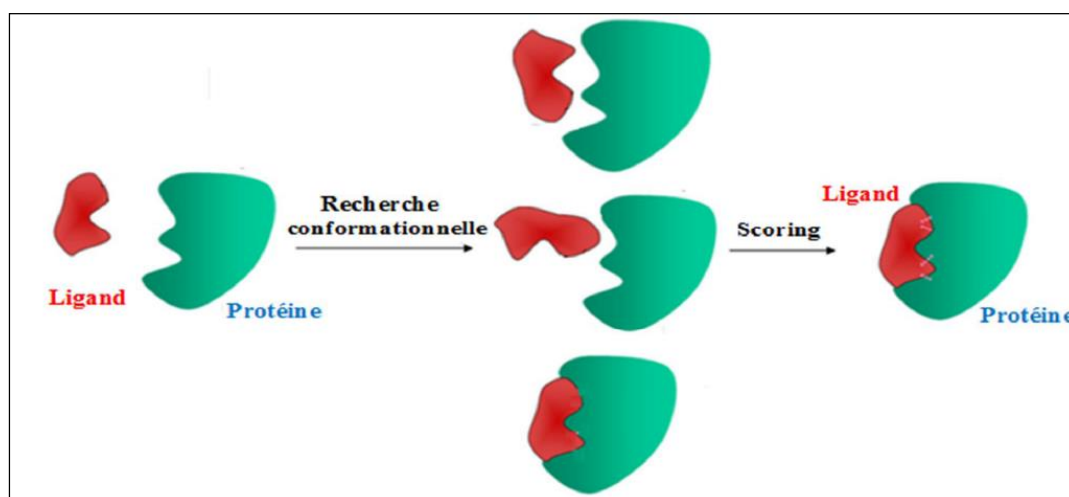
## 2. Définition de docking moléculaire

Le docking moléculaire est une méthode qui permet d'anticiper l'orientation favorable du ligand par rapport au récepteur (protéine) afin de créer un complexe stable. L'orientation préférée peut être utilisée pour prédire la force de la connexion ou l'affinité des liaisons entre le ligand et la protéine en utilisant des fonctions de scoring. Le docking est souvent appliqué pour anticiper l'orientation de la liaison des candidats médicaments contre les cibles protéiques afin de prédire l'affinité et l'activité du médicament, il joue donc un rôle central dans le processus de conception et de découverte de médicaments. L'objectif principal du docking moléculaire est de simuler par le calcul le processus d'identification moléculaire et d'obtenir une conformation optimisée[61].

### 3. Principe de docking

Une simulation de docking comprend essentiellement deux étapes :

- La première étape dite de **Docking**, consiste à positionner le ligand et lui adopter plusieurs conformations et plusieurs positions possibles au niveau du site actif de la protéine, en ne retenant que celle qui représentent les modes d'interactions les plus favorables.
- La deuxième étape dite de **Scoring** consiste à évaluer l'affinité du ligand à l'égard de la protéine et de donner un score aux poses obtenues lors de la phase de docking. Ce score permettra de retenir la meilleure pose parmi toutes celles proposées [62].



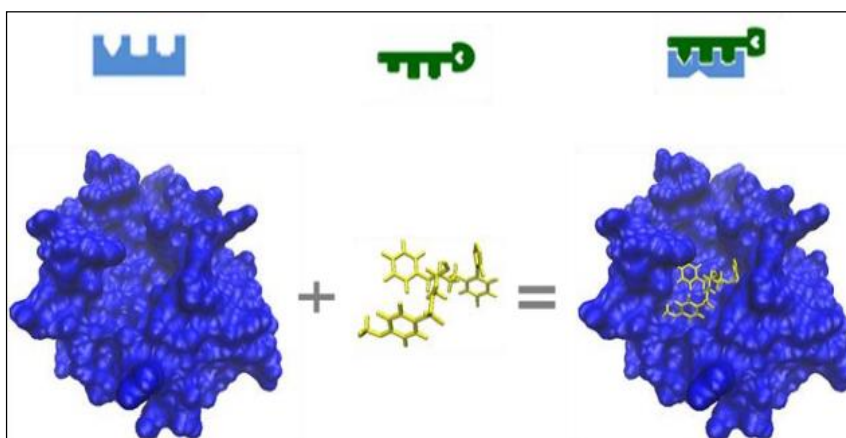
**Figure 15** : Principe générale d'un programme de docking [63].

### 4. Types de docking

Les méthodes de docking peuvent être classées en trois catégories selon la prise en compte de la flexibilité du ligand.

- ❖ **Docking rigide** : C'est le docking le plus simple et le plus rapide, il est généralement employé pour l'amarrage entre deux macromolécules (docking protéine-protéine ou protéine/acide nucléique structuré). De plus, l'amarrage rigide ne nécessite pas de calculs multiples[64] ( Figure16).

- ❖ **Docking semi-flexible** : Pendant le processus d'amarrage semi-flexible, la conformation du récepteur fixe reste inchangée, tandis que la conformation du ligand est altérée. L'amarrage semi-flexible convient à l'amarrage de petites molécules et macromolécules, telles que les protéines ou les acides nucléiques et les petites molécules de ligand [64].
- ❖ **Docking flexible** : L'amarrage flexible est généralement utilisé pour étudier avec précision l'interaction entre les molécules. Il permet de modifier facilement la conformation du système d'amarrage (récepteur et ligand). le processus de ce type est trop compliqué [64].



**Figure 16** : Modèle clé-serrure [65].

## 5. Les outils de docking moléculaire

Les outils qui sont très utilisés en docking sont : une petite molécule qui est appelée ligand, une cible biologique d'intérêt thérapeutique appelée protéine et un programme de docking.

### 5.1. Les Récepteurs 3D (cible)

Pour l'obtention des récepteurs, on utilise les structures 3D des protéines qui sont accessibles gratuitement dans la banque de donnée PDB « Protein Data Bank ». Il s'agit d'une grande archive de données structurales de macromolécules biologiques : protéines et acides nucléiques (ADN, ARN) [66]. En général, les sources d'information de ces structures proviennent majoritairement de la cristallographie aux rayons X et de la RMN. En juin 2021, plus de 156 944 structures protéiques ont été résolues et injectées dans la

PDB grâce à la cristallographie aux rayons X. La RMN fournit, quant à elle, environ 13 408 structures, contre seulement 7806 dans le cas de la microscopie électronique [67].



**Figure 17** : Interface d'accueil du PDB.

Les coordonnées atomiques des structures se trouvent dans un format de référence, appelé pdb. Lu par différents logiciels de visualisation, de modélisation, de criblage virtuel...etc.

Le fichier pdb contient diverses informations sur la macromolécule telle que le nom du récepteur, l'équipe qui a résolu la structure, la méthode expérimentale, ...etc. On y retrouve aussi des informations sur la structure primaire, les hétéro-atomes (ligand, métaux, résidus modifiés, ...etc.), la structure secondaire et les coordonnées atomiques X, Y et Z qui déterminent la position exacte de chaque atome dans une conformation donnée. Enfin, la dernière section du fichier fournit des informations sur les connections chimiques entre les atomes [68].

## 5.2 Ligands 3D (Chimiothèques)

Pour un docking moléculaire, le ligand doit être également sous forme 3D. Pour y obtenir, il existe deux moyens :

Le premier moyen consiste à utiliser des programmes de construction moléculaire 3D tels que ChemDraw, Arguslab, Titan, Marvin, Sybyl...etc. Chacun de ces programmes contient une banque d'atomes permettant de dessiner le ligand tout en tenant compte de l'état d'hybridation de chaque atome et du type de liaisons simple, double ou triples. Ensuite, la géométrie du ligand doit être optimisée à l'aide d'un algorithme semi-empirique [69].

Le second moyen est souvent d'aspect commercial consistant à consulter des bases de données de structures chimiques appelées chimiothèques [68]. Parmi les chimiothèques virtuelles nous citons :

- ❖ **PubChem** : 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 Institutes of Health (NCI). Actuellement, la PubChem répertorie plus de 109.920.523 composés en mettant en ligne, gratuitement. Cette chimiothèque est consultée gratuitement via une interface utilisateur Web : <http://pubchem.ncbi.nlm.nih.gov> [70].



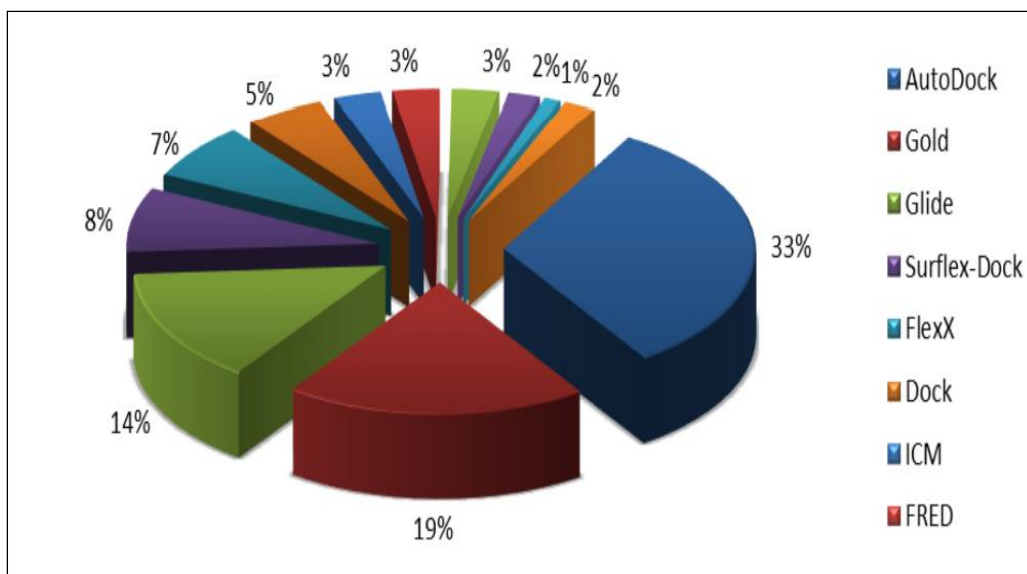
**Figure 18** : Interface d'accueil du PubChem.

- ❖ **Chimiothèque nationale Française** : cette chimiothèque a été créée en 2003. Elle fédère les collections de produits de synthèse et d'extraits naturels existants dans les laboratoires publics français et en assure la valorisation scientifique et industrielle. En 2021, la chimiothèque nationale française regroupe plus de 76.276 molécules. Cette chimiothèque est consultée gratuitement via le lien : <http://chimiotheque-nationale.cn.cnrs.fr/> [71].

### 5.3 Programmes (Logiciels)

Dans le domaine de Docking moléculaire, plusieurs logiciels ont été utilisés pour étudier les différentes interactions et les conformations les plus favorables du ligand au sein de son récepteur. Actuellement, un grand nombre de logiciels de Docking moléculaire sont disponibles tels que: GOLD, AUTODOCK, AUTODOCK VINA, FLEXX, SURFLEX [72]. (Figure 17)





**Figure 19:** Utilisation des différents logiciels de docking en 2010-2011 [68].

### 5.3.1 AutoDock

AutoDock utilisé comme optimiseur global et la minimisation d'énergie comme méthode de recherche locale. Les orientations possibles sont évaluées par le modèle AMBRE de champ de force en conjonction avec les fonctions de score d'énergie libre et un grand ensemble de complexes de protéine-ligand avec des constantes connues.

### 5.3.2 Dock

Le dock est l'un des logiciels de docking ligand-protéine le plus ancien et le plus connu. La version initiale utilise des ligands rigides, la flexibilité a été incorporée plus tard par l'intermédiaire de la construction par incrémentation du ligand dans la poche de liaison.

### 5.3.3 FlexX

FlexX est une autre méthode basée sur des fragments en utilisant des ligands flexibles et des protéines rigides [72].

**Tableau 1 :** Les programmes de Docking moléculaire [73].

Programme	Editeur	Site internet
Auto Dock	Scripps	<a href="http://www.scripps.edu/mb/olson/doc/autodock/">http://www.scripps.edu/mb/olson/doc/autodock/</a>
Dock	UCSF	<a href="http://dock.compbio.ucsf.edu/">http://dock.compbio.ucsf.edu/</a>

<b>FlexX</b>	BioSolveIT	<a href="http://www.biosolveit.de/FlexX/">http://www.biosolveit.de/FlexX/</a>
<b>Fred</b>	Open Eyes	<a href="http://www.eyesopen.com/products/applications/fred.html">http://www.eyesopen.com/products/applications/fred.html</a>
<b>Glide</b>	Schrödinger	<a href="http://www.schrodinger.com/Products/glide.html">http://www.schrodinger.com/Products/glide.html</a>
<b>Gold</b>	CCDC	<a href="http://www.ccdc.cam.ac.uk/products/life_sciences/gold/">http://www.ccdc.cam.ac.uk/products/life_sciences/gold/</a>
<b>Ligand Fit</b>	Accelrys	<a href="http://www.accelrys.com/cerius2/c2ligandfit.html">http://www.accelrys.com/cerius2/c2ligandfit.html</a>
<b>Surflex</b>	Biopharmics	<a href="http://www.biopharmics.com/products.html">http://www.biopharmics.com/products.html</a>

## *Chapitre 4 : Materials and Methods*

## 1. Softwares and tools used

### 1.1 Surflex

Surflex is one of the most successful docking programs. It is widely used and able to dock ligands in an environment consisting of amino acids with good precision [74].

### 1.2 ChemDraw 3D

ChemDraw 3D has allowed us the molecular construction of different inhibitors from the literature, many formats are available such as Pdb, Mol, Mol2, etc.

### 1.3 Open Babel

We have used this program to convert the *.pdb* format of the protein and ligand in *.mol2* format and also to add hydrogen atoms to the structure of the protein and the ligand.

### 1.4 BIOVIA Discovery Studio

Biovia is a free tool for viewing, which allows a 3D view of a structure of biological molecule. We have used this program for the visualization of different interactions formed between the active site of ALK and its inhibitors.

### 1.5 Excel Microsoft 2016

Excel Microsoft 2016 was used in the present work to plot the linear correlation curve in the reliability test of the Surflex.

## 2. PDB (Protein Data Bank)

The 3D structure of the ALK kinase, which is the subject of this work, comes from the PDB. The latter is a repository for the 3-D structural data of large biological molecules, such as proteins and nucleic acids. The data, typically obtained by X-ray crystallography or NMR spectroscopy and submitted by biologists and biochemists from around the world, can be accessed at no charge over the internet [75].

## 3. Programs evaluation

### 3.1. The root mean square deviation (RMSD)

It is the ability of a program to reproduce the experimental binding modes of a ligand. For this, a ligand in its structure obtained by X-ray crystallography is docked in its site of binding and then the binding mode docked is compared with that obtained experimentally

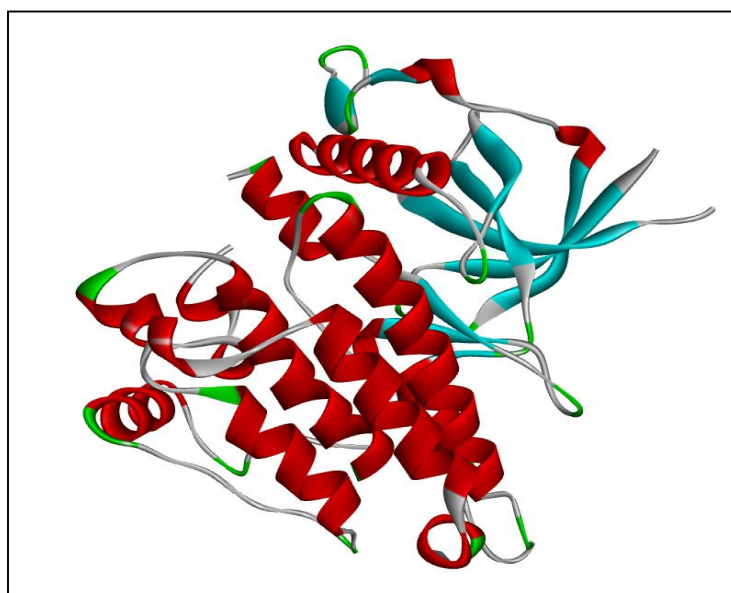
and a RMSD between the two is calculated. A good RMSD is that less than or equal to 2 Å [76]. In the present work this test was performed on 100 complexes (protein-ligand) randomly drawn from the PDB.

### 3.2 Correlation coefficient (R)

The correlation coefficient (R) indicated the degree of linear relationship between two data sets with values between -1 and 1. If there was no linear relationship between the two sets of data, the coefficient correlation was very close to zero: In that case we considered the two variables not being correlated [74]. This test was performed with 30 ALK inhibitors from the literature. This test allows the evaluation of the degree of correlation between the scores simulated by Surflex and the experimentally determined  $IC_{50}$  values of the studied inhibitors.

## 4. Preparation of molecules for molecular docking

The protein-ligand complex is downloaded in the *.pdb* format from the Protein Data Bank (PDB) by introducing its code ID (5IUI). We have eliminated all the molecules in required of the structure of protein (the water molecules, metal ions, the different ligands of molecules, the information of cristallographies). The Surflex software requires the format *.mol2*. Therefore the two molecules (protein, ligand) are processed in the format *.mol2* via Open Babel program.



**Figure 20:** structure of the complex 5IUI.

## 5. Calculation of interactions “enzyme- similar”

To study the interactions between the anaplastic lymphoma kinase (ALK) and the similar we use a specific PubChem library.

### 5.1 PubChem

Pubchem is an American database of chemical molecules managed by the National Center for Biotechnology Information (NCBI), a branch of the National Medical Library of the United States under the authority of the National Institutes of Health (NIH). PubChem currently lists several million compounds whose structure and physicochemical properties are accessible free of charge via the internet (<http://pubchem.ncbi.nlm.nih.gov>) [77]. This bank was used to obtain the structural analogues of our ligand 45Q to a percentage of similarity 89% in order to dock them by our software and to search among these compounds one that binds more strongly than the 45Q and propose it as a new, more effective inhibitor.

## 6. Lipinski rule

Lipinski's Rule of Five is a rule of thumb to determine if a chemical compound with a certain pharmacological or biological activities has properties that would make it a likely orally active drug in humans.

The rule made by Christopher Lipinski. Each drug must comply with several basic criteria, such as low cost of production, be soluble, stable, but must also conform to the schedules associated with its pharmacological properties of absorption, distribution, metabolism, excretion and toxicity(ADME/Tox).

Lipinski's Rule of Five states that an orally active drug must complete three of these five proprieties:

- No more than 5 hydrogen bond donors.
- No more than 10 hydrogen bond acceptors.
- No more than 15 rotatable bonds.
- The Molecular weight is over 500 g/mol<sup>-1</sup>.
- The Log P is over 5 [78].

## ***Chapitre 5 : Results and discussion***

With the aim of developing new inhibitors of the ALK protein (anaplastic lymphoma kinase) involved in lung cancer disease, we have undertaken in this part a study of the modeling by docking of the structures of this enzyme.

Using the molecular docking technique, we investigated the most likely mode of interaction of the 45Q inhibitor (from PDB) in the active site of the ALK enzyme.

## 1. Reliability of the programs used

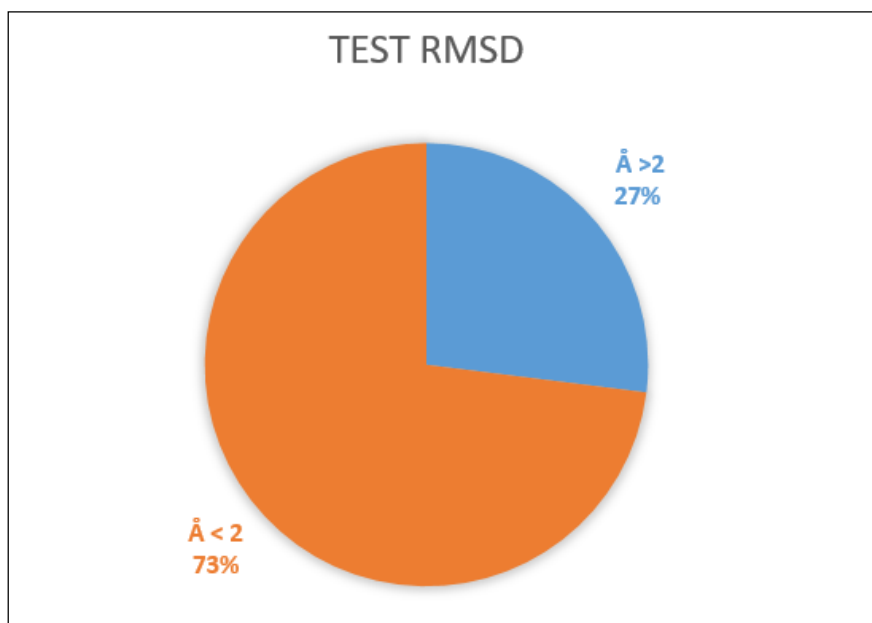
The performance of the Surflex mooring program is judged by three criteria:

- The mean square deviation (RMSD).
- Visual analysis.
- The correlation coefficient ( $r$ ).

### 1.1. The RMSD

The RMSD reliability test of the Surflex program was performed using 100 protein-ligand complexes taken at random from PDB. The complexes are downloaded in *pdb* format after entering the code corresponding to each complex into the PDB. The prediction is acceptable if the RMSD values do not exceed 2Å.

The results obtained from this test are illustrated in (Figures 3).



**Figure 21:** Results as a percentage of the RMSD test at two intervals.



The figure 3 shows that 73% of RMSD values are less than or equal to 2 Å and that only 27% of results are higher than 2 Å. The RMSD values were consistent with the results obtained by Chikhi and Bensegueni, (2008)[79], Teniou (2012)[80], and Boucherit et al., (2014)[81], who have shown that the docking program Surflex produces reasonable binding modes.

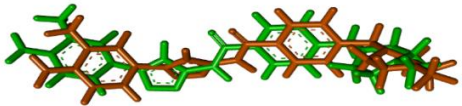
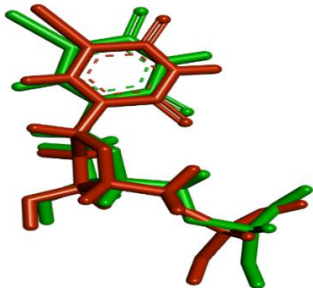
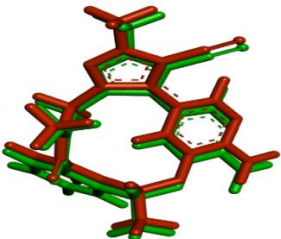
### 1.2. Visual analysis

The visual analysis following the RMSD test is an essential and complementary step to judge the performance of the Surflex program. It allows verification of the superposition of the experimental ligand with respect to that docked.

The visualization of the results of molecular docking by the Biovia Discovery software is carried out with 3 complexes with different RMSD results (table 1).

The comparison of the experimental conformations of the ligand and the optimal conformations calculated by the Surflex software of the 3 complexes is shown in figure 4.

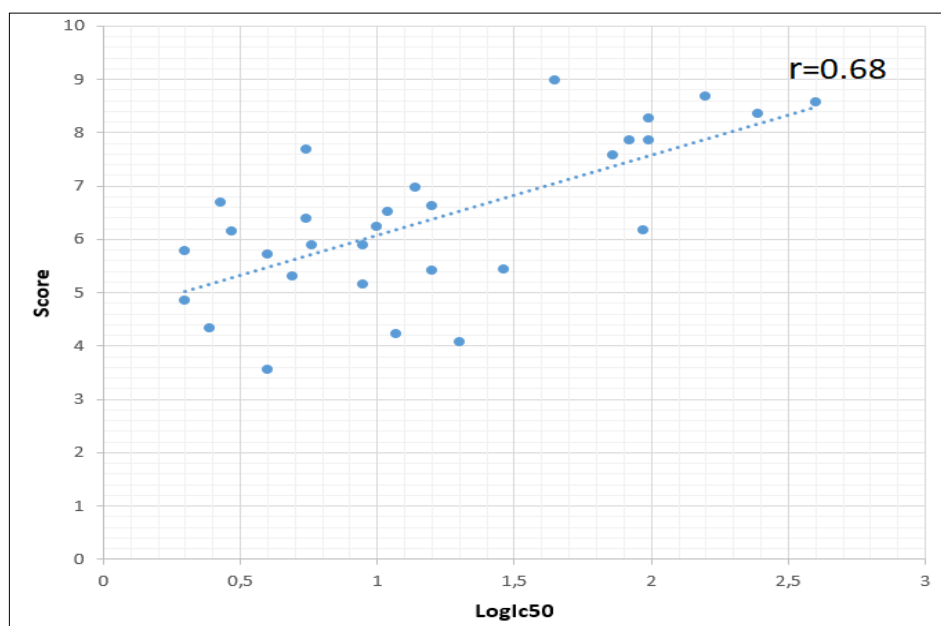
**Table 2:** RMSD values of the 3 complexes studied and superpositions of the **reference** ligands given by PDB (colored in green) and the final-0 calculated by Surflex (colored in brown).

PDB Code	Ligands PDB	Rmsd (Å)	Figures
5IUI	45Q	1.465	
2QCH	5IU	0.160	
4UXL	5P8	0.625	

In light of the results obtained with the RMSD test followed by visual analysis, the Surfex program is performing well enough since it reproduces the experimental results fairly well. We can therefore use this program, without too much risk of error, to simulate the formation of ALK complexes.

### 1.3 The correlation coefficient

The correlation coefficient indicates the degree of linear relationship between the affinity obtained by molecular docking with surflex and  $IC_{50}$  determined experimentally. In this study, PDB ALK inhibitors were examined. A total of 30 complexes (ALK inhibitors) were tested by the Surflex software.



**Figure 22:** Linear correlation curve between Log  $IC_{50}$  of the 30 inhibitors and their scores.

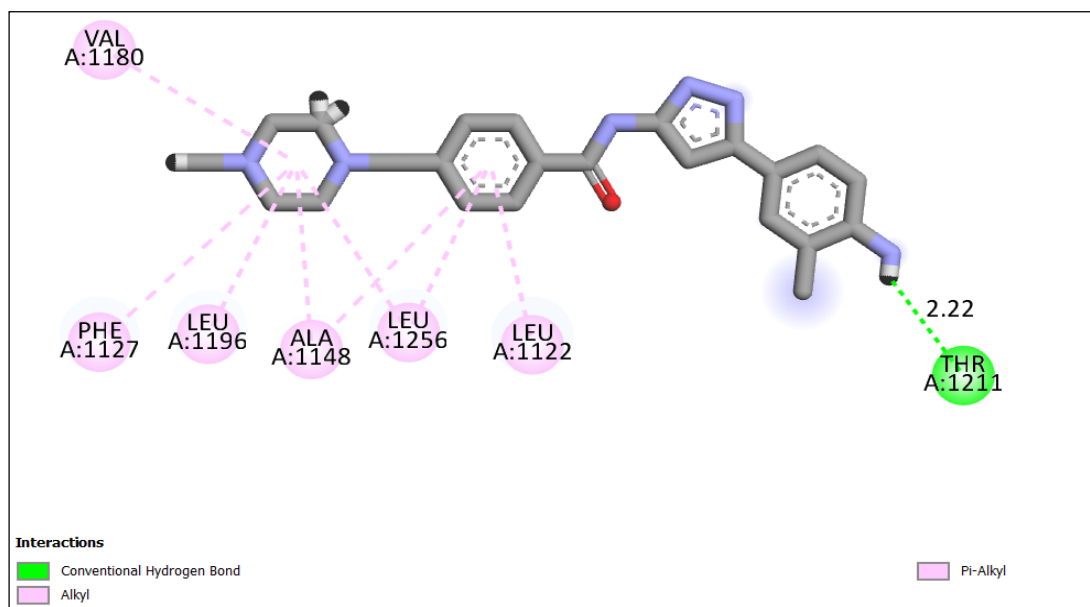
Linear regression analysis represented in figure 5 shows that there is a strong correlation between the affinity scores and the Log  $IC_{50}$  values of the inhibitors studied. A good correlation was observed with  $r=0.68$ , which is in agreement with the results of Chikhi A and Bensegueni A., (2010)[82], Kamel et al., (2010)[83], and Boucherit et al., (2014)[81]. It demonstrated a good linear relationship and the trend could indicate that the docking program produced reasonable binding modes. It could be used to predict enzyme-inhibitors interactions.

## 2. Study of the interactions involved in the inhibition of the human ALK

### 2.1 Interaction 5IUI-45Q

Surflex program was used to perform molecular docking aiming to analyze the interaction modes and the orientations of various inhibitors within ALK active site. 5IUI being the complex with the best resolution (1.88 Å) available in PDB, its co-crystallized

inhibitor 45Q was chosen as reference ligand. Our approach is to study the interaction of the kinase ALK with the reference ligand 45Q. The result is shown in figure 6:



**Figure 23:** Representation of interactions formed by 45Q with the active site of 5IUI.

The figure 6 shows the active site of the kinase ALK complexed with the inhibitor 45Q, the inhibitor is represented in “stick”. Visualization of results of the docking shows that the inhibitor 45Q form with the active site of the kinase ALK one hydrogen bond with Thr1211 of the active site. The hydrogen bond is observed between the NH group of the inhibitor 45Q and the carbonyl of Thr1211 (N-H .....O Thr1211) (distance=2.22 Å). 45Q is also stabilized by hydrophobic interactions in the binding pocket, specifically with Val1180, Phe1127, Leu1196, Ala1148, Leu1256, Leu1122 amino-acids residues.

### 3. Proposed new inhibitors of human ALK

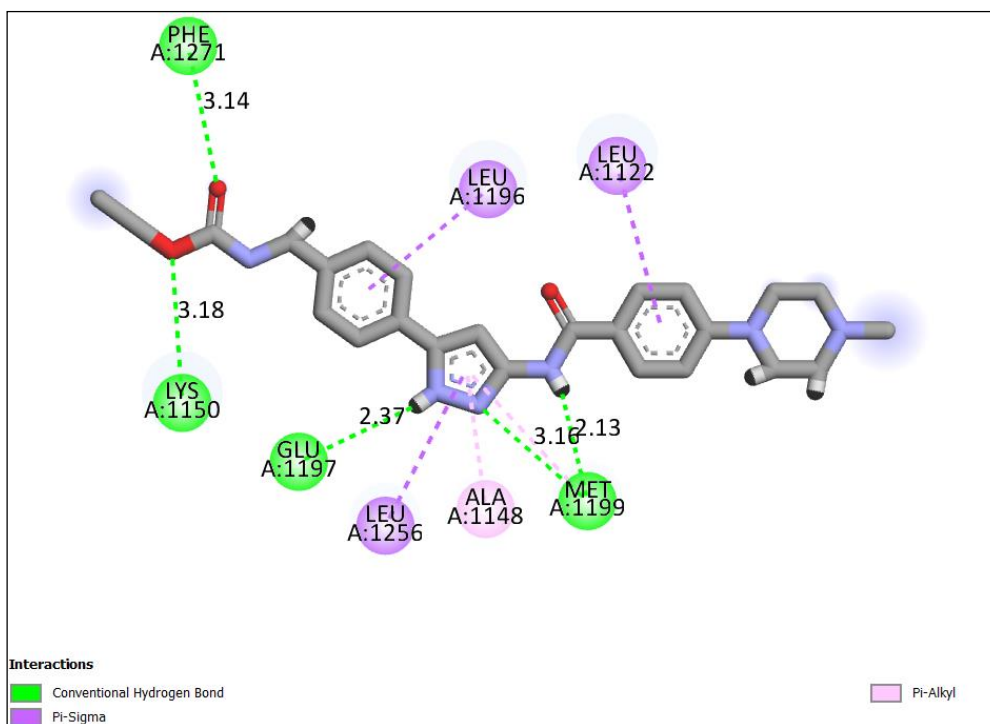
#### 3.1 Inhibition of the human ALK by the similar of compound 45Q

In the aim of finding new more potent inhibitors of human ALK, we performed the molecular docking of a collection of 449 molecules similar of the 45Q with 89% similarity. These chemical compounds are downloaded from the PubChem database.

**Table 3:** Docking results of the reference ligand and the two similar compounds with the highest affinity.

Inhibitors	Affinity ( $M^{-1}$ )
45Q ( <b>Reference</b> )	6.88
CID_ 70654694	9.16
CID_70655002	9.70

The visual analysis shows that the similar (CID\_ 70654694) establishes five hydrogen bonds as shown in figure 7.

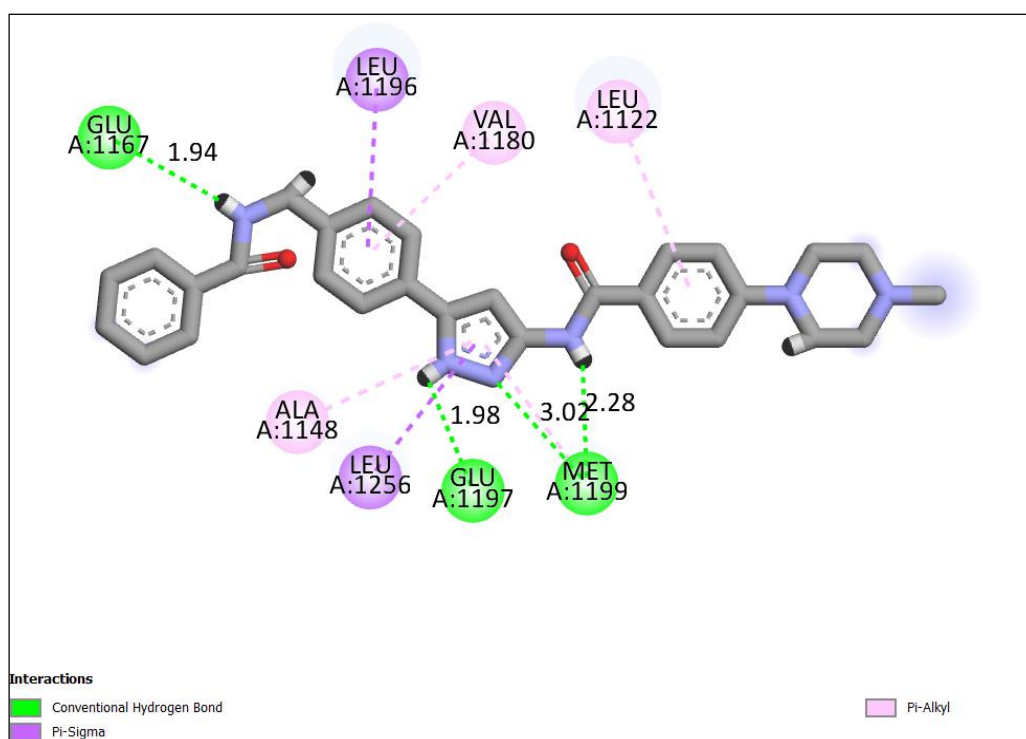


**Figure 24 :** Representation of interactions formed by the similar CID\_ 70654694.

- Two hydrogen bonds are observed , between the NH group of CID\_ 70654694 and the carbonyl function of Met1199 ( N-H .....O Met1199) ( distance =2.13 Å), and between the nitrogen atom of CID\_ 70654694 and nitrogen atom of the Met1199 residue (N.....N Met1199) ( distance=3.16 Å);
- A hydrogen bond is formed between the NH group of CID\_ 70654694 and the carbonyl function of Glu1197 ( N-H .....O Glu1197) ( distance=2.37 Å) ;

- A hydrogen bond is formed between the oxygen atom of CID\_ 70654694 and the nitrogen atom of the Lys1150 residue (O .....N Lys1150) (distance=3.18 Å);
  - The last hydrogen bond is observed between the carbonyl function of CID\_ 70654694 and the nitrogen atom of the Phe1271 (O.....N Phe1271) (distance=3.14 Å).
- In addition , this similar CID\_ 70654694 establishes several hydrophobic interactions with the residues Ala1148 , Leu1256 , Leu1196 , Leu1122.

The similar (CID\_70655002) establishes four hydrogen bonds as shown in figure 8.



**Figure 25:** Representation of interactions formed by the similar CID\_70655002.

- Two hydrogen bonds are observed , between the NH group of CID\_70655002 and the carbonyl function of Met1199 ( N-H .....O Met1199) ( distance =2.28 Å) , and between the nitrogen atom of CID\_ 70654694 and nitrogen atom of the Met1199 residue (N.....N Met1199) ( distance=3.02 Å);
  - A hydrogen bond is formed between the NH group of CID\_70655002 and the carbonyl function of Glu1197 (N-H .....O Glu1197) (distance=1.98 Å);
  - The last hydrogen bond is observed between the NH group of CID\_70655002 and the carboxyl function of Glu1167 (N-H .....O Glu1167) (distance=1.94 Å);
- Several hydrophobic interactions stabilizing the similar CID\_70655002 are carried out

with the residues Leu1256 , Leu1196 , Ala1148 , Val1180 , Leu1122 .

#### 4. Drug-likeness Prediction

In the subsequent part of our work, we verified the pharmacokinetic and toxicity proprieties of the most promising products (CID\_ 70654694 and CID\_ 70655002) using SuissADME and PreADMET . These properties consist of their blood-brain barrier (BBB) penetration, gastrointestinal absorption (GI), Cytochrome P450 (CYP) inhibition, Lipinski's rule of 5 catagorization, and toxicity (table 2). The same parameters of 45Q were also studied for comparison.

**Table 4:** The predicted pharmacokinetic and toxicity proprieties of the most promising compounds.

Properties	45Q	CID_ 70654694	CID_ 70655002
Lipinski's rule of 5	Suitable	Suitable	Suitable
Synthetic accessibility	2.82	3.52	3.52
BBB permeant	Yes	No	No
GI absorption	High	High	High
CYP inhibitor	CYP1A2 inhibition CYP2D6 inhibition CYP3A4 inhibition	CYP2C19 inhibition CYP2C9 inhibition CYP2D6 inhibition CYP3A4 inhibition	CYP1A2 inhibition CYP2C19 inhibition CYP2C9 inhibition CYP2D6 inhibition CYP3A4 inhibition
Toxicity	hERG inhibition (medium risk) carcinogenicity in mice ( positive)	hERG inhibition (medium risk) carcinogenicity in mice ( positive)	hERG inhibition (medium risk) carcinogenicity in mice ( positive)

As indicated in Table 3, both CID\_ 70654694 and CID\_ 70655002 follow the criteria for orally available drugs with no Lipinski's rule of 5 violation, the synthetic accessibility of these compounds seems to be possible with values below 5. Moreover, the studied products have no BBB penetration, this may protect the brain from a probable harmful effect of these compounds, contrary to 45Q which has a high BBB penetration, However, these 2 compounds have shown a high GI absorption which makes it possible to access the blood. The results also show that the compound CID\_ 70654694 and CID\_ 70655002 have an inhibitory effect on the different iso-forms of Cytochrome P450 (essential enzymes for the metabolism of numerous medicaments in the liver). Furthermore, their toxicity problems (including carcinogenicity in mice and hERG inhibition) can be improved during their further optimization to obtain clinically useful drugs.



## *Conclusion et perspectives*

## Conclusion et perspectives

Notre travail de recherche avait pour but essentiel d'acquies des compétences en simulation informatique, en particulier le docking moléculaire par Surflex afin de contribuer au développement *in silico* de nouveaux inhibiteurs de la kinase ALK, cette dernière est une cible thérapeutique potentielle du cancer du poumon non à petites cellules.

Dans un premier temps, nous avons testé la performance du programme Surflex utilisé dans cette étude, par le test RMSD (*root-mean-square deviation*) sur 100 complexes pris aléatoirement de la PDB, 73% ont présenté un RMSD inférieure à 2 Å. L'analyse visuelle faisant suite à ce test nous a permis d'avoir une meilleure appréciation des valeurs numériques de RMSD. Avec un coefficient de corrélation égal à 0.68, des valeurs de RMSD inférieures à 2 Å et une bonne superposition des ligands simulés et expérimentaux, ces tests nous ont permis de conclure que Surflex est un programme de docking moléculaire très fiable et hautement performant qu'on peut utiliser en toute confiance pour l'étude *in silico* de l'inhibition de la kinase ALK.

Dans un deuxième temps, nous avons étudié l'inhibition de la kinase ALK par diverses molécules provenant de la littérature et aussi de la PDB, on a aussi visualisé les liaisons qu'elles impliquent avec le site actif de cette enzyme, les résultats ont montré que le ligand issu de la PDB portant le code 45Q (Affinité=6.88 M<sup>-1</sup>, I<sub>c50</sub>= 10000 nM) possède le meilleur effet inhibiteur de la kinase ALK. En prenant ce composé comme structure de départ, la PubChem nous a proposé 449 composés similaires avec un taux de similarité allant jusqu'à 89%, Le docking moléculaire de ces composés similaires envers le site actif de la kinase ALK fait ressortir les composés CID\_70655002 et CID\_70654694 comme nouveaux inhibiteurs potentiellement puissants avec des affinités égale à 9.70 M<sup>-1</sup>, 9.16 M<sup>-1</sup> respectivement.

Enfin, l'étude *in silico* portant sur l'évaluation des différentes propriétés physicochimiques et pharmacocinétiques a montré que les composés CID\_70655002 et CID\_70654694 sont, en général, conformes aux critères imposés par la règle de 5 de Lipinski.

Comme perspectives d'avenir, des tests *in vitro* et/ou *in vivo* de l'activité biologique des composés CID\_70655002 et CID\_70654694 sont recommandés pour compléter les résultats théoriques et vérifier l'efficacité de l'approche *in silico*.

## *Références bibliographiques*

## Références bibliographiques

- [1] D. Duhovny, R. Nussinov, and H. J. Wolfson, "Efficient unbound docking of rigid molecules," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2452, pp. 185–200, 2002, doi: 10.1007/3-540-45784-4\_14.
- [2] G. Coureau and F. Delva, "Lung cancer screening among the smoker population," *Bull. Cancer*, vol. 106, no. 7–8, pp. 693–702, 2019, doi: 10.1016/j.bulcan.2018.12.010.
- [3] T. L. D. E. G. MD, "ALK alterations and inhibition in lung cancer," *Semin. Cancer Biol.*, vol. 42, pp. 81–88, 2017, doi: 10.1016/j.semcancer.2016.08.007.
- [4] E. Moerschel, "Formation Physiopathologie en ST2S," pp. 1–15, 2010.
- [5] I. Mimeault and V. Lord, "Qu' est-ce que le cancer ? Une maladie de notre civilisation Pour en savoir plus Santé des femmes," 2011.
- [6] a Biyi, N. Benrais, A. Albouzidi, and S. Benomar, "Mécanisme de la carcinogenèse à la lumière des données de la biologie moléculaire," *Médecine du Maghreb*, vol. 69, pp. 1–7, 1998.
- [7] L. Pecorino, *Molecular Biology of Cancer: Mechanisms, Targets, and Therapeutics*, 3rd ed. 2012.
- [8] "What Is Cancer? - National Cancer Institute."
- [9] M. Riquet, "Anatomie du poumon humain," *EMC - Pneumol.*, vol. 9, no. 3, pp. 1–11, 2012.
- [10] D. Youcef, Z. Cherif, and H. Cherif, "Application en Téléradiologie dédiée à la détection Du Cancer des poumons," 2018.
- [11] A. Waugh and A. Grant, *Anatomie et physiologie normales et pathologiques*, 12th ed. 62, rue Camille-Desmoulins, 92442 Issy-les-Moulineaux Cedex, 2012.
- [12] J. E. Larsen and J. D. Minna, "Molecular Biology of Lung Cancer: Clinical Implications," *Clin. Chest Med.*, vol. 32, no. 4, pp. 703–740, 2011.
- [13] W. A. Cooper, D. C. L. Lam, S. A. O'Toole, and J. D. Minna, "Molecular biology of lung cancer," *J. Thorac. Dis.*, vol. 5, no. SUPPL.5, 2013.

- [14] E. Kaznowska *et al.*, “The classification of lung cancers and their degree of malignancy by FTIR, PCA-LDA analysis, and a physics-based computational model,” *Talanta*, vol. 186, pp. 337–345, 2018.
- [15] J. A. Barta, C. A. Powell, and J. P. Wisnivesky, “Global epidemiology of lung cancer,” *Ann. Glob. Heal.*, vol. 85, no. 1, pp. 1–16, 2019.
- [16] H. Sung *et al.*, “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA. Cancer J. Clin.*, vol. 0, no. 0, pp. 1–41, 2021.
- [17] Globocan, “cancers cases according to WHO,” vol. 536, pp. 2020–2021, 2021.
- [18] Ziamni IdirZiane Nabil, “Dédicaces :,” *Essai d’estimation des coûts directs médicaux du cancer du poumon dû au tabagisme*, p. 21, 2016.
- [19] Baptiste GURY, “These:enqijete s[jr le tabagisme des footballeurs professionne1,s nu grand est de la france (auxerre, dijon, metz, nancy, reims, sedan, sochaux, strasbourg, troyes),” pp. 68–70, 2008.
- [20] R. Solidarit and D. Sant, “Ministère des Solidarités et de la Santé de la franceÉdition 2018,” 2018.
- [21] la Commission canadienne de sûreté nucléaire (CCSN), “Le radon et la santé,” p. 20, 2011.
- [22] E. Maxence, “Les causes du cancer,” p. 21, 2012.
- [23] S.-G. Prof. dr méd. Martin Brutsche and K. Dr méd. Markus Heitz, “Le cancer du poumon,” *Ligue Pulm. suisse*, vol. 19.
- [24] C. P. THOMAS, “Les cancers du poumon,” *Rev. Med. Liege*, vol. 3, no. 9, pp. 205–214, 1948, [Online]. Available: [www.ligue-cancer.net](http://www.ligue-cancer.net).
- [25] S. Chaleat and L. Greillier, “Cancer du poumon,” *Rev. des Mal. Respir. Actual.*, p. 28, 2019, doi: 10.1016/s1877-1203(19)30161-2.
- [26] G. Ferretti, A. Jankowski, A. Calizzano, D. Moro-Sibilot, and J. P. Vuillez, “Imaging and PET/CT of lung cancer,” *Journal de Radiologie*. Elsevier Masson SAS, pp. 387–402, 2008, doi: 10.1016/s0221-0363(08)89016-6.

- [27] A. Khalil, M. Majlath, V. Gounant, A. Hess, J. P. Laissy, and M. P. Debray, “Place de l’IRM dans le cancer bronchopulmonaire,” *J. Radiol. Diagnostique Interv.*, p. 424, 2016, doi: 10.1016/j.jradio.2016.09.002.
- [28] M. Wissler, “THESE: Bilan de l’analyse du statut mutationnel EGFR de 1000 patients atteints d’adenocarcinomes pulmonaires pris en charge par la plateforme d’oncologie moléculaire du CHU-CAV de NANCY,” 2012.
- [29] T. Saidi *et al.*, “Non-invasive prediction of lung cancer histological types through exhaled breath analysis by UV-irradiated electronic nose and GC/QTOF/MS,” *Sensors Actuators, B Chem.*, vol. 311, no. September 2019, 2020.
- [30] H. Hoy, T. Lynch, and M. Beck, “Surgical Treatment of Lung Cancer,” *Crit. Care Nurs. Clin. North Am.*, vol. 31, no. 3, pp. 303–313, 2019.
- [31] Y. Watanabe, “TNM classification for lung cancer,” *Ann. Thorac. Cardiovasc. Surg.*, vol. 9, no. 6, pp. 343–350, 2003.
- [32] S. Hammerschmidt and H. Wirtz, “Lungenkarzinom - Aktuelle diagnostik und therapie,” *Dtsch. Arztebl.*, vol. 106, no. 49, pp. 809–820, 2009.
- [33] C. Zappa and S. A. Mousa, “Non-small cell lung cancer: Current treatment and future advances,” *Transl. Lung Cancer Res.*, vol. 5, no. 3, pp. 288–300, 2016.
- [34] Didier Vander Steichel, “Les cancers du poumon,” 2009.
- [35] J. R. Molina, P. Yang, S. D. Cassivi, S. E. Schild, and A. A. Adjei, “Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship,” *Mayo Clin. Proc.*, vol. 83, no. 5, pp. 584–594, 2008.
- [36] Institut National du Cancer, “Les traitements des cancers du poumon - Guide patients,” 2017.
- [37] B. Golding, A. Luu, R. Jones, and A. M. Vilorio-Petit, “The function and therapeutic targeting of anaplastic lymphoma kinase (ALK) in non-small cell lung cancer (NSCLC),” *Mol. Cancer*, vol. 17, no. 1, pp. 1–15, 2018, doi: 10.1186/s12943-018-0810-4.
- [38] T. Iwahara *et al.*, “Molecular characterization of ALK, a receptor tyrosine kinase

- expressed specifically in the nervous system,” *Oncogene*, vol. 14, no. 4, pp. 439–449, 1997, doi: 10.1038/sj.onc.1200849.
- [39] R. Chiarle, C. Voena, C. Ambrogio, R. Piva, and G. Inghirami, “The anaplastic lymphoma kinase in the pathogenesis of cancer,” *Nat. Rev. Cancer*, vol. 8, no. 1, pp. 11–23, 2008, doi: 10.1038/nrc2291.
- [40] A. Kruczynski *et al.*, “Oncogène ALK et translocations en hématologie,” vol. I, pp. 98–104, 2012.
- [41] F. C. Kelleher and R. McDermott, “The emerging pathogenic and therapeutic importance of the anaplastic lymphoma kinase gene,” *Eur. J. Cancer*, vol. 46, no. 13, pp. 2357–2368, 2010, doi: 10.1016/j.ejca.2010.04.006.
- [42] R. H. Palmer, E. Vernersson, C. Grabbe, and B. Hallberg, “Anaplastic lymphoma kinase: Signalling in development and disease,” *Biochem. J.*, vol. 420, no. 3, pp. 345–361, 2009, doi: 10.1042/BJ20090387.
- [43] C. J. Tartari *et al.*, “Characterization of some molecular mechanisms governing autoactivation of the catalytic domain of the anaplastic lymphoma kinase,” *J. Biol. Chem.*, vol. 283, no. 7, pp. 3743–3750, 2008, doi: 10.1074/jbc.M706067200.
- [44] T. G. Karabencheva, C. C. Lee, G. W. Black, R. Donev, and C. Z. Christov, “How does conformational flexibility influence key structural features involved in activation of anaplastic lymphoma kinase?,” *Mol. Biosyst.*, vol. 10, no. 6, pp. 1490–1495, 2014, doi: 10.1039/c4mb00141a.
- [45] N. James and K. Ramanathan, “Discovery of Potent ALK Inhibitors Using Pharmacophore-Informatics Strategy,” *Cell Biochem. Biophys.*, vol. 76, no. 1–2, pp. 111–124, 2017, doi: 10.1007/s12013-017-0800-y.
- [46] N. James, V. Shanthi, and K. Ramanathan, “Density Functional Theory and Molecular Simulation Studies for Prioritizing Anaplastic Lymphoma Kinase Inhibitors,” *Appl. Biochem. Biotechnol.*, vol. 190, no. 4, pp. 1127–1146, 2019, doi: 10.1007/s12010-019-03156-1.
- [47] C. H. Tu *et al.*, “Pyrazolylamine Derivatives Reveal the Conformational Switching between Type i and Type II Binding Modes of Anaplastic Lymphoma Kinase (ALK),”

- J. Med. Chem.*, vol. 59, no. 8, pp. 3906–3919, 2016, doi: 10.1021/acs.jmedchem.6b00106.
- [48] R. Roskoski, “Anaplastic lymphoma kinase (ALK): Structure, oncogenic activation, and pharmacological inhibition,” *Pharmacol. Res.*, vol. 68, no. 1, pp. 68–94, 2012, doi: 10.1016/j.phrs.2012.11.007.
- [49] M. Soda *et al.*, “Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer,” *Nature*, vol. 448, no. 7153, pp. 561–566, 2007, doi: 10.1038/nature05945.
- [50] S. R. Sabir, S. Yeoh, and G. Jackson, “EML4-ALK variants: Biological and molecular properties, and the implications for patients,” *Cancers (Basel)*, vol. 9, no. 9, 2017, doi: 10.3390/cancers9090118.
- [51] A. García-Regalado and C. H. González-De la Rosa, “The Role of Anaplastic Lymphoma Kinase in Human Cancers,” *Oncol. Hematol. Rev.*, vol. 09, no. 02, p. 149, 2013, doi: 10.17925/ohr.2013.09.2.149.
- [52] K. Ando *et al.*, “Brigatinib and Alectinib for ALK Rearrangement-Positive Advanced Non-Small Cell Lung Cancer with or without Central Nervous System Metastasis: A Systematic Review and Network Meta-Analysis,” *Cancers (Basel)*, vol. 12, no. 4, p. 942, Apr. 2020, doi: 10.3390/cancers12040942.
- [53] M. Viala, S. Brosseau, D. Planchard, B. Besse, and J. C. Soria, “Second generation ALK inhibitors in non-small cell lung cancer: Systemic review,” *Bull. Cancer*, vol. 102, no. 4, pp. 381–389, 2015, doi: 10.1016/j.bulcan.2015.02.016.
- [54] A. Passaro *et al.*, “Personalized treatment in advanced ALK-positive non-small cell lung cancer: From bench to clinical practice,” *Onco. Targets. Ther.*, vol. 9, pp. 6361–6376, 2016, doi: 10.2147/OTT.S98347.
- [55] L. Friboulet *et al.*, “The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer,” *Cancer Discov.*, vol. 4, no. 6, pp. 662–673, 2014, doi: 10.1158/2159-8290.CD-13-0846.
- [56] W. Chen, X. Guo, C. Zhang, D. Ke, G. Zhang, and Y. Yu, “Discovery of 2-aminopyridines bearing a pyridone moiety as potent ALK inhibitors to overcome the



- crizotinib-resistant mutants,” *Eur. J. Med. Chem.*, vol. 183, 2019, doi: 10.1016/j.ejmech.2019.111734.
- [57] G. Recondo, G. Recondo, R. Mechanisms, T. Kinase, and I. Tkis, “Resistance Mechanisms to ALK Tyrosine Kinase Inhibitors ( TKIs ) in NSCLC To cite this version : HAL Id : tel-02337169 Resistance mechanisms to ALK tyrosine kinase inhibitors in NSCLC Thèse de doctorat de l ’ Université Paris -Saclay Gonzalo Recondo,” 2019, [Online]. Available: <https://tel.archives-ouvertes.fr/tel-02337169>.
- [58] H. Ayachi, “Compréhension du mécanisme enzyme-substrat par modélisation moléculaire : Cas du diabète de type 2. Thèse de doctorat .,” Université Abou-bekr Belkaid - TLEMCEN.
- [59] C. Empereur-mot, “Développement d’outils statistiques d’évaluation de méthodes de criblage virtuel : courbes de prédictivité & Screening Explorer. Thèse de doctorat,” Conservatoire national des arts et metiers, 2018.
- [60] A. Rayar, “In silico drug design et chimie médicinale : développement de nouvelles molécules coumariniques, sélectives de la cyclooxygénase-2. Thèse de doctorat,” Conservatoire national des arts et metiers, 2017.
- [61] K. K. Chaudhary and N. Mishra, “A Review on Molecular Docking: Novel Tool for Drug Discovery,” *JSM Chem*, vol. 4, no. 3, 2016.
- [62] H. Boucherit, “Etude theorique des interactions intervenant dans l’inhibition de la methionine aminopeptidase de MYCOBACTERIUM TUBERCULOSIS par diverses molecules .Mémoire du magister en biochimie Option: Technologies des explorations biochimiques,” Université des Frères Mentouri Constantine-1, 2012.
- [63] Amina MERZOUG, “Recherche de nouvelles molécules bioactives par criblage virtuel et applications à l’inhibition de la peptide déformylase.Thèse de doctorat de l’Université : Biochimie .,” Université des Frères Mentouri Constantine-1, 2018.
- [64] X. Tao *et al.*, “Etude theorique des interactions intervenant dans l’inhibition de la methionine aminopeptidase de MYCOBACTERIUM TUBERCULOSIS par diverses molecules .Mémoire du magister en biochimie Option: Technologies des explorations biochimiques,” Université des Frères Mentouri Constantine-1, 2012.

- [65] Z. Rania, D. Abdelkarim, and Constantine, “Conception in silico de nouveaux inhibiteurs de l’ADAMTS-5 pour le traitement de l’arthrose, Mémoire du magister en biochimie. Université des Frères Mentouri Constantine,” 2019.
- [66] A. Grosdidier, “Conception d’un logiciel de docking et applications dans la recherche de nouvelles molécules actives Aurélien Grosdidier To cite this version : HAL Id : dumas-01115990,” 2015.
- [67] RCSB PDB, “PDB Statistics: Growth of Structures from 3DEM Experiments Released per Year,” 2019. <https://www.rcsb.org/stats/growth/growth-em> (accessed Jun. 07, 2021).
- [68] M. El-hassen, “Contribution à l’amélioration de l’activité biologique des inhibiteurs de la dipeptidyl-peptidase 4 dans le diabète de type 2 par simulation informatique .,” pp. 2011–2012, Mémoire du magister en biochimie Université des Frères Mentouri Constantine-1 2012.
- [69] B. A. B. Mohsane, “Conception in silico de nouveaux inhibiteurs De la Bêta-Lactamase de classe C par Docking Moléculaire 2017,” Mémoire du magister en Biochimie Moléculaire et Santé Université des Frères Mentouri Constantine-1 2017.
- [70] PubChem, “pubchem docs,” 2021, [Online]. Available: <https://pubchemdocs.ncbi.nlm.nih.gov/about>.
- [71] “ChemBioFrance - Infrastructure de recherche,” 2021. <https://chembiofrance.cn.cnrs.fr/fr/composante/chimiotheque> (accessed Jun. 07, 2021).
- [72] ZERDAOUI Sara BOUNASLA Sara, “Mémoire :Docking moléculaire de deux flavones naturels contre l’acétylcholinestérase et le butylcholinestérase,” Mémoire du magister en Chimie Pharmaceutique UNIVERSITÉ MOHAMED BOUDIAF – M’SILA 2020.
- [73] C. Esmâ and G. Selma, “inhibition de la beta-secretase 1 , enzyme clef de la maladie d’alzheimer , par la methode du docking,” 2015.
- [74] A. Merzoug, S. Mermoum, S. Khaled, A. Chikhi, A. Bensegueni, and H. Boucherit, “Theoretical Study of the Interactions Involved in the Inhibition of Staphylococcus aureus Peptide Deformylase by GSK1322322 Derivatives,” 2015.

- [75] E. H. Mokrani, Y. Krid, and A. Bensegueni, "Highlighting of a New Flavonol Derivatives as a Potent Antihypertensive Compound using Molecular Docking," *J. Acad.*, vol. 3, no. 2, 2013.
- [76] B. Hanane, Nouioua Amira, Serrar Asma, Chikhi Abdelouahab, Bensegueni Abderrahmane, and Merzoug Amina, "Research of new inhibitors of staphylococcus aureus methionine aminopeptidase by computer stimulation," 2016.
- [77] H. Boucherit, N. Zouaghi, Z. Sara, A. Chikhi, A. Bensegueni, and A. Merzoug, "Use of molecular docking in the search for new anti-angiogenic agents," *Cancer Reports Rev.*, 2017.
- [78] A. Merzoug, M. Lalaoui, sana Matoug, A. Chikhi, Abderrahmane Bensegueni, and Hanane Boucherit, "Research of new anti-tuberculosis agents by molecular docking ' s method," vol. 6, no. 2, 2016.
- [79] A. Chikhi and Abderrahmane Bensegueni, "Docking Efficiency Comparison of Surfex, a Commercial Package and Arguslab, a Licensable Freewar," *J. Comput. Sci. Syst. Biol.*, 2008.
- [80] Teniou Soumia, "Conception in silico de nouveaux inhibiteurs de la topoisomérase I de cellules cancéreuses. Thèse de Magistère en biochimie. Algérie: Université Mentouri Constantine.," 2012.
- [81] B. Hanane, Chikhi Abdelouahab, Bensegueni Abderrahmane, Merzoug Amina, Hioual khadidja Soulef, and Mokrani El Hassen, "Theoretical Study of the Interactions Involved in the Inhibition of Mycobacterium Tuberculosis Methionine Aminopeptidase by Several Molecules," *Comput. Biol. Bioinforma.*, 2014.
- [82] A. Chikhi and A. Bensegueni, "In silico study of the selective inhibition of Bacterial Peptide Deformylases by several drugs," *J. Proteomics Bioinforma.*, 2010.
- [83] M. M. Kamel, H. I. Ali, M. M. Anwar, N. A. Mohamed, and A. M. M. Soliman, "Synthesis, antitumor activity and molecular docking study of novel Sulfonamide-Schiff's bases, thiazolidinones, benzothiazinones and their C-nucleoside derivatives," *Eur. J. Med. Chem.*, 2010.



***ANNEXES***

**Annexe 1** : Tableau RMSD de 100 complexes Protéine – ligand.

<b>N</b>	<b>Protéine</b>	<b>RMSD</b>
1	4FNY	3.156
2	4CLI	1.883
3	5A9U	0.087
4	2WKM	0.883
5	4QUW	1.044
6	4FOB	1.201
7	3NNU	0.245
8	7BTT	2.466
9	3L9P	1.184
10	5IMX	1.221
11	5CEP	1.184
12	1N7I	1.291
13	4UXL	0.625
14	4ANS	1.948
15	2WGJ	1.089
16	1RWQ	0.107
17	1ZHM	1.084
18	2OPH	1.800
19	5CEQ	0.861
20	2PEV	1.046
21	2QCH	0.160
22	1AH3	0.836
23	1EKO	2.608
24	4NCM	1.439
25	3FPM	1.780
26	2QCG	2.063
27	9HVP	1.616
28	1KAV	0.295
29	3FHR	0.123
30	4NGN	1.266

## ANNEXES

31	3FXW	0.137
32	5URS	0.609
33	3CCC	0.529
34	6LUD	1.372
35	5XGM	1.490
36	5FI2	1.629
37	3ENE	0.041
38	6D8E	1.671
39	5XDL	2.023
40	5XDK	1.290
41	5FI7	1.840
42	5X2K	2.376
43	5AUT	0.203
44	1QXK	1.733
45	1XOE	1.569
46	1HGG	9.962
47	6VHN	1.732
48	1SEU	1.073
49	4HJO	1.688
50	5IUI	1.465
51	4CTB	1.921
52	6MX8	3.378
53	4CM0	1.525
54	4CMT	1.225
55	4CNH	0.698
56	3LCS	0.838
57	4JOA	2.443
58	4ANP	1.999
59	2FWD	1.232
60	4RC5	3.808
61	2F02	1.311
62	3ZBF	0.491
63	6TF2	4.901

## ANNEXES

64	5AAA	0.434
65	5AAC	0.483
66	2XB7	0.669
67	4CCB	1.953
68	5FTQ	0.480
69	4DCE	0.236
70	2XP2	0.262
71	4CO3	0.860
72	6BRC	0.613
73	3ZBP	0.491
74	4CO3	0.509
75	2WOO	1.186
76	6DUK	0.667
77	5D41	1.489
78	5FED	1.719
79	4DPF	2.156
80	4RJ3	2.225
81	4R5Y	1.008
82	5D12	2.067
83	6LUB	3.280
84	2QLQ	9.472
85	2WLS	0.332
86	4UY8	7.667
87	2C5G	10.538
88	1H22	2.532
89	3ZV7	6.099
90	5QQ6	2.194
91	6CD5	4.877
92	1P7C	1.492
93	5DA3	0.271
94	6HAZ	3.513
95	2IZR	0.083
96	4O75	5.004

**ANNEXES**

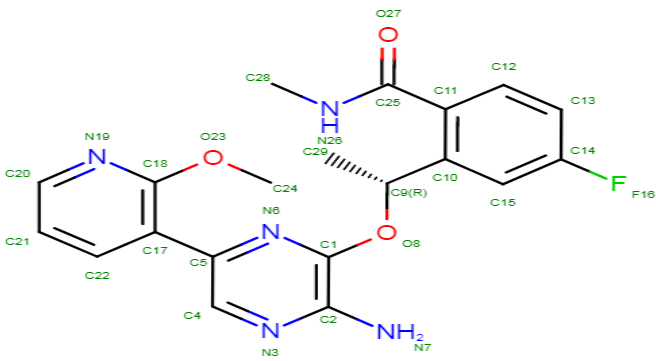
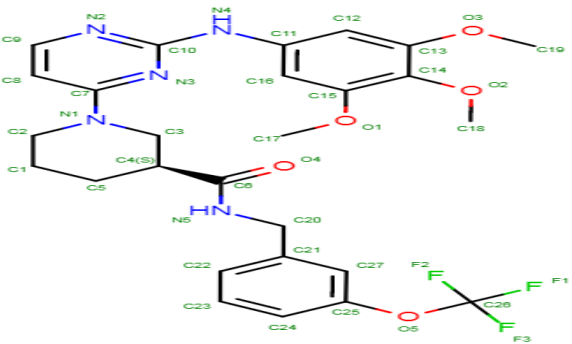
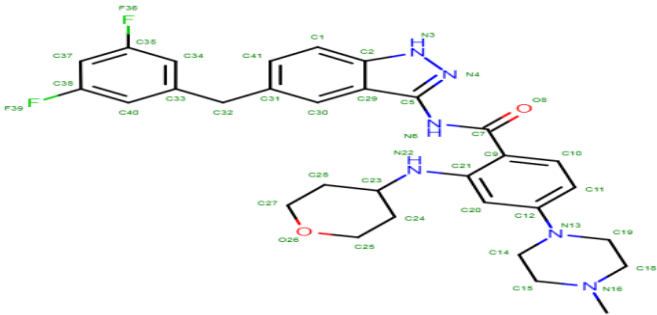
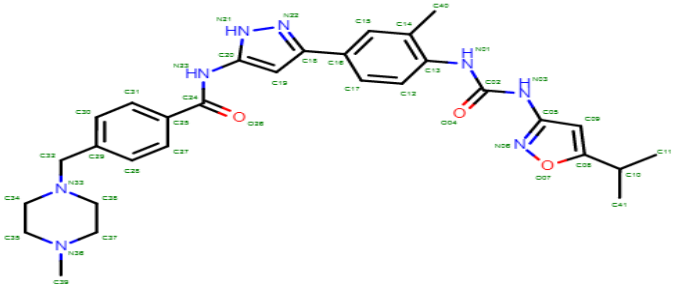
<b>97</b>	4I72	9.292
<b>98</b>	3CXR	5.481
<b>99</b>	6FNM	0.197
<b>100</b>	5E3I	4.419

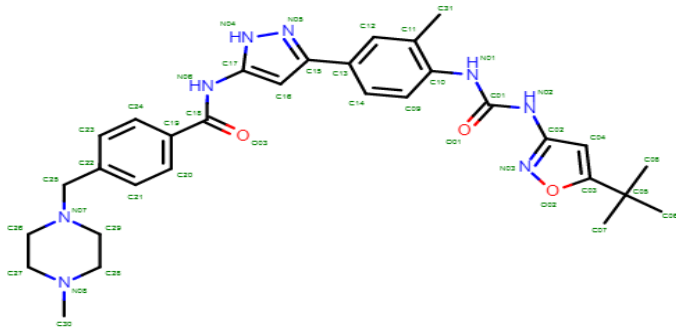
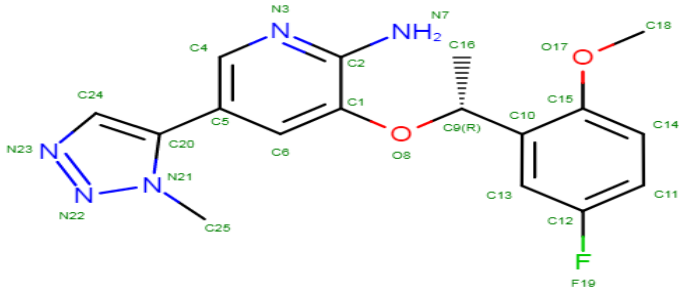
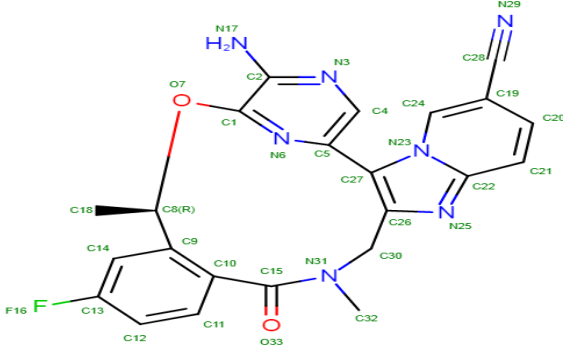
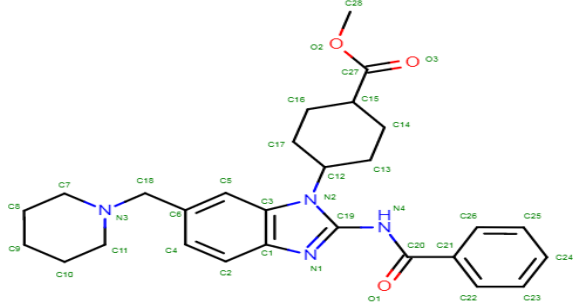


**Annexe 2:** Les 30 inhibiteurs de l'ALK ainsi leurs scores ET leurs IC50.

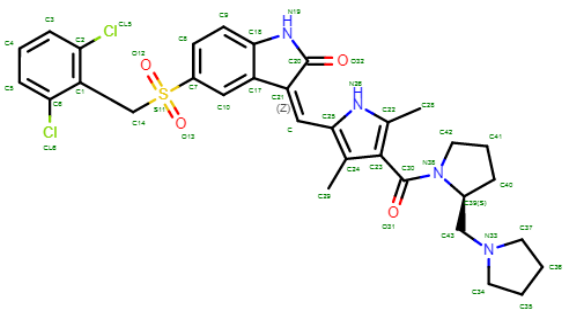
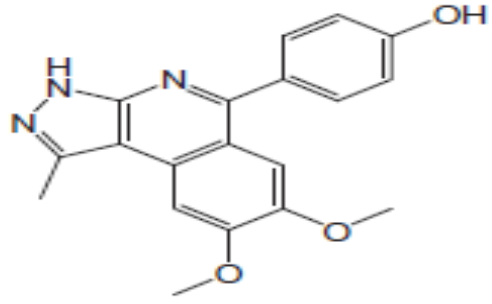
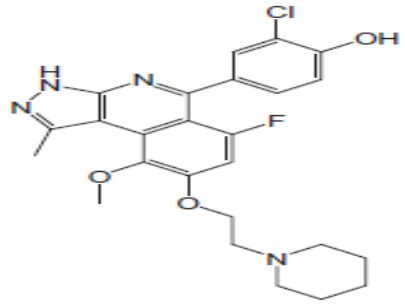
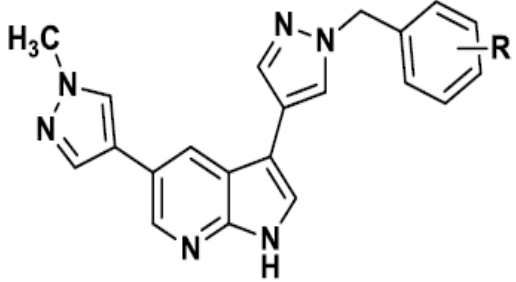
Composé	IC50 (nM)	LogIC50	Score	Structure
1	2.5	0.39	4.34	
2	2	0.30	4.86	
3	20	1.3	4.07	

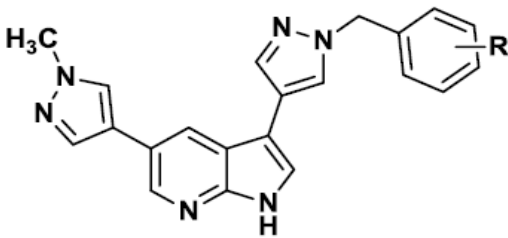
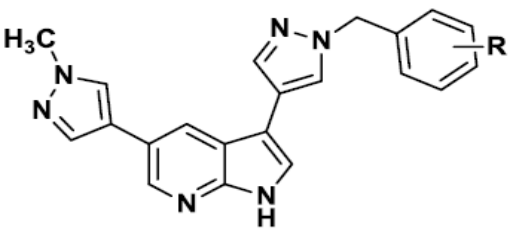
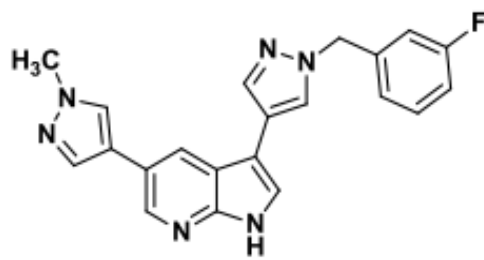
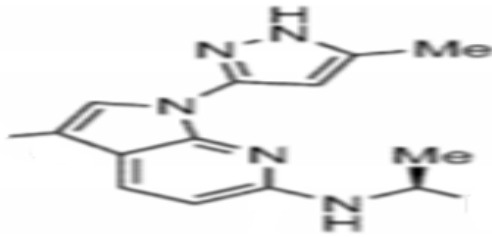
4	14	1.14	6.96	
5	5.8	0.76	5.9	
6	2.7	0.43	6.68	
7	4	0.60	3.56	

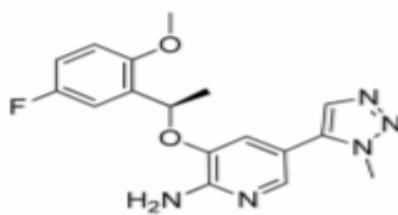
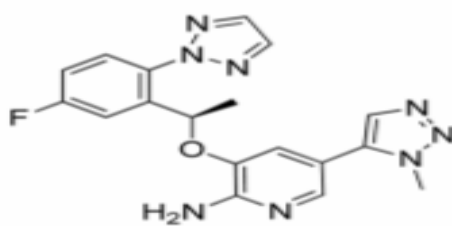
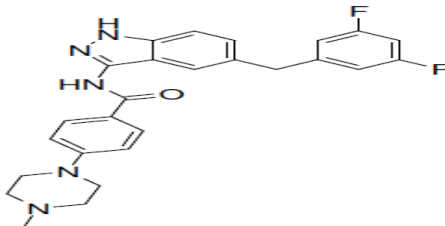
8	16	1.20	5.41	 <p>Chemical structure of compound 8, showing a complex molecule with multiple rings and functional groups. The structure is labeled with atom identifiers: C1-C22, N1-N19, O1-O27, and F1-F16. It features a pyridine ring (C17-C22, N19), a pyrazole ring (C1-C5, N3, N6), and a benzene ring (C10-C15) with a fluorine atom (F16). Other rings include a piperazine-like ring (C23-C28, N26, N29) and a carbonyl group (C25, O27).</p>
9	16	1.20	6.63	 <p>Chemical structure of compound 9, showing a complex molecule with multiple rings and functional groups. The structure is labeled with atom identifiers: C1-C27, N1-N4, O1-O4, and F1-F3. It features a pyridine ring (C6-C10, N2, N3), a piperazine ring (C1-C5, N1, N4), and a benzene ring (C11-C15) with a fluorine atom (F1). Other rings include a piperazine-like ring (C16-C21, N5, N6) and a carbonyl group (C6, O4).</p>
10	12	1.07	4.23	 <p>Chemical structure of compound 10, showing a complex molecule with multiple rings and functional groups. The structure is labeled with atom identifiers: C1-C41, N1-N4, O1-O8, and F1-F36. It features a pyridine ring (C31-C35, N3, N4), a benzene ring (C1-C5), and a piperazine ring (C16-C21, N13, N16). Other rings include a piperazine-like ring (C22-C27, N6, N8) and a carbonyl group (C5, O8).</p>
11	402	2.6	8.56	 <p>Chemical structure of compound 11, showing a complex molecule with multiple rings and functional groups. The structure is labeled with atom identifiers: C1-C41, N1-N22, O1-O4, and F1-F36. It features a pyridine ring (C31-C35, N21, N22), a benzene ring (C1-C5), and a piperazine ring (C16-C21, N13, N16). Other rings include a piperazine-like ring (C22-C27, N6, N8) and a carbonyl group (C5, O8).</p>

12	177	2.2	8.67	
13	95	1.97	6.17	
14	45	1.65	8.97	
15	3	0.47	6.15	

16	5	0.69	5.31	
17	11	1.04	6.52	
18	29	1.46	5.43	
19	10	1	6.24	

20	9	0.95	5.16	
21	99	1.99	7.85	
22	99	1.99	8.26	
23	85	1.92	7.86	 <p>R : 2-F</p>

24	5.6	0.74	7.69	 <p>R : 3-F</p>
25	247	2.39	8.35	 <p>R : 3-CF<sub>3</sub></p>
26	5.6	0.74	6.39	
27	2	0.30	5.78	
28	4	0.60	5.72	

				
29	9	0.95	5.88	
30	73	1.86	7.58	

**Annexe 3:** Numéros d'accèsion et scores des 449 similaires.

N	CODE	SCORE
1	CID_100665031	9.12
2	CID_10251289	7.51
3	CID_104616147	7.43
4	CID_104616423	6.67
5	CID_104618063	6.65
6	CID_104618259	7.62
7	CID_104618450	7.25
8	CID_104618848	7.85
9	CID_108451298	0.43



## ANNEXES

10	CID_108451669	8.64
11	CID_108451682	7.01
12	CID_108451683	6.43
13	CID_108454909	7.22
14	CID_108458034	6.66
15	CID_108458317	4.56
16	CID_108459520	2.40
17	CID_108460111	4.56
18	CID_108460939	5.66
19	CID_108462087	5.95
20	CID_108462158	4.07
21	CID_108462416	6.18
22	CID_108462430	6.23
23	CID_108485662	4.92
24	CID_11249084	1.32
25	CID_11305328	6.39
26	CID_11305344	5.56
27	CID_113433321	6.83
28	CID_11408844	3.50
29	CID_11487958	6.28
30	CID_11545405	6.06
31	CID_115595333	7.24
32	CID_115603638	7.47
33	CID_115603642	6.93
34	CID_115610381	6.88
35	CID_11561187	6.82
36	CID_115621109	7.72
37	CID_117631931	6.83
38	CID_117636414	4.21
39	CID_118699735	8.25
40	CID_118723697	4.42
41	CID_119025616	1.47
42	CID_119025618	8.86

## ANNEXES

43	CID_119067735	4.51
44	CID_11994444	3.26
45	CID_11994445	5.74
46	CID_122555784	6.36
47	CID_122584268	6.91
48	CID_122584270	2.27
49	CID_122584278	5.77
50	CID_122584294	2.07
51	CID_122584307	2.60
52	CID_122584313	6.51
53	CID_122584315	4.54
54	CID_122584316	2.04
55	CID_122584317	6.51
56	CID_122584326	6.90
57	CID_122584331	5.24
58	CID_122584343	5.95
59	CID_122584347	5.71
60	CID_122584350	5.37
61	CID_122584353	3.49
62	CID_122584379	6.44
63	CID_122584384	5.33
64	CID_122584395	3.49
65	CID_122584398	4.33
66	CID_122584401	0.89
67	CID_122584417	5.49
68	CID_122584425	3.66
69	CID_122584455	3.09
70	CID_122584460	7.06
71	CID_122584469	6.75
72	CID_122584516	4.13
73	CID_122584523	4.35
74	CID_122584542	4.32
75	CID_122584549	4.19

## ANNEXES

76	CID_122584561	2.27
77	CID_122584566	5.16
78	CID_122584574	3.09
79	CID_122597187	6.17
80	CID_123326880	4.05
81	CID_123360512	5.73
82	CID_124111000	6.49
83	CID_124111020	6.13
84	CID_124111179	6.69
85	CID_124111181	7.23
86	CID_124111240	5.50
87	CID_124118342	5.82
88	CID_124118344	6.65
89	CID_124118346	7.07
90	CID_124118382	7.11
91	CID_124118383	3.28
92	CID_124118386	6.36
93	CID_126440431	4.51
94	CID_126440432	3.71
95	CID_127045515	8.22
96	CID_127373847	6.84
97	CID_129094005	2.51
98	CID_129094014	4.01
99	CID_129107839	5.23
100	CID_129186314	2.38
101	CID_129198313	5.12
102	CID_132359836	9.08
103	CID_134262922	3.07
104	CID_134341421	6.34
105	CID_134341469	5.48
106	CID_134341471	5.79
107	CID_134341516	6.23
108	CID_134341545	7.65

## ANNEXES

109	CID_134341576	5.69
110	CID_134341595	6.49
111	CID_136952455	3.48
112	CID_137653101	5.96
113	CID_137974346	7.88
114	CID_138995886	7.83
115	CID_139012022	3.55
116	CID_139558757	5.61
117	CID_140609179	5.73
118	CID_140697433	5.03
119	CID_141458906	6.67
120	CID_142965322	6.83
121	CID_144015793	4.33
122	CID_144123699	6.62
123	CID_145192770	6.85
124	CID_145898042	9.10
125	CID_146127181	7.50
126	CID_146143897	7.45
127	CID_147677454	3.48
128	CID_149231681	6.64
129	CID_149894116	5.77
130	CID_150412772	7.43
131	CID_150872190	5.47
132	CID_151436481	9.04
133	CID_152263403	6.02
134	CID_152276299	4.75
135	CID_152432300	6.48
136	CID_152516999	4.95
137	CID_153113905	6.76
138	CID_153249899	2.61
139	CID_154805736	5.69
140	CID_154837768	3.01
141	CID_154851017	5.40

## ANNEXES

142	CID_154856764	7.97
143	CID_154857672	1.11
144	CID_154864275	6.99
145	CID_154864276	7.61
146	CID_154864279	5.95
147	CID_154873584	7.01
148	CID_154873586	8.37
149	CID_154873623	8.02
150	CID_154873630	6.80
151	CID_155059172	5.96
152	CID_155208895	8.98
153	CID_155701135	6.67
154	CID_16002421	6.22
155	CID_20910766	6.76
156	CID_21106481	5.19
157	CID_22176705	5.39
158	CID_22176765	6.59
159	CID_22176920	6.76
160	CID_22176953	6.25
161	CID_22478720	7.96
162	CID_24022867	5.35
163	CID_24044094	8.14
164	CID_24893958	3.40
165	CID_24946432	5.10
166	CID_24946434	5.35
167	CID_24947510	5.70
168	CID_24947514	3.20
169	CID_24961557	7.36
170	CID_24987271	5.49
171	CID_24987624	4.48
172	CID_24988746	0.57
173	CID_25174894	5.00
174	CID_25175551	5.49

## ANNEXES

175	CID_25175709	6.80
176	CID_25175710	5.14
177	CID_25175711	4.23
178	CID_25175712	3.46
179	CID_25175717	3.28
180	CID_25175859	4.47
181	CID_25176013	7.50
182	CID_28330750	5.54
183	CID_29257234	6.38
184	CID_3139653	8.51
185	CID_3139654	8.58
186	CID_3686738	7.27
187	CID_41077940	5.70
188	CID_45106915	5.08
189	CID_45517848	4.88
190	CID_45517998	5.36
191	CID_45518149	6.48
192	CID_45518293	5.60
193	CID_47038562	6.14
194	CID_47079615	6.75
195	CID_49806879	4.70
196	CID_50816218	6.31
197	CID_50816262	5.92
198	CID_52690382	6.68
199	CID_52906492	5.56
200	CID_52906493	5.72
201	CID_52906498	5.57
202	CID_53009791	5.91
203	CID_53039092	6.03
204	CID_53098711	4.93
205	CID_53098713	4.92
206	CID_53098715	5.01
207	CID_53098720	4.50

## ANNEXES

208	CID_53098734	6.46
209	CID_53098736	5.23
210	CID_53098741	5.16
211	CID_53180948	3.11
212	CID_53180951	2.99
213	CID_53180965	4.76
214	CID_53180967	3.87
215	CID_55219555	5.85
216	CID_55862532	6.37
217	CID_55864567	6.16
218	CID_56552392	6.89
219	CID_56552617	3.97
220	CID_56663280	8.24
221	CID_56666718	5.49
222	CID_56676976	2.84
223	CID_56826856	5.54
224	CID_56944168	5.38
225	CID_58810434	7.10
226	CID_59222397	2.43
227	CID_59557866	5.87
228	CID_59577330	4.42
229	CID_59577363	3.51
230	CID_60198831	7.61
231	CID_60198893	8.19
232	CID_60199002	5.99
233	CID_60322569	5.06
234	CID_60339547	7.39
235	CID_60519257	6.72
236	CID_60547740	8.53
237	CID_60547770	8.08
238	CID_60547797	7.87
239	CID_63821728	6.02
240	CID_66589757	7.33

## ANNEXES

241	CID_67212900	3.82
242	CID_67452580	6.04
243	CID_67454111	5.67
244	CID_68234021	4.69
245	CID_68234128	5.92
246	CID_68440585	3.46
247	CID_68441433	8.35
248	CID_99756595	3.07
249	CID_68443078	4.91
250	CID_68443172	3.49
251	CID_68670811	6.80
252	CID_68671821	1.03
253	CID_68672404	3.82
254	CID_68884394	1.30
255	CID_69761927	7.08
256	CID_70654390	5.46
257	CID_70654476	5.36
258	CID_70654628	7.40
259	CID_70654674	6.46
260	CID_70654694	9.16
261	CID_70654695	8.38
262	CID_70654711	6.45
263	CID_70654866	5.24
264	CID_70654869	6.35
265	CID_70654875	1.60
266	CID_70654897	6.82
267	CID_70654968	5.43
268	CID_70654989	3.49
269	CID_70655001	4.07
270	CID_70655002	9.70
271	CID_70655032	4.40
272	CID_70655057	8.69
273	CID_70660503	6.44



## ANNEXES

274	CID_70660507	8.00
275	CID_70661429	7.89
276	CID_70661446	7.50
277	CID_70688403	2.82
278	CID_71654821	4.09
279	CID_71654879	4.64
280	CID_71654880	4.52
281	CID_71654956	4.45
282	CID_71658742	3.21
283	CID_71658743	4.05
284	CID_71659100	7.33
285	CID_71778645	6.47
286	CID_72000881	6.76
287	CID_72011530	5.36
288	CID_73951507	5.37
289	CID_75358577	6.74
290	CID_75376098	6.92
291	CID_75422379	6.74
292	CID_75527216	4.65
293	CID_75527217	5.57
294	CID_75527222	5.30
295	CID_75527229	5.36
296	CID_75527232	6.50
297	CID_75527234	5.60
298	CID_75527235	6.09
299	CID_75527241	4.72
300	CID_75527250	5.76
301	CID_75527251	4.77
302	CID_75527252	4.59
303	CID_75527260	5.53
304	CID_75527266	4.69
305	CID_75527331	4.75
306	CID_75527333	4.95

## ANNEXES

307	CID_75527334	4.94
308	CID_75527337	5.54
309	CID_75527346	5.55
310	CID_75527348	5.46
311	CID_75527350	5.26
312	CID_75527352	6.05
313	CID_75527354	5.47
314	CID_75527358	4.79
315	CID_75527360	4.23
316	CID_75528127	3.42
317	CID_75528138	3.07
318	CID_75528148	3.18
319	CID_75528150	3.58
320	CID_75528151	2.21
321	CID_75528152	5.66
322	CID_75528156	2.29
323	CID_75528162	4.40
324	CID_75528173	5.57
325	CID_75528230	4.90
326	CID_75528232	3.43
327	CID_75528235	5.92
328	CID_75528403	2.51
329	CID_75528404	4.61
330	CID_75528409	3.08
331	CID_75528414	4.05
332	CID_75528424	3.75
333	CID_75528432	2.31
334	CID_75528451	3.58
335	CID_75528463	4.25
336	CID_75528478	6.03
337	CID_75528482	2.91
338	CID_75528488	5.38
339	CID_7593274	5.81

## ANNEXES

340	CID_7593298	5.77
341	CID_76285482	7.33
342	CID_84554661	6.91
343	CID_84565737	5.30
344	CID_84579450	7.63
345	CID_84579632	2.55
346	CID_84579942	5.65
347	CID_84581189	7.08
348	CID_84581459	6.19
349	CID_84582309	6.31
350	CID_84582375	6.95
351	CID_846005	7.88
352	CID_8670400	5.23
353	CID_8670406	7.10
354	CID_90862360	4.49
355	CID_91155643	4.64
356	CID_91231758	7.59
357	CID_91631959	4.17
358	CID_91887604	2.93
359	CID_91887671	4.50
360	CID_91887681	4.09
361	CID_91887682	3.46
362	CID_91887683	4.72
363	CID_91887684	3.05
364	CID_91887746	4.80
365	CID_91887750	5.30
366	CID_91887751	5.38
367	CID_91887758	2.48
368	CID_91887765	3.64
369	CID_91887766	0.58
370	CID_91887768	3.61
371	CID_91887771	3.07
372	CID_91887785	3.83

## ANNEXES

373	CID_91887786	3.84
374	CID_91887792	3.93
375	CID_91887849	3.51
376	CID_91887860	3.63
377	CID_91887861	4.24
378	CID_91887867	5.09
379	CID_91887868	3.54
380	CID_91887872	2.82
381	CID_91887895	1.53
382	CID_91887907	1.70
384	CID_91904736	5.48
385	CID_91904745	5.67
386	CID_91904750	2.26
387	CID_91904775	3.05
388	CID_91904816	3.90
389	CID_91904817	2.84
390	CID_91904851	6.77
391	CID_91904852	7.14
392	CID_91904859	4.53
393	CID_91904860	5.57
394	CID_91904871	5.62
395	CID_91914496	2.52
396	CID_91914514	5.35
397	CID_91914517	3.35
398	CID_91914532	3.77
399	CID_91914560	2.45
400	CID_91914567	3.67
401	CID_91914568	5.34
402	CID_91914569	4.38
403	CID_91943065	7.21
404	CID_95755345	7.69
405	CID_95755346	7.56
406	CID_95755356	7.70

## ANNEXES

407	CID_95755359	6.97
408	CID_95755380	7.29
409	CID_95755385	5.59
410	CID_95755412	7.49
411	CID_95755418	7.44
412	CID_95755440	7.48
413	CID_9888885	7.51
414	CID_99740768	2.86
415	CID_99740769	3.83
416	CID_99740770	3.84
417	CID_99740771	1.49
418	CID_99740792	2.21
419	CID_99740793	4.80
420	CID_99740800	3.10
421	CID_99740801	5.30
422	CID_99740802	2.74
423	CID_99740803	5.38
424	CID_99740812	2.88
425	CID_99740813	2.48
426	CID_99740824	3.64
427	CID_99740825	2.88
428	CID_99740826	2.10
429	CID_99740827	0.58
430	CID_99740835	2.27
431	CID_99740836	3.61
432	CID_99740845	3.93
433	CID_99740846	3.14
434	CID_99741009	3.13
435	CID_99741010	3.51
436	CID_99741029	3.29
437	CID_99741030	3.34
438	CID_99741039	2.95
439	CID_99741040	5.09

440	CID_99741043	3.56
441	CID_99741044	3.54
442	CID_99741053	2.71
443	CID_99741054	2.82
444	CID_99741109	2.92
445	CID_99741110	1.53
446	CID_99741137	1.70
447	CID_99756206	3.08
448	CID_99756208	4.24
449	CID_99756594	2.26

**Annexe 4:** Résultats de l'affinité de l'ALK et ses deux meilleur dérivés obtenus par le programme Surfex.

**Similaire N° 260 :** CID\_70654694

```

C:\Windows\system32\cmd.exe
Polishing: .....
Times (total 21): search 0 align 9 merge+chain 10 polish 2
Nconfs (final 99): 3 50 2970 2800
[ret_conf -2.33] [ret_conf -5.51] [ret_conf -7.72] [ret_conf 0.91] [ret_conf -1.18] [ret_conf -4.39] [ret_conf -8.96] [ret_conf -0.75] [ret_conf 1.89] [ret_conf -7.93] [ret_conf -8.67] [ret_conf 1.33] [ret_conf 0.81] [ret_conf -2.51] [ret_conf -7.15] [ret_conf -0.78] [ret_conf -10.87] [ret_conf -8.43] [ret_conf -10.75] [ret_conf -2.35] [ret_conf 2.16] [ret_conf 2.32] [ret_conf -6.91] [ret_conf -0.43] [ret_conf 3.64] [ret_conf -4.72] [ret_conf -6.68] [ret_conf -8.45] [ret_conf 0.36] [ret_conf -3.18] [ret_conf 0.46] [ret_conf -5.54] [ret_conf -8.37] [ret_conf -2.37] [ret_conf -0.34] [ret_conf 1.62] [ret_conf -6.78] [ret_conf -7.99] [ret_conf -8.22] [ret_conf -5.61] [ret_conf 5.47] [ret_conf -6.98] [ret_conf -2.88] [ret_conf -2.31] [ret_conf -5.90] [ret_conf 5.36] [ret_conf -6.47] [ret_conf -4.66] [ret_conf -2.23] [ret_conf 3.64] final-0: 9.16 (crash -1.22 (self -0.40))
final-1: 8.93 (crash -2.07 (self -0.27))
final-2: 8.60 (crash -1.25 (self -0.18))
final-3: 7.10 (crash -1.60 (self -0.30))
final-4: 6.72 (crash -2.86 (self -0.62))
final-5: 6.00 (crash -3.69 (self -0.37))
final-6: 5.76 (crash -3.69 (self -0.34))
final-7: 4.33 (crash -4.12 (self -0.57))
final-8: 3.41 (crash -5.81 (self -1.41))
final-9: 3.04 (crash -6.34 (self -1.34))
C:\Users\mp\Desktop\pub\S_260>

```

Similaire N° 270 : CID\_70655002

```
C:\Windows\system32\cmd.exe
Polishing: .....
Times (total 21): search 0 align1 10 merge+chain 10 polish 1
Nconfs (final 99): 3 50 2970 2800
[ret_conf -0.17] [ret_conf -2.96] [ret_conf 0.88] [ret_conf -4.07] [ret_conf 0.7
2] [ret_conf -3.36] [ret_conf 0.33] [ret_conf -3.66] [ret_conf -0.10] [ret_conf
-1.31] [ret_conf -4.64] [ret_conf -1.47] [ret_conf -3.68] [ret_conf -1.85] [ret_
conf -1.53] [ret_conf -0.80] [ret_conf 2.50] [ret_conf 2.53] [ret_conf 6.87] [re
t_conf 3.03] [ret_conf 4.99] [ret_conf -2.87] [ret_conf -0.14] [ret_conf -3.37]
[ret_conf 0.19] [ret_conf -4.95] [ret_conf -1.97] [ret_conf -3.75] [ret_conf 0.1
9] [ret_conf -1.17] [ret_conf 1.74] [ret_conf -3.94] [ret_conf -3.11] [ret_conf
1.45] [ret_conf 4.23] [ret_conf -1.31] [ret_conf 6.93] [ret_conf -2.61] [ret_con
f 2.83] [ret_conf 0.19] [ret_conf 2.71] [ret_conf 2.73] [ret_conf 4.85] [ret_con
f -1.34] [ret_conf -3.83] [ret_conf -1.23] [ret_conf 2.61] [ret_conf -5.81] [ret
_conf 3.72] [ret_conf -6.16] final-0: 9.70 (crash -2.45 (self -0.21))
final-1: 9.33 (crash -2.35 (self -0.22))
final-2: 8.91 (crash -2.67 (self -0.25))
final-3: 5.74 (crash -3.93 (self -0.51))
final-4: 5.15 (crash -4.35 (self -0.93))
final-5: 5.10 (crash -2.66 (self -0.44))
final-6: 4.88 (crash -5.59 (self -0.26))
final-7: 4.81 (crash -2.86 (self -0.63))
final-8: 4.42 (crash -4.58 (self -0.29))
final-9: 4.41 (crash -4.91 (self -0.48))
C:\Users\mp\Desktop\pub\S_270>
```



# ***Résumé***



## **Résumé**

Le cancer du poumon est devenue un problème de santé publique majeur au niveau mondial et représente la tumeur la plus fréquente chez les hommes et la première cause de mortalité masculine par cancer. Dans ce travail, nous avons tenté de proposer de nouveaux inhibiteurs de la kinase ALK (Anaplastic lymphoma kinase), cible thérapeutique du cancer du poumon non à petites cellules. Pour cela, nous avons fait recours à l'approche par docking moléculaire à l'aide du programme Surflex. Nous avons réalisé un docking moléculaire d'une collection de 449 similaires du composé 45Q de la base de données PubChem. À la suite de ces travaux, le composé CID\_70655002 (9,70 M-1) et le composé CID\_70654694 (9,16 M-1) ont démontré la plus grande affinité. Enfin, l'application de la règle de 5 de lipinski nous renseigne de manière positive sur les composés CID\_70655002 et CID\_70654694, qui se présentent comme de nouveaux inhibiteurs potentiellement plus actifs envers la kinase ALK.

**Mots Clés :** Anaplastic lymphoma kinase (ALK), cancer du poumon non à petites cellules, Docking moléculaire, Surflex.

## **Abstract**

Lung cancer has become a major public health problem worldwide and represents the most frequent tumor in men and the first cause of male cancer mortality. In this work, we tried to propose new inhibitors of ALK kinase, a therapeutic target for non-small cell lung cancer. For that, we used the molecular docking approach with the Surflex program. We performed a molecular docking of a collection of 449 similar compounds of compound 45Q from the PubChem database. As a result of this work, compound CID\_70655002 (9.70 M-1) and compound CID\_70654694 (9.16 M-1) showed the highest affinity. Finally, the application of the lipinski's rule of 5 gives us positive information about the compounds CID\_70655002 and CID\_70654694, which present themselves as new and potentially more active inhibitors of ALK.

**Key Words:** Anaplastic lymphoma kinase (ALK), Non-small cell lung cancer (NSCLC), Molecular docking, Surflex.

## الملخص

لقد أصبح سرطان الرئة مشكلة صحية عامة رئيسية في جميع أنحاء العالم ويمثل الورم الأكثر شيوعًا لدى الرجال ويعتبر السبب الرئيسي للوفاة عند الرجال. في هذا العمل، حاولنا اقتراح مثبطات جديدة لـ ALK kinase وهو هدف علاجي لسرطان الرئة ذو الخلايا غير الصغيرة. لذلك استخدمنا نهج الالتحام الجزيئي باستخدام برنامج SurFlex. أجرينا الالتحام الجزيئي لمجموعة من 449 نظيرًا كيميائيًا للمركب 45Q من قاعدة بيانات PubChem. نتيجة لهذا العمل، أظهر المركب (CID\_70655002 (9.70 M-1) والمركب (CID\_70654694 (9.16 M-1) أعلى طاقة ارتباط. أخيرًا، بمنحنا تطبيق قاعدة Lipinski معلومات إيجابية عن المركبين CID\_70655002 و CID\_70654694، اللذان يبدو أنهما مثبطات جديدة. يحتمل أن تكون أكثر نشاطًا ضد ALK kinase.

**الكلمات المفتاحية:** سرطان الرئة ذو الخلايا غير الصغيرة ALK، الالتحام الجزيئي، Surflex.

**Année universitaire : 2020/2021**

**Présenté par : BERLAT Ahmed Samy  
BOUANAKA Ahmed**

**Conception *in silico* de nouveaux inhibiteurs de l'Anaplastic lymphoma kinase (ALK)  
une cible thérapeutique pour le traitement du cancer du poumon non à petits cellules  
(CPNPC)**

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

Le cancer du poumon est devenu un problème de santé publique majeur au niveau mondial et représente la tumeur la plus fréquente chez les hommes et la première cause de mortalité masculine par cancer. Dans ce travail, nous avons tenté de proposer de nouveaux inhibiteurs de la kinase ALK (Anaplastic lymphoma kinase), cible thérapeutique du cancer du poumon non à petites cellules. Pour cela, nous avons fait recours à l'approche par docking moléculaire à l'aide du programme Surflex. Nous avons réalisé un docking moléculaire d'une collection de 449 similaires du composé 45Q de la base de données PubChem. À la suite de ces travaux, le composé CID\_70655002 (9,70 M-1) et le composé CID\_70654694 (9,16 M-1) ont démontré la plus grande affinité. Enfin, l'application de la règle de 5 de lipinski nous renseigne de manière positive sur les composés CID\_70655002 et CID\_70654694, qui se présentent comme de nouveaux inhibiteurs potentiellement plus actifs envers la kinase ALK.

**Mots clés :** Anaplastic lymphoma kinase (ALK), cancer du poumon non à petites cellules, Docking moléculaire, Surflex.

**Laboratoire de recherche :** Laboratoire de Biochimie Appliquée

**Jury d'évaluation :**

**Président du jury :** Mme GUENDOUZE.A (Maître de conférences B, UFM - Constantine 1.)  
**Examineurs :** Mme TENIOU.S (Maître assistanteA, UFM - Constantine 1.)  
**Encadreur :** Mme DEMMAK.R-G (Maître de conférences B, USB - Constantine 3.)

**Date de soutenance :** 14/07/2021