



*Université Mohamed Khider de Biskra
Faculté des Sciences Exactes et des Sciences de la Nature de
la Vie*

Département des Sciences de la Matière

MÉMOIRE DE MASTER

Domaine des Sciences de la matière

Filière de Chimie

Spécialité Chimie Pharmaceutique

Réf :

Présenté et soutenu par :

Ghacha Abdelhak

Fekroune Abdelhamid

Le : 11/07/ 2021

***Etude QSAR Des Nouvelles Dérivé De 7-Azaindole Comme C-Met
Kinase Inhibiteurs***

Jury :

Meklid Abdelhak	MCA	Université Mohamed Khider de Biskra	Président
Sellami Mohamed	MCB	Université Mohamed Khider de Biskra	Examineur
Mellaoui Maika	MCB	Université Mohamed Khider de Biskra	Encadreur

Année universitaire : 2020/2021

Dédicace

Je dédie cet humble travail à :

ma chère maman

fekroune Abdhamid

Dédicace

*A ceux qui me sont les plus chers au monde
Mes Parents, que Dieu les protège et bénisse;
A mes frères et sœurs qui m'étaient support et aide;
A tous ceux qui m'ont encouragés et conseillés;
Je dédie ce travail*

ÇACHA ABDELHAK

REMERCIEMENTS

Le travail présenté dans cette thèse a été dirigé par Dr.Mallaoui Malika .

Tout d'abord, nous remercions Dieu le tout puissant pour la volonté, la santé et la patience qu'il nous a données durant toutes ces années d'études.

Nous tenant à remercier sincèrement l'encadreur, Mme Mallaoui Malika pour sa patience, ses encouragements, son orientation, et ses conseils précieux, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu nous consacrer et sans qui ce mémoire n'aurait jamais vu le jour.

Nous souhaitant adresser nos remerciements les plus sincères aux personnes qui nous ont apporté ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de cette formidable année universitaire.

Nous exprimons nos profonds remerciements aux membres de jury qui vont juger notre recherche.

J'adresse, enfin et surtout, ma plus profonde gratitude et tout mon amour à mon père , mes sœurs et mes frères, qui ont su me faire confiance et me soutenir en toutes circonstances, a mes amis Abdeghani ,Samir, Nouaman, Abdelhalim, , Islam, Akram, Lakhdar ,Omar ; Hamza , Fouad, Youssef, Nadji et n'obélie pas Mr.Touahir Abdelraouf et sa souris qui été une grande aide a se travaille .

merci

Sommaire

Introduction générale	08
Références Bibliographique	11
CHAPITRE I : Généralité sur la réactivité de 7- AZAINDOLE	12
I .1. Introduction	13
I .2. 7-Azaindoles :	15
I .3. C-Met Tyrosine Kinases Et Leur Inhibiteurs	16
I .3.1. C-Met Tyrosine Kinase	16
I .3.1. Les Inhibiteurs De La Kinase C-Met	16
I .4. Propretés De 7-Azaindoles	18
I .5. La Synthèse Des Composes Souhaites	19
I .5.1. Synthèse De 7-Azaindoles	19
I .5.2. Synthèse de -4 - ((1H-pyrrolo [2,3-b]pyridin-4-yl) oxy) aniline :	21
I .5.3. Synthèse Des Composés Cibles A Base De Dihydropyridazine7-Azaindole:	22
I .6. Références Bibliographiques	23
CHAPITRE II : Représentation des méthodes de calculs	27
II.1. Intrduction	28
II .2. Chimie Quantique	29
II .2.1.Bases De La Chimie Quantique: Equation De Schrödinger	29
II .3. Mécanique Quantique et Mécanique Moléculaire	30
II .4. Méthodes De Calculs	30
II .4.1. Méthodes Quantiques	30
II.4.2. Différents champs de force en mécanique moléculaire :	32
II.4.3. Dynamique moléculaire	33
II.5. Les relations quantitatives structure-activité/propriété (QSAR)	34
II.5.1. Définition	34
II.5.2. Principe et Méthodologie des méthodes QSAR :	34
II.6. Les Régressions linéaires	41
II.6.1. Régression linéaire multiple:	42
II.7. Références Bibliographiques	44
CHAPITRE. Etude QSAR d'une série de dérivés de7- AZAINDOLE	48
III.1. Introduction	49
III .2. MATÉRIELS ET MÉTHODES	49
III.3. Elaboration Du Modèle QSAR	50
III.3.1. Une Base De Données :	50
III.3.2. Calcul Des Descripteurs	56
III.4.Interprétation D'un Modèle QSAR	60
III.5.Validation De Modèle	62
III.6.Références Bibliographiques	65
Conclusion Générale	66

Liste Des Figures

Chapitre I:

Figure.1: structures de quatre dérivés de Variolin.....	10
Figure.2 : formation de 7-Azaindole.....	12
Figure.3 : Les Types Des Inhibiteurs De Kinase C-Met	14
Figure.4: l'échange de proton.....	15
Figure.5: . La formation de l'indole par la méthode de Fischer.....	16
Figure.6 :formation de 7-azaindole -2-.....	16
Figure.7: la séquence de Madelung.....	17
Figure.8: la synthèse de Wibberley.....	17
Figure 9: la synthèse des 7-azaindoles à partir de la β -picoline.....	18
Figure.10: La synthèse de 4 - ((1-1H-pyrrolo [2,3-b] pyridin-4-yl) oxy) aniline.....	18
Figure.11:la Synthèse des composés cibles.....	19

Chapitre II

Figure.1:Méthodologie QSAR.....	40
---------------------------------	----

Chapitre III

Figure. 1 : Corrélation Entre L'activité Biologique Expérimentale Et Prédite.....	61
Figure 2 Corrélation Entre L'activité Biologique Expérimentale Et Résiduell.....	61

Liste Des Tableaux

Chapitre I:

Tableau 1: Évolution de référence de 2003 à 2013.....	11
---	----

Chapitre III

Tableaux.1 :les structures chimiques et L'activité biologique des dérivés de 7-azaindole	47
Tableaux 2 : les descripteurs obtenu par Hyper-chem 8.0.....	54
Tableaux 3 : les descripteurs obtenu par gaution 09. 01.....	55
Tableaux 4 : les descripteurs obtenu par gaution 09. 02.....	56

Tableaux 5 : Modèles sélectionnés et paramètres statistiques des corrélations entre les propriétés moléculaires et l'activité biologique antiprolifératif.....	57
Tableau 6 : Paramètres de validation croisée.....	59
Tableau 7 : Valeurs Expérimentales , Prédites Et Résiduelles De P-IC50 Des Dérivées De 7-Azaindole.....	60

Liste des Abréviations

RTK: récepteurs tyrosine kinases.

ADN: Acide Désoxyribonucléique.

DFT: Density Functional Theory.

MD : dynamique moléculaire .

MLR: *Multivariate Linear Regression* .

PCR: *Principale Component Regression* .

PLS : *Partial Least Squares* -

ANN : *Artificial Neural Network* .

B3LYP: Lee-Yang-Parr à trois paramètres de Becke.

HF: Hartree-Fock.

HOMO: Orbitale Moléculaire plus Haut Occupée.

LOO : Leave One Out .

LUMO: Orbitale Moléculaire plus bas Occupée .

MLR : Régression linéaire multiple .

P-IC50 : Inhibiteur concentration of 50% .

PRESS: Somme des carrés des erreurs de prédiction .

QSAR : Relations quantitatives structure-activités .

QSPR : Relations quantitatives structure-propriétés .

SPRESS: La capacité prédictive des modèles est évaluée par l'erreur de racine carrée moyenne.

SSE : Erreur de somme des carrés .

SSR : Régression de la somme des carrés .

SSY : Somme des carrés des distances des valeurs observées .

Log P : Coefficient de partage.

R : Coefficient de Corrélation.

R² adj : Coefficient de Détermination Ajusté.

S: Erreur type résiduel.

Introduction générale

Avec l'avènement de la technologie scientifique moderne et de l'industrialisation, l'utilisation de produits chimiques s'accroît parallèlement à l'augmentation du nombre de nouveaux produits chimiques synthétisés chaque jour [1]. Le développement important des moyens informatiques au cours de ces dernières années a permis à la chimie de s'enrichir d'outils informatiques spécialement dédiés à la représentation des molécules. Ainsi un nouvel outil est apparu permettant à tout chimiste de mieux appréhender les problèmes liés au domaine moléculaire : la modélisation moléculaire [2,3]. La modélisation moléculaire regroupe les techniques de visualisation, de manipulation, d'analyse et de calcul de la structure spatiale [4-7]. Aussi elle est de plus en plus utilisée aujourd'hui pour étudier la réactivité chimique [8].

La modélisation moléculaire est un terme général qui englobe différentes techniques de graphisme moléculaire et de chimie computationnelle permettant d'afficher, simuler, analyser, calculer et stocker les propriétés physicochimiques des molécules. La chimie pharmaceutique est une discipline qui s'intéresse à la conception, à la préparation et à l'interprétation du mode d'action des médicaments. C'est une science qui tente, entre autres, d'établir des relations entre la structure chimique des composés et leur activité biologique [9].

Les relations entre les structures des molécules et leurs propriétés ou activités sont généralement établies à l'aide de méthodes de modélisation par apprentissage statistique. Les techniques usuelles reposent sur la caractérisation des molécules par un ensemble de descripteurs, nombres réels mesurés ou calculés à partir des structures moléculaires. Il est alors possible d'établir une relation entre ces descripteurs et la grandeur modélisée [10].

La comparaison de l'activité biologique de certaines molécules et de leurs structures a permis d'établir dans de nombreux cas des corrélations entre les paramètres structuraux et les propriétés d'une molécule. L'association des variations de l'activité biologique aux paramètres structuraux permet d'obtenir un système d'équations qui donne, pour une série chimique donnée et pour une activité définie, une équation de corrélation. L'intérêt essentiel de cette équation est qu'elle doit permettre de déterminer les valeurs des paramètres qui correspondent à une activité biologique maximale et de prédire l'activité des molécules qui n'ont pas encore été synthétisées [11]. La prédiction de l'activité biologique est basée sur des calculs mathématiques et statistiques, c'est pour cette raison qu'on met le choix d'une méthode statistique, appelée la régression multilinéaire (RML) [9].

Qidong Tang et al. [12] ont étudié l'activité biologique d'une série de 7-azaindole avec c-met tyrosine kinase. Dans cette thèse, notre objectif est d'appliquer l'ensemble des méthodes fournies par la chimie informatique pour déterminer les propriétés structurales et électroniques de cette série ; l'effet de

Introduction Générale

substitution sur ces propriétés est également discuté et aussi des corrélations entre la structure chimique et l'activité biologique .

Pour faire ce travail d'une manière convenable on a divisé en trois chapitres; on commence par une introduction général et se termine par une conclusion générale.

- ✚ **Le premier chapitre** : nous exposons des généralités sur la réactivité de 7- AZAINDOLE
- ✚ **Le deuxième chapitre** : contient une Représentation des méthodes de calculs
- ✚ **Le troisième chapitre** : consacré aux résultats obtenus et leur discussion

Références Bibliographiques

1. J.R. Seward, E. Hamblen, T.W. Schultz, Regression comparisons of Tetrahymena pyriformis and Poecilia reticulata toxicity, Chemosphere, 47 (2002) 93-101.
2. E. Derty, J. Mol. Structur. (theochem), 1999, 459, 273
3. S. Belaidi, M. Omari, T. Lanez et A. Dibi, J. Soc. Alger. Chim, 2004, 14, 27.
4. A.R. Leach, Molecular Modelling: Principals and Applications, Edimbourg : Addison Wesley Longman Ltd, 1996, 1-24.
5. H. Chermette, Spectra, 171, 1993, 15-22.
6. G.R. Marshall, Burger's Medicinal Chemistry and Drug Discovery, tm. 1, chap. 15, New York: John Wiley and Sons, 1995, 573-659.
7. L. P. Graham, Chimie pharmaceutique, de boeck, 2003, 319-334.
8. Andricioaei, M. Karplus, Chem. Phys., 115, 6292, 2001.
9. Mehellou Mohammed Nadjib ; Etude des relations structure/activité quantitatives (QSAR/2D) d'une série de dérivés de Triazolothiadiazoles ; Université echahide hamma lakhder d'el-oued; 2018.
10. G-S-A. Aurélie, Thèse de doctorat, Université de Paris 6, 2008.
11. J. Acar, P. Courvalin. La recherche. 314, 1998, 50-52.
12. Qidong Tang, Linxiao Wang, Yongli Duan, Wenhui Wang, Shunmin Huang, Jia Zhi, Shuang Jia, Wufu Zhu, Ping Wang, Rong Luo, Pengwu Zheng, Discovery of novel 7-azaindole derivatives bearing dihydropyridazine moiety as c-Met kinase inhibitors; 2017.

CHAPITRE I :
Généralité sur la réactivité de
7-AZAINDOLE

I.1. Introduction

Le cycle indole est présent dans une gamme de produits naturels tels que les alcaloïdes, les métabolites fongiques et les produits naturels marins, La principale source d'indole elle-même est l'extraction du goudron de houille, bien que la préparation industrielle à partir de réactifs tels que l'aniline, également été démontrée. [1].

Bien que les occurrences naturelles d'azaindoles soient rares par rapport aux indoles, quelques composés naturels contenant de l'azaindole ont été isolés. Les variolins en sont un bon exemple (figure 1). Les variolines sont un groupe d'alcaloïdes qui ont été isolés de l'éponge antarctique *Kirkpatrickia variolosa* [2]

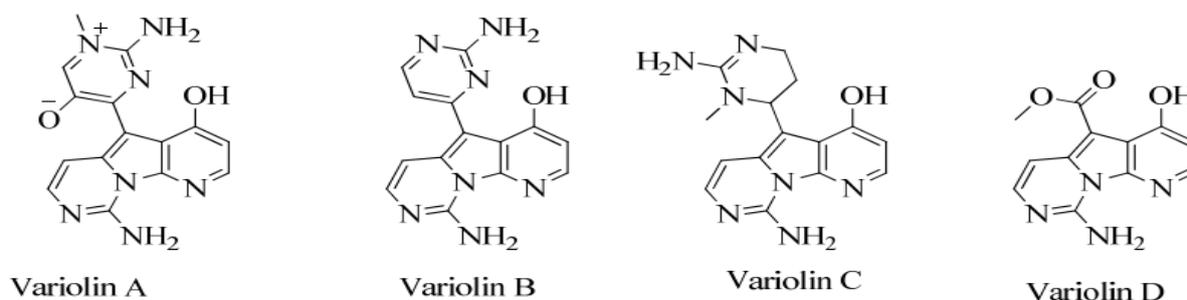


Figure.1: structures de quatre dérivés de Variolin

Même si les azaindoles manquaient d'abondance naturelle, ils ont suscité un intérêt synthétique considérable en tant que bioisostères d'indole[3]. dans lesquels l'adoption d'un atome d'azote supplémentaire dans le cycle aromatique confère des propriétés uniques aux azaindoles. En raison de l'indisponibilité naturelle des azaindoles, la plupart des azaindoles ont été préparé traditionnellement par des méthodes classiques telles que les procédures de Fischer, Madelung et Reissert. [4]. Les propriétés intéressantes affichées par les azaindoles en font des composés attrayants pour les chercheurs en chimie.

Les azaindoles et leurs dérivés présentent des activités biologiques importantes et l'utilisation de ce cadre a contribué à la génération de nouveaux agents thérapeutiques. Les quatre isomères d' azaindole, qui associent un cycle pyridine et pyrrole par une liaison C-C fusionnée, possèdent tous les critères nécessaires pour être d'excellents bioisostères des systèmes indole ou purine. Ces trois entités ne diffèrent que par la substitution d'un fragment CH sp² par un atome d'azote sp² et vice versa. Rares dans la nature [5-8], les azaindoles sont intéressants en termes de stratégies d'optimisation des médicaments. La modification de la règle de Lipinski de cinq, la solubilité, la pK_A et la lipophilicité, la liaison à la cible et les propriétés ADME-tox peuvent être modulées et finement ajustées en utilisant le noyau azaindole au lieu d'autres

CHAPITRE I : Généralité Sur La Réactivité De 7- AZAINDOLE

hétérocycles fusionnés bicycliques [9-18]. Les azaindoles ont été reconnus comme des structures privilégiées dans la modulation des processus biologiques, dans les programmes de chimie médicinale et de découverte de médicaments [19-24]. Leur disponibilité commerciale n'a cessé d'augmenter. L'innovation et synthétique a été continuellement mise à jour. Les portes de recherche en ligne spécialisées pour les structures chimiques donnent facilement une vue globale du domaine de la recherche sur l'azaindole avec de nouvelles synthèses et des structures inconnues émergeant tout le temps. Lorsque l'on considère l'utilisation d'un échafaudage azaindole au lieu d'un échafaudage indole dans les médicaments actifs, l'isomère 5-aza semble d'abord être le plus couramment rencontré en raison de sa forte homologie avec les 5-hydroxy indoles, les principaux métabolites de l'indole présents dans plusieurs biomolécules (par exemple, sérotonine, mélatonine, 5-HIA, 5-OHDPAT), mais cela est trompeur. L'azaindole le plus populaire est indubitablement l'isomère N-7 qui a généré plus de 100 000 structures, le nombre de dérivés disponibles dans le commerce est le double de celui de tous les autres isomères individuellement. C'est donc aussi la structure la plus brevetée (tableau 1). Évolution de référence de 2003 à 2013 [25].

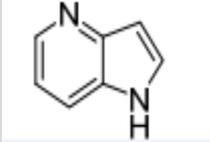
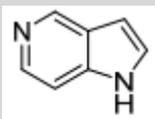
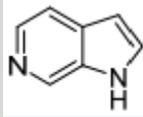
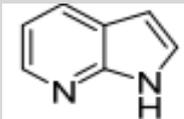
Cadre Azaindole	4-Azaindole	5-Azaindole	6-Azaindole	7-Azaindole
recherche de sous-structures en ligne a				
structures (chimiques)	16,505	13,632	34,762	100,384
disponible dans le commerce	2517	2006 >	2749	4273
Références (brevets)	2158 (1187)	2923 (961)	10,497 (1756)	6576 (2863)

Tableau 1: Évolution de référence de 2003 à 2013.

I.2. 7-Azaindoles :

L'un des azaindoles les plus courants, le 7-azaindole (également connu sous le nom de pyrrolo[1,2-b]pyridine) a été signalé pour la première fois par Perkin[26]. Plus tard, le 7-azaindole a été obtenu à partir de goudron de houille par Kruber, qui a également réussi à préparer des dérivés et obtenu des preuves de la constitution du composé par oxydation de son benzènesulfonamide en acide 2-benzènesulfonamido-nicotinique et conversion de ce dernier en acide 2-aminonicotinique connu[27]. Le 7-azaindole (1-23) contient un atome d'azote dans le cycle pyrrole et un atome d'azote dans le cycle pyridine où les deux cycles sont fusionnés pour donner le composé du titre. Certaines études théoriques ont été réalisées sur les 7-azaindoles et il a été constaté que le cycle pyridine agissait comme un et un -accepteur tandis que le cycle pyrrole agissait comme un donneur et un -accepteur. Des études de spectroscopie de fluorescence ont montré que les réactions de transfert de protons avaient lieu à température ambiante, conduisant à des tautomères de 7-azaindole (1-23) (Figure 2) qui se sont formés très rapidement et aucune émission dimère n'a pu être observée. Cependant, lorsque la température a été abaissée, les émissions du dimère ont été observées lorsque le dimère 7-azaindole 1-24 a donné l'isomère 7-azaindole 1-25[28].

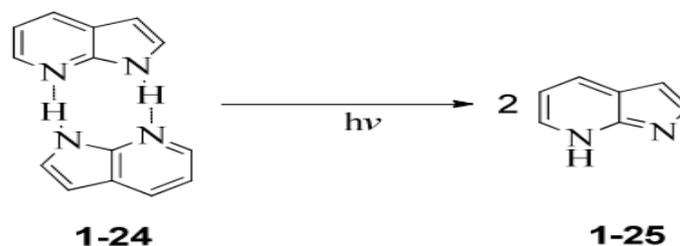


Figure.2 : formation de 7-Azaindole.

7-azaindoles sont l'analogue azoté du système cyclique indolique le plus étudié [29]. se sont révélés être thérapeutiquement actifs pour un large éventail de maladies. Saify et coll. ont synthétisé une série de dérivés de 7-azaindole ayant une réponse antimicrobienne importante contre les bactéries gram positives et gram négatives; Il a été utilisé comme bioisostère d'indole pour améliorer les propriétés physico-chimiques et pharmacocinétiques de plusieurs candidats médicaments tels que les ligands de la dopamine D4 [30-33]. C'est également un pharmacophore émergent dans les inhibiteurs de kinase compétitifs de l'ATP car il contient le motif typique pour s'ancrer dans la poche de liaison à l'adénine [34-35].

Basé en étude sur petites molécules qui inhibes c-met tyrosine kinase un groupe des scientifiques synthétisé une nouvel série de 24 dérivés de 7-azaindole qui sont avérés actifs comme inhibiteurs de c-met kinase avec des valeurs IC₅₀ allant de 1,06 à 56,25 nM [36].

I.3. C-Met Tyrosine Kinases Et Leur Inhibiteurs

I.3.1. C-Met Tyrosine Kinase

Les récepteurs tyrosine kinases (RTK) sont les récepteurs de surface cellulaire de haute affinité pour de nombreux facteurs de croissance polypeptidiques, cytokines et hormones. Sur les 90 gènes uniques de tyrosine kinase identifiés dans le génome humain, 58 sont protéines de récepteur tyrosine kinase[37].

La kinase c-Met est un récepteur tyrosine kinase qui s'exprime dans les cellules endothéliales et épithéliales. Le facteur de croissance des hépatocytes (HGF) est le ligand naturel de c-Met, également connu sous le nom de facteur de diffusion. La voie de signalisation HGF/c-Met joue un rôle important dans le développement normal, l'organogenèse et l'homéostasie [38]. L'activation de c-Met par HGF provoque la dimérisation des récepteurs et l'autophosphorylation de la tyrosine 1234 et 1235. La c-Met phosphorylée déclenche en outre l'activation des voies de signalisation en aval telles que les voies Ras/MAPK, c-Src et PI3K/Akt. Pendant ce temps, c-Met induit un programme invasif consistant en la prolifération, la migration, l'invasion et la survie cellulaires qui sont essentiels pour les événements physiologiques au cours des processus normaux tels que l'orphogénèse, la régénération du foie et la cicatrisation des plaies. La signalisation aberrante HGF/c Met par activation constitutive, amplification génique, mutations et activation d'une boucle autocrine se produit dans pratiquement tous les types de tumeurs solides et, en tant que telle, est impliquée dans de multiples processus oncogènes tumoraux, tels que la mitogènes, la survie, l'angiogénèse et croissance invasive, en particulier dans le processus métastatique [39-42]. De plus, il a été démontré que la surexpression de c-Met et de HGF était en corrélation avec un mauvais pronostic ou une progression métastatique dans un certain nombre de cancers humains majeurs, notamment les cancers du poumon, de la prostate, du rein, de l'ovaire, de l'estomac et du foie [43]. En ce sens, c-Met a récemment suscité un intérêt considérable en tant que cible thérapeutique pour divers cancers [44].

I.3.1. Les Inhibiteurs De La Kinase C-Met

Les inhibiteurs de la kinase c-Met sont regroupés selon leur mode de liaison dans le site de liaison. Les composés avec un mode de liaison en forme de U dans le site de liaison de l'ATP sont des inhibiteurs de la kinase c-Met de type I. En général, les composés de ce type présentent des profils de sélectivité cellulaire relativement bons avec un nombre limité de hits de kinase hors cible, tels que Crizotinib et PF-04217903. Ils rivalisent avec l'ATP pour se lier dans le site de liaison de l'ATP et imitent les interactions entre l'ATP et le site de liaison, y compris les interactions de liaison hydrogène avec les résidus dans la région charnière (Pro1158, Tyr1159 et Met1160). L'interaction avec la boucle d'activation est également

CHAPITRE I : Généralité Sur La Réactivité De 7- AZAINDOLE

importante pour les inhibiteurs de kinase, y compris le motif DFG et Tyr1230 dans la kinase c-Met. Le motif DFG (Asp1222, Phe1223 et Gly1224) est conservé dans la kinase, dont les interactions sont bénéfiques pour améliorer l'activité inhibitrice de la kinase. De plus, Tyr1230 dans le domaine c-Met kinase est unique, de sorte que les composés ayant des interactions avec lui ont montré une bonne sélectivité. Les inhibiteurs de la kinase c-Met de type II sont des inhibiteurs de la kinase c-Met multi-cibles qui passent le portier et occupent la poche arrière hydrophobe profonde. Un mouvement important de la boucle A est nécessaire pour entrer dans la poche arrière c-Met qui est compensé par des interactions supplémentaires avec la poche arrière hydrophobe, conduisant souvent à des composés avec des poids moléculaires élevés et des lipophilies élevées. Mais les interactions de liaison hydrogène avec la région charnière sont les interactions protéine-ligand partagées des inhibiteurs de la kinase c-Met de type I et II. Les inhibiteurs représentatifs de c-Met de type I, II et d'autres types dans différentes études de phase clinique ont été présentés dans la figure 3. En particulier, le crizotinib (c-Met/ALK) et le cabozantinib (VEGFR2/c-Met) ont été approuvés pour le traitement du carcinome pulmonaire non à petites cellules. En outre, d'autres inhibiteurs de la kinase c-Met ont également fait l'objet d'études cliniques, notamment PF-04217903 et JNJ-38877605 de type I, BMS-777607, GSK1363089 et MGCD-265 de type II [45].

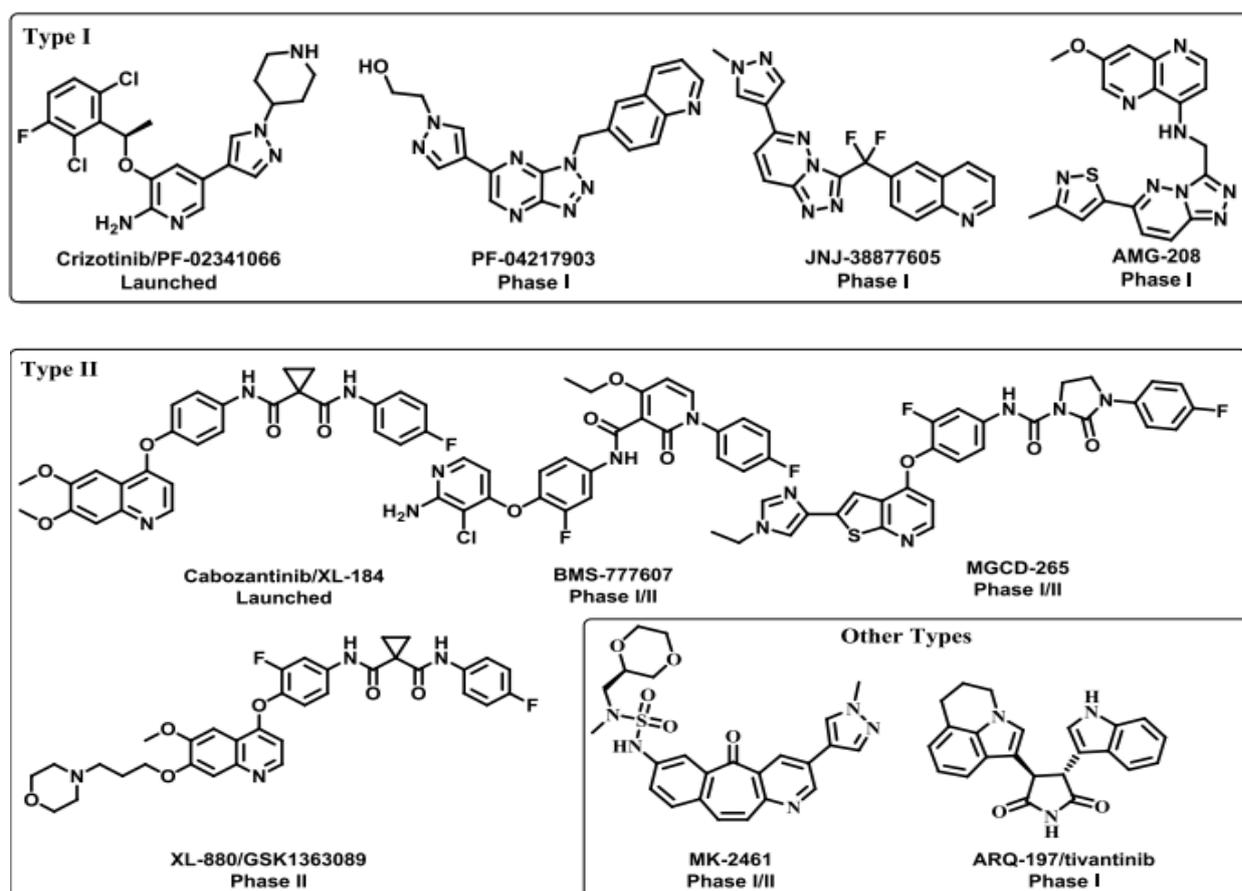


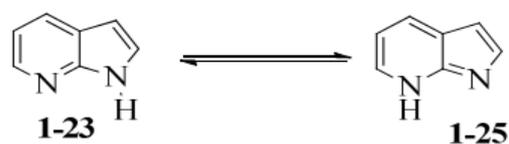
Figure.3 : Les Types Des Inhibiteurs De Kinase C-Met .

I.4. Propriétés De 7-Azaindoles

Les 7-azaindoles sont généralement comparés aux indoles, mais il existe quelques différences dans leurs propriétés physiques et chimiques. Par exemple, 7 azindoles sont des bases plus fortes que les indoles avec un pKa de 4,59 pour l'azaindole 1-23 et ceci est influencé par l'implication dans le système bicyclique condensé de deux cycles hétéroaromatiques azotés avec des propriétés électroniques opposées[28]. Il a été constaté que l'état excité double le transfert de protons du 7-azaindole (1-23) se produit bien dans les solvants hydrocarbures et alcools par la formation de dimères à liaison hydrogène et d'interactions soluté/solvant respectivement. Cela a abouti à l'observation de deux bandes de fluorescence dans les alcools et les solutions d'hydrocarbures concentrées à température ambiante. D'autre part, le comportement spectroscopique de l'azaindole 1-23 dans l'eau était différent de celui des alcools et des hydrocarbures, ne montrant qu'une seule bande de fluorescence et cela pourrait être attribué aux propriétés uniques des 7-azaindoles dans l'eau, leur acidité qui conduit à une molécule d'eau hébergeant jusqu'à quatre liaisons hydrogène et, par conséquent, l'eau se comporte nettement différemment des alcools et des hydrocarbures[46]. Il a également été constaté que la fluorescence du 7-azaindole était relativement faible en raison de la formation de dimères qui à son tour provoquait un double transfert de état excité. La méthylation de l'azote N-1 a empêché la formation de dimères et a augmenté significativement la fluorescence[47].

L'étude des liaisons hydrogène est d'une grande importance en raison de l'importance des liaisons hydrogène en chimie et en biologie. Le dimère 1-24 du 7-azaindole présenté dans le schéma 2 est reconnu comme un modèle simple pour la paire de bases à liaison hydrogène de l'ADN et pourrait fournir des informations sur le rôle possible de la tautomérie dans les mutations[48]. Parce que le 7-azaindole contient deux azotes sur deux cycles différents, la réactivité de l'un ou l'autre de l'azote dépendrait de la sélectivité du site qui à son tour dépendrait des conditions de réaction. Des études ont montré que le 7-azaindole existe sous forme de tautomère comme le montre le Figure 4, mais la forme la plus dominante est le 7-azaindole (1-23) ; Par exemple, dans des conditions vigoureuses, on obtient un échange de protons qui entraîne une dialkylation sur les deux azotes. . La désalkylation ultérieure conduit à l'équilibre avec la formation préférentielle conséquente du produit pyrrole N-alkylé. En revanche, l'alkylation des systèmes anioniques correspondants dans des conditions alcalines douces se produit en douceur à température ambiante, favorisant principalement la N-alkylation du pyrrole[49].

Figure.4: l'échange de proton



I.5. La Synthèse Des Composés Souhaités

I.5.1. Synthèse De 7-Azaindoles :

Les synthèses plus classiquement mises en œuvre pour la préparation des indoles ont été les séquences de Fischer ou de Madelung [48]. Il était donc tentant de les utiliser pour la synthèse du 7-azaindole et de ses dérivés fonctionnels. La formation de l'indole par la méthode de Fischer est réalisée par chauffage ou catalyse en milieu acide de phénylhydrazones (Figure 5).

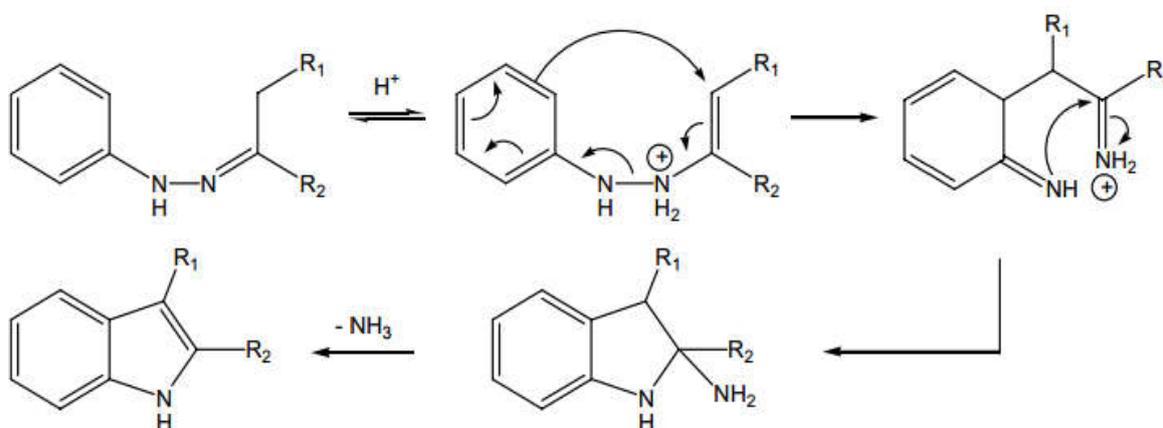


Figure.5: . La formation de l'indole par la méthode de Fischer

Il a été constaté qu'en changeant l'éthylate de sodium en anilide de sodium en présence de formiate de potassium, l'azaindole a été isolé avec des rendements modérés allant jusqu'à 51% [27] (figure 6).

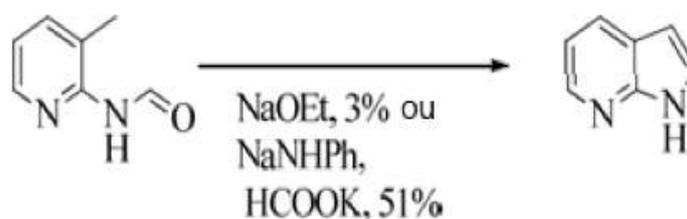


Figure.6 :formation de 7-azaindole -2-

En série indolique, la séquence de Madelung met en jeu des dérivés de la 2-méthylaniline qui sont cyclisés en milieu basique (figure 7) [50].

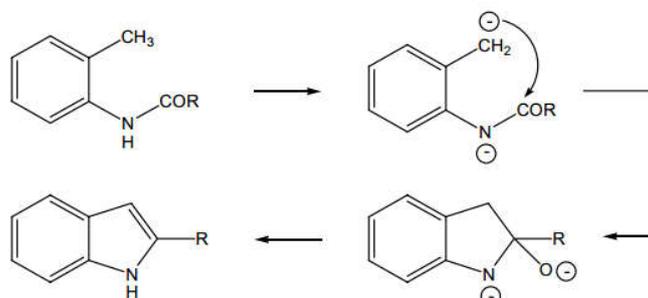


Figure.7: la séquence de Madelung

Le noyau pyrrole constitue un squelette de choix pour la synthèse de 7-azaindoles permettant la formation de composés substitués à la fois sur le noyau pyrrole et sur la partie pyridine, cependant les rendements observés sont modérés. Wibberley ou Quéguiner reportent la synthèse de divers 7-azaindoles à partir de 2-aminopyrroles. Les rendements globaux demeurent moyens (Figure 8) [51-52].

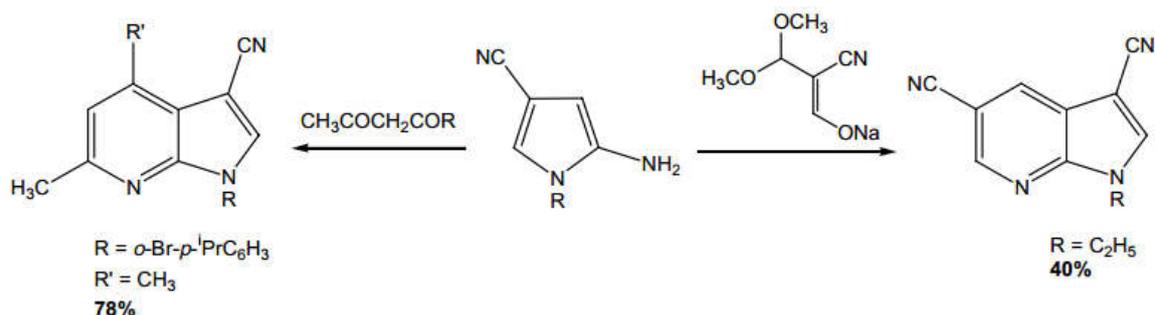


Figure.8: la synthèse de Wibberley

Une autre méthode de synthèse des 7-azaindoles consiste à partir de la β -picoline (1-27). Le groupe méthyle est faiblement acide et peut être déprotoné par le diisopropylamide de lithium (LDA) et ensuite traité avec des électrophiles. Par exemple, le traitement de la β -picoline (1-27) avec du benzonitrile (1-28) en présence de LDA entraîne la formation de 1-phényl-2-(pyridin-3-yl)éthanimine (1-29).

CHAPITRE I : Généralité Sur La Réactivité De 7- AZAINDOLE

Un traitement supplémentaire de 1-29 avec du LDA a donné un intermédiaire 1-30 qui a été traité avec un excès de LDA pour donner un intermédiaire 1-31 qui a été cyclisé en dilithio-7-azaindole 1-32, qui après traitement a donné 2 phényl-7- azaindole (1-33) avec de bons rendements d'environ 90 % (figure 9). Les rendements supérieurs allant jusqu'à 90 % ont été obtenus en utilisant un excès de LDA. Le bilan a impliqué l'élimination de l'hydrure de lithium[53].

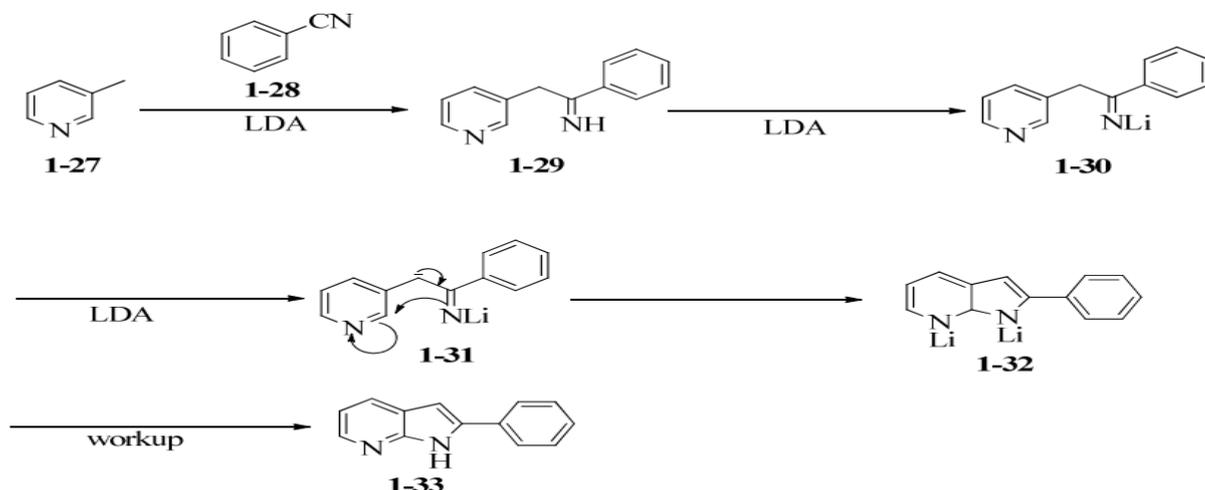


Figure 9: la synthèse des 7-azaindoles à partir de la β -picoline

I.5.2. Synthèse de 4 - ((1H-pyrrolo [2,3-b]pyridin-4-yl) oxy) aniline :

La synthèse des intermédiaires clés du 4 - ((1H-pyrrolo [2,3-b] pyridin-4-yl) oxy) aniline a été obtenue à partir de la 4-chloro-1H-pyrrolo[2,3-b]pyridine ou 4-chloro-1-méthyl-1H-pyrrolo[2,3-b]pyridine qui sont disponible sur la marché comme montré dans Figure 10 [54-55].

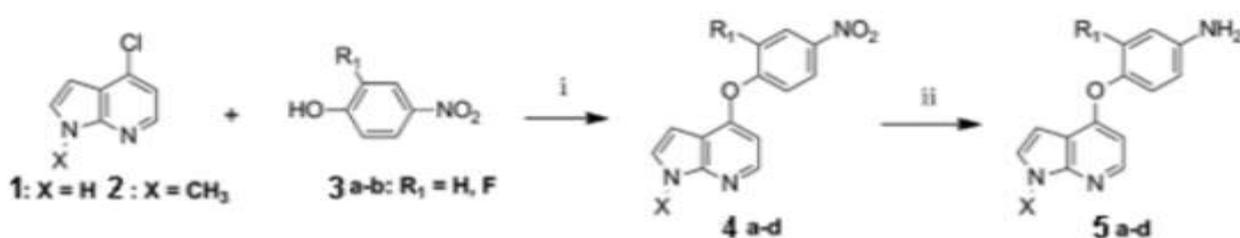


Figure.10: La synthèse de 4 - ((1H-pyrrolo [2,3-b] pyridin-4-yl) oxy) aniline.

Reagents and conditions:

(i) Diphenyl Ether, 190 C°, 50-55% .

(ii) FeCl₃, N₂H₄,H₂O, Carbon active, 80 C°, 77-80% .

I.5.3. Synthèse Des Composés Cibles A Base De Dihydropyridazine7-Azaindole:

Notre 24 composés ont été préparés comme illustré dans Figure 11. En utilisant de l'aniline 6a-l comme matière de départ, une diazo-réaction a été effectuée pour donner l'intermédiaire 7a-l en rendement élevé, qui a été condensé avec de l'acétylacétate d'éthyle pour obtenir intermédiaire 8a-l. La cyclisation a été effectuée pour convertir 8a-l à la pyridazinone 9a-l correspondante, en présence de (éthoxycarbonylméthylène) triphénylphosphorane sous conditions basique [56].

Procédures simples telles que l'hydrolyse et l'acylechloration ont été utilisées pour convertir l'éthyl 4-méthyl-6-oxo-1 carboxylate de phényl-1,6-dihydropyridazine-3 carboxylate 9a-l à le chlorure d'acyle 11a-l correspondant ; les réactions se sont déroulées avec 20% de H₂SO₄ et de chlorure de thionyle, respectivement. Réaction des anilines 5a-d avec le chlorure d'acyle 11a-l favorisée par la DIPEA dans le dichlorométhane à température ambiante a donné les composés nous besoin [57].

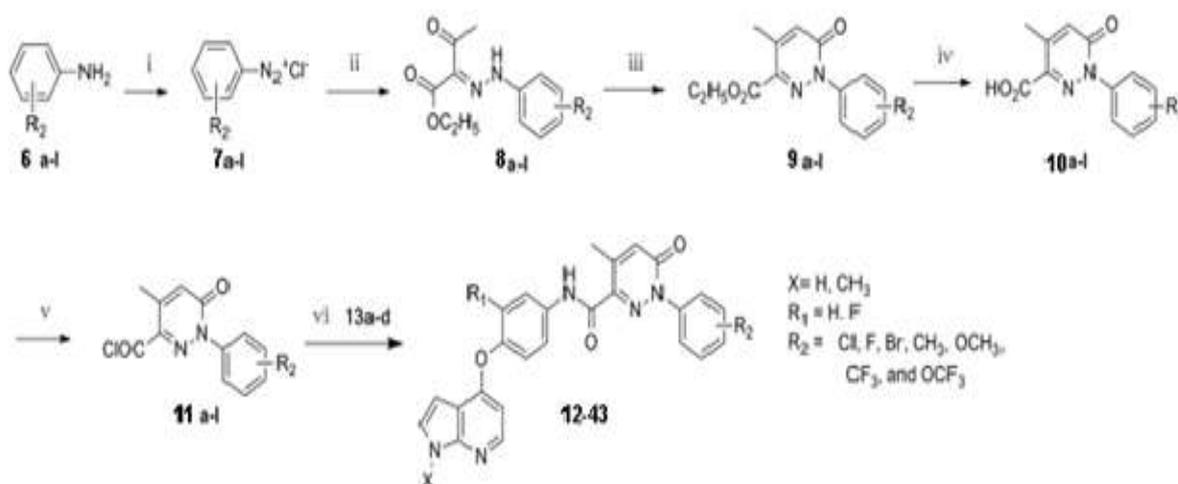


Figure.11:la Synthèse des composés cibles

Réactifs et conditions :

- (i) HCl, NaNO₂, 0 C, 1 h, 90e95%
- (ii) Acétylacétate d'éthyle, CH₃COONa, EtOH/H₂O, 0 C, 0,5 h, r.t., 1,5 h, 86e94 %
- (iii) Ph₃P CHCOOC₂H₅, dioxane, 120 C, 10e15 h, 53e59%
- (iv) 20 % H₂SO₄, 100 C, 5e10 h, 70e76 %; (v) SOCl₂, reflux, 6 h, 92e94%
- (vi) aniline appropriée, chlorure de carbonyle, DIPEA, 0 C, 1 h, rt., 7e10 h, 51e59%

I.6. Références Bibliographiques

1. Sundberg R. J., *Indoles*, Academic Press Limited, New York, 1996
2. Perry N. B., Ettouati L., Litaudon M., Blunt J. W., Munro M. H. G., *Tetrahedron*, 1994, 50, 3987.
(b) Trimurtulu G., Faulkner J. D., *Tetrahedron*, 1994, 50, 3993.
3. Mérour J.-Y., Joseph B., *Curr. Org. Chem.*, 2001, 5, 471.
4. Chi S. M., Choi J.-K., Yum E. K., Chi D. Y., *Tetrahedron Lett.*, 2000, 41, 919
5. Yang, S.-W. ; Abdel-Kader, M. ; Malone, S.; Werkhoven, M.C.M. ; Wisse, J.H. ; Bursuker, I.; Neddermann, K. ; Fairchild, C.; Raventos-Suarez, C.; Menendez, A.T. ; et al. Synthèse et évaluation biologique d'analogues de la cryptolépine, un alcaloïde isolé de la forêt tropicale du Suriname1. *J. Nat. Prod.* 1999.
6. Paulo, A.; Gomes, E.T. ; Houghton, P.J. Nouveaux alcaloïdes de *Cryptolepissanguinolenta*. *J. Nat. Prod.* 1995.
7. Ahaidar, A. ; Fernández, D.; Danelón, G.; Cuevas, C.; Manzanares, I. ; Albericio, F.; Joule, J.A. ; Alvarez, M. Synthèses totales de la varioline B et de la désoxyvarioline B. *J. Org. Chem.* 2003.
8. Choshi, T. ; Yamada, S. ; Sugino, E.; Kuwada, T.; Hibino, S. Synthèse totale des grossularines-1 et -2. *J. Org. Chem.* 1995.
9. Simone, M.; Erba, E.; Damia, G.; Vikhanskaya, F.; di Francesco, A.M.; Riccardi, R.; Bailly, C.; Cuevas, C.; Fernandez Sousa-Faro, J.M.; D'Incalci, M. Variolin B and its derivate deoxy variolin ; B: New marine natural compounds with cyclin-dependent kinase inhibitor activity. *Eur. J. Cancer Oxf. Engl.* 1990 2005.
10. Prudhomme, M. Rebeccamycin analogues as anti-cancer agents. *Eur. J. Med. Chem.* 2003.
11. Zhang, H.-C.; Ye, H.; Conway, B.R.; Derian, C.K.; Addo, M.F.; Kuo, G.-H.; Hecker, L.R.; Croll, D.R.; Li, J.; Westover, L.; et al. 3-(7-Azaindoly)-4-arylmaleimides as potent, selective inhibitors of glycogen synthase kinase-3. *Bioorg. Med. Chem. Lett.* 2004.
12. Grant, E.R.; Errico, M.A.; Emanuel, S.L.; Benjamin, D.; McMillian, M.K.; Wadsworth, S.A.; Zivin, R.A.; Zhong, Z. Protection against glutamate toxicity through inhibition of the p44/42 mitogen-activated protein kinase pathway in neuronally differentiated P19 cells. *Biochem. Pharmacol.* 2001.
13. Dugar, S. Azaindole Derivatives as Inhibitors of p38 Kinase. Patent WO2004032874, 9 October 2003.
14. Showell, G.A.; Emms, F.; Marwood, R.; O'Connor, D.; Patel, S.; Leeson, P.D. Binding of 2,4-disubstituted morpholines at human D4 dopamine receptors. *Bioorg. Med. Chem.* 1998.
15. Kulagowski, J.J.; Broughton, H.B.; Curtis, N.R.; Mawer, I.M.; Ridgill, M.P.; Baker, R.; Emms, F.; Freedman, S.B.; Marwood, R.; Patel, S.; et al. 3-((4-(4-Chlorophenyl)piperazin-1-yl)-methyl)-

CHAPITRE I : Généralité Sur La Réactivité De 7- AZAINDOLE

- 1Hpyrrolo-2,3-b-pyridine: An antagonist with high affinity and selectivity for the human dopamine D4 receptor. *J. Med. Chem.* 1996.
16. Fonquerna, S.; Miralpeix, M.; Pagès, L.; Puig, C.; Cardús, A.; Antón, F.; Vilella, D.; Aparici, M.; Prieto, J.; Warreallow, G.; et al. Synthesis and structure-activity relationships of piperidinylpyrrolopyridine derivatives as potent and selective H1 antagonists. *Bioorg. Med. Chem. Lett.* 2005.
17. Gong, Y.; Becker, M.; Choi-Sledeski, Y.M.; Davis, R.S.; Salvino, J.M.; Chu, V.; Brown, K.D.; Pauls, H.W. Solid-phase parallel synthesis of azarene pyrrolidinones as factor Xa inhibitors. *Bioorg. Med. Chem. Lett.* 2000.
18. Sanderson, P.E.J.; Stanton, M.G.; Dorsey, B.D.; Lyle, T.A.; McDonough, C.; Sanders, W.M.; Savage, K.L.; Naylor-Olsen, A.M.; Krueger, J.A.; Lewis, S.D.; et al. Azaindoles: Moderately basic P1 groups for enhancing the selectivity of thrombin inhibitors. *Bioorg. Med. Chem. Lett.* 2003, 13, 795–798. 6. N.R. Mohamed, M.T. El-Saidi, Y.M. Ali, M.H. Elnagdi, *J. Heterocycl. Chem.* 44 (2007) .
19. Horton, D.A.; Bourne, G.T.; Smythe, M.L. The Combinatorial Synthesis of Bicyclic Privileged Structures or Privileged Substructures. *Chem. Rev.* 2003 .
20. De Sá Alves, F.R.; Barreiro, E.J.; Fraga, C.A.M. From nature to drug discovery: The indole scaffold as a “privileged structure”. *Mini Rev. Med. Chem.* 2009.
21. Sharma, V.; Kumar, P.; Pathak, D. Biological importance of the indole nucleus in recent years: A comprehensive review. *J. Heterocycl. Chem.* 2010.
22. Perry, N.B.; Ettouati, L.; Litaudon, M.; Blunt, J.W.; Munro, M.H.G.; Parkin, S.; Hope, H. Alkaloids from the antarctic sponge *Kirkpatrickia variolosa*: Part 1: Variolin b, a new antitumour and antiviral compound. *Tetrahedron* 1994.
23. Welsch, M.E.; Snyder, S.A.; Stockwell, B.R. Privileged scaffolds for library design and drug discovery. *Curr. Opin. Chem. Biol.* 2010.
24. Walker, S.R.; Carter, E.J.; Huff, B.C.; Morris, J.C. Variolins and related alkaloids. *Chem. Rev.* 2009.
25. Article de Pascal Bonnet et de Jean YVES Merour :The Azaindole Framework in the Design of Kinase Inhibitors Molecules ;Université d'Orléans décembre ;2014.
26. Perkin Jr. W. H., Robinson J. R., *J. Chem. Soc.*, 1912, 1175.
27. Robinson M. M., Robinson B. L., *J. Am. Chem. Soc.*, 1955, 77, 457.
28. Mérour J.-Y., Joseph B., *Curr. Org. Chem.*, 2001, 5, 471.
29. Saify ZS, Mozzam SM, Nisa M, Khan SA, Ahmed A, Haider S, Aryne A, Khanam M, Arshad N, Ghani M. 7-Azaindole derivatives as potential antibacterial agents. *Pak. J. Sci. Ind. Res.* 2009;52(1):1- 7.

CHAPITRE I : Généralité Sur La Réactivité De 7- AZAINDOLE

30. Saify ZS. 7-Azaindole derivatives as potential anti-inflammatory agents. *Pak. Journal of Pharmacology*. 1986;2:43-46.
31. Saify ZS. Synthesis of some azaindole derivatives of potential biological interest; Significance of physical parameters in drug design. *J. of Pharmacy, University of Karachi*. 1984;2(2):99-103.
32. Saify ZS. The effect of substituted nicotinamide and azaindoles on lipid peroxidation and platelet aggregation. *Proceedings of second International Conference on Pharmaceutical Sciences, Karachi*; 1997.
33. Wang T, Zhang Z, Wallace OB, Deshpande M, Fang H, Yang Z, Zadjura L, Tweedie DL, Huang S, Zhao F, Ranadive S, Robinson B, Gong YF, Ricardi K, Spicer TP, Deminie C, Rose R, Wang H, Blair WS, Shi PY, Ling P, Colonna RJ, Meanwell NA. Discovery of 4-benzoyl-1-[(4-methoxy-1H-pyrrolo[2,3-b]pyridin-3-yl)oxoacetyl]-2-(R)-methylpiperazine (BMS-378806): a novel HIV-1 attachment inhibitor that interferes with CD4-gp120 interactions. *J. Med. Chem.* 2003;46:4236. (b) Curtis NR, Kulagowski JJ, Leeson PD, Ridgill MP, Emms F, Freedman SB, Patel S, Pate S. Synthesis and SAR of 2- and 3- substituted 7-azaindoles as potential dopamine D4 ligands. *Bioorg. Med. Chem. Lett.* 1999;9:585.
34. Henry JR, Dodd JH. Synthesis of RWJ 68354: A potent inhibitor of the MAP kinase p38. *Tetrahedron Lett.* 1998;48: 8763. (b) Dumas, J. *Exp. Opin. Ther. Patents*. 2001;11:405.
35. 7-azaindoles, use thereof as phosphodiesterase 4 inhibitors and method for producing the same. US 7169787 B2.
36. Qidong Tang, Linxiao Wang, Yongli Duan, Wenhui Wang, Shunmin Huang, Jia Zhi, Shuang Jia, Wufu Zhu, Ping Wang, Rong Luo, Pengwu Zheng, Discovery of novel 7-azaindole derivatives bearing dihydropyridazine moiety as c-Met kinase inhibitors; 2017
37. Robinson DR, Wu YM, Lin SF (November 2000). "The protein tyrosine kinase family of the human genome". *Oncogene*. 19 (49): 5548–57. doi:10.1038/sj.onc.1203957. PMID 11114734
38. M.F. Di Renzo, R.P. Narsimhan, M. Olivero, S. Bretti, S. Giordano, E. Medico, P. Gaglia, P. Zara, P.M. Comoglio, Expression of the Met/HGF receptor in normal and neoplastic human tissues, *Oncogene* 6 (1991) 1997-2003.
39. E. Gherardi, W. Birchmeier, C. Birchmeier, G. Vande Woude, Targeting MET in cancer: rationale and progress, *Nat. Rev. Cancer* 12 (2012) 89-103.
40. C. Birchmeier, W. Birchmeier, E. Gherardi, G.F. Vande Woude, Met, metastasis, motility and more, *Nat. Rev. Mol. Cell Biol.* 4 (2003) 915-925.
41. C. Boccaccio, P.M. Comoglio, Invasive growth: a MET-driven genetic programme for cancer and stem cells, *Nat. Rev. Cancer* 6 (2006) 637-645.
42. P.C. Ma, M.S. Tretiakova, A.C. MacKinnon, N. Ramnath, C. Johnson, S. Dietrich, T. Seiwert, J.G. Christensen, R. Jagadeeswaran, T. Krausz, E.E. Vokes, A.N. Husain, R. Salgia, Expression

and mutational analysis of MET in human solid cancers, *Genes Chromosomes Cancer* 47 (2008) 1025-1037.

43. J.J. Cui, Targeting receptor tyrosine kinase MET in cancer: small molecule inhibitors and clinical progress, *J. Med.Chem.* 57 (2014) 4427-4453.
44. S. Peters, A.A. Adjei, MET: a promising anticancer therapeutic target, *Nat. Rev. Clin. Oncol.* 9 (2012) 314-326.
45. Haoliang Yuan, Qiufeng Liu,cde, Li Zhang,a, Shihe Hu, Tiantian Chen, Huifang Li, Yadong Chen, Yechun Xuc and Tao Luab,Discovery, Optimization and Biological Evaluation for Novel c-Met kinase Inhibitors; 2017.
46. Chou P.-T., Martinez M. L., Cooper W. C., *J. Phys. Chem.*, 1992, 96, 5203
47. Wang K., Stringfellow S., Dong S., Jiao Y., *Spectrochim. Acta Part A*, 2002, 58, 2595.
48. Fuke K., Yoshiuchi H., Kaya K., *J. Phys. Chem.*, 1984, 88, 5840.
49. Mahadevan I., Rasmussen M., *Tetrahedron*, 1993, 49, 7337.
50. Lorenz, R. R. ; Tullar, B. F. ; Koelsch, C. F.; Archer, S. J. *Org. Chem.* 1965, 30, 2531.
51. Brodrick, A. ; Wibberley, D. G. *J. Chem. Soc. Perkin Trans. I* 1975, 1910 ;
52. (b) Zhang, Y. ; Donahue, J. P. ; Li C.-J. *Org. Lett.* 2007, 9, 627 ; (c) Lee, K. L. ; Foley, M. A. ; Chen, L. ; Behnke, M. L. ; Lovering, F. E. ; Kirincich, S. J. ; Wang, W. ; Shim, J.
53. Davies M. L., Wakefield B. J., Wardell J. A., *Tetrahedron*, 1992, 48, 939.
54. S. Raepfel, F. Raepfel, S.W. Claridge, L.J. Zhan, F. Gaudette, M. Mannion, N. Sato, Y. Yuki, M. Kishida, A. Vaisburg, US 2011.
55. X.W. Chen, Y.Y. Sun, L. Fu, J.Q. Li, *Eur. J. Med. Chem.* 123 (2016).
56. N.R. Mohamed, M.T. El-Saidi, Y.M. Ali, M.H. Elnagdi, *J. Heterocycl. Chem.* 44 (2007) .
57. Q.D. Tang, Y.F. Zhao, X.M. Du, L.E. Chong, P. Gong, C. Guo, *Eur. J. Med. Chem.* 69 (2013) .

CHAPITRE II :

Représentation des méthodes de calculs

II.1. Introduction

La modélisation moléculaire implique l'utilisation de méthodes de calculs théoriques (mécanique moléculaire, dynamique moléculaire, mécanique quantique abinitio ou semi-empirique, ...) permettant de déterminer la représentation graphique de la géométrie et d'évaluer les propriétés physico-chimiques de la molécule étudiée. La modélisation moléculaire permet d'interpréter des phénomènes physico-chimiques, de suggérer des nouvelles expériences et d'analyser ainsi des résultats d'une façon plus critique que les expériences classiquement utilisées.

Les fondements de la Chimie Quantique sont apparus et se sont beaucoup développés dans les années 1920 grâce à des scientifiques comme Bohr, Schrödinger, Born, Oppenheimer, Hartree ou encore Slater. En 1930, Hartree et Fock développèrent la méthode du champ autocohérent qui permit d'effectuer les premiers calculs ab initio sur des systèmes diatomiques. Il fallut attendre les années 1950 pour que l'ordinateur fasse son apparition. En 1964, Hohenberg et Kohn ont défini un théorème, la théorie de la fonctionnelle de la densité (DFT). En 1970, Pople créa Gaussian qui est aujourd'hui le logiciel de modélisation de très loin le plus utilisé. Les méthodes semi empiriques et l'utilisation des bases de fonction gaussiennes se sont particulièrement développées dans les années 1970 et 1980. Avec la forte augmentation de la puissance de calcul, la modélisation s'est invitée dans nos ordinateurs à partir des années 1990. En 1993 apparaît la méthode B3LYP, méthode hybride qui permet d'effectuer des calculs DFT. en 1998, le prix Nobel de chimie a été décerné à John. A. Pople et Walter Kohn pour leurs travaux dans le domaine de la chimie informatique et la modélisation moléculaire (chimie quantique) [1-2].

Et avec Le Développement important des moyens informatiques (mémoire plus importante, calcul scientifique, infographie, ...) au cours de ces dernières années a permis à la chimie de s'enrichir d'outils informatiques spécialement dédiés à la représentation des molécules[3].

Ainsi un nouvel outil est apparu permettant à tout chimiste de mieux appréhender les problèmes liés au domaine moléculaire "la Modélisation Moléculaire".

II.2. Chimie Quantique

Différentes approches sont envisageables dans le cadre des outils de modélisation moléculaire. Si ceux de mécanique classique, économiques en termes de temps de calcul, permettent de traiter des systèmes moléculaires de grande taille, les méthodes quantiques (ab-initio, semi empiriques ou théorie de la fonctionnelle de la densité) sont, quant à elles, capables de calculer les propriétés électroniques des systèmes. C'est pour cette raison que ces approches ont été employées dans le cadre de cette étude. Des compléments sur les méthodes décrites ici peuvent être trouvés dans différents ouvrages de références [4-10].

II.2.1. Bases De La Chimie Quantique: Equation De Schrödinger

Les méthodes de chimie quantique, présentées par la suite, reposent toutes sur l'équation de Schrödinger, non relativiste, indépendante du temps.

$$\hat{H}\Psi = E\Psi$$

où Ψ est la fonction d'onde décrivant le système de noyaux et d'électrons, \hat{H} est l'opérateur Hamiltonien relatif à ce même système et E est l'énergie correspondante, valeur propre de l'équation.

L'Hamiltonien rend compte des différentes contributions à l'énergie totale du système à partir d'opérateurs pour les énergies cinétiques des électrons et des noyaux ainsi que les interactions noyau-électron, électron-électron et noyau-noyau.

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_k \frac{1}{2} \nabla_k^2 - \sum_i \sum_k \frac{Z_k}{r_{ik}} + \sum_i \sum_{i < j} \frac{1}{r_{ij}} + \sum_k \sum_{k < l} \frac{Z_k Z_l}{r_{kl}}$$

où ∇^2 est l'opérateur d'énergie cinétique, Z_k est le numéro atomique de l'atome k , r_{ik} est la distance entre un électron i et un noyau k , r_{ij} est la distance entre deux électrons i et j et r_{kl} est la distance entre deux noyaux k et l .

Il n'est pas possible de résoudre cette équation pour des systèmes d'intérêt chimique (au-delà de H_2), de manière exacte. Il faut alors introduire différentes approximations : Approximation de Born-Oppenheimer (considérer les noyaux fixes et leur énergie cinétique est négligeable) ; Approximation orbitale [11].

II .3. Mécanique Quantique et Mécanique Moléculaire

Les méthodes théoriques qui permettent de calculer les données relatives à la structure et aux propriétés d'une molécule peuvent être subdivisées en deux catégories : la mécanique moléculaire et la mécanique quantique (semi-empirique, ab initio et DFT) [12] .

Il existe deux approches différentes pour calculer l'énergie d'une molécule. Premièrement, la mécanique quantique, une procédure basée sur des principes premiers. Dans cette approche, les noyaux sont disposés dans l'espace et les électrons correspondants sont répartis dans tout le système dans une densité électronique continue et calculés en résolvant l'équation de Schrödinger. Lorsque les réactions chimiques n'ont pas besoin d'être simulées, la mécanique classique peut décrire le comportement d'un système biomoléculaire . Ce modèle mathématique est connu sous le nom de mécanique moléculaire et peut être utilisé pour calculer l'énergie de systèmes contenant un grand nombre d'atomes, tels que des molécules ou des systèmes complexes d'intérêt biochimique et biomédical .

Contrairement à la mécanique quantique, la mécanique moléculaire ignore les électrons et ne calcule l'énergie d'un système qu'en fonction des positions nucléaires. Ensuite, il est possible de prendre en compte de manière implicite la composante électronique du système par un paramétrage adéquat de la fonction énergie potentielle. L'ensemble des équations et des paramètres qui définissent la surface potentielle d'une molécule est appelé champ de force [13] .

II .4. Méthodes De Calculs

II .4.1. Méthodes Quantiques

➤ Les Méthodes Ab-Initio :

Les méthodes ab-initio résolvent exactement l'équation [12]. dans une base d'orbitales atomiques données. Alors que les éléments de la matrice de cœur H_{cc} se calculent aisément, le nombre des intégrales multicentriques (qui croît comme la puissance 4 du nombre d'orbitales atomiques) et leur forme mathématique posent des problèmes délicats, notamment en termes de temps de calcul.

Plus le choix des fonctions de base est judicieux et plus leur nombre est grand, plus le calcul est exact. Avec une station de travail moyenne, on traite, en un temps raisonnable, des systèmes contenant au maximum une vingtaine d'atomes lourds. Ce nombre varie évidemment avec le type de base utilisé et les caractéristiques de l'ordinateur.

CHAPITRE II : Représentation Des Méthodes De Calculs

➤ Méthodes Semi-Empiriques :

Les méthodes semi-empiriques sont dérivées des méthodes ab-initio. Elles considèrent deux approximations dont la première consiste à prendre en considération que les électrons de valence. La deuxième approximation, nommée RDN (Recouvrement Différentiel Nul) ou NDO (Neglect of Differential Overlap) a été proposée en 1952 par Pann dans le but de réduire le nombre d'intégrales à calculer, en particulier les intégrales de Répulsion bioélectroniques multicentrées [14-16]. On distingue plusieurs variantes :

CNDO : (Complete Neglect of Differential Overlap) [17] proposée par Pople, Segal et Santry en 1965. Elles négligent complètement les intégrales de répulsions entre les atomes Non chimiquement liés. Son défaut majeur réside dans son échec pour le traitement des Systèmes à couche ouverte.

INDO : (Intermediate Neglect of Differential Overlap) [17]: proposée par Pople, Beveridge et Dobosh en 1967. C'est une correction de la méthode CNDO. Elle permet de distinguer entre les états singulets et les états triplets d'un système en conservant les intégrales d'échange.

MINDO/3 : (Modified Intermediate Neglect of Differential Overlap) [18-19] proposée par Bingham, Dewar et Lo en 1975. La paramétrisation est effectuée en se référant aux résultats expérimentaux et non pas aux résultats ab-initio. Cependant, elle surestime le calcul exact des chaleurs de formation, des potentiels d'ionisation et des distances interatomiques dans la plupart des molécules de dimension moyenne.

MNDO : (Modified Neglect of Diatomic Overlap): Proposée par Dewar et Thiel en 1977 [20] Cette méthode est basée sur l'approximation NDDO (Neglect of diatomic Differential Overlap). Elle néglige le recouvrement différentiel entre orbitales atomiques sur des atomes différents. Elle a été élaborée de façon à reproduire les grandeurs physicochimiques caractérisant les états fondamentaux d'un grand nombre de composés organiques et minéraux. Les enthalpies de formation, les chaleurs de réactions, les barrières de rotation, les potentiels d'ionisation, les moments dipolaires et les géométries moléculaires calculées par MNDO sont en bon accord avec les données expérimentales.

AM1 : (Austin Model 1): Proposée par Dewar en 1985 [21] c'est une amélioration de la Méthode MNDO. Elle diffère de MNDO dans l'expression pour le calcul des répulsions cœur cœur (noyau-électrons internes).

PM3 : (Parametric Method 3): Proposée par Stewart en 1989 [22]. PM3 est la troisième paramétrisation de MNDO [22]; AM1 étant considérée la seconde. Elles négligent généralement le calcul de certaines intégrales et les remplacent par des paramètres expérimentaux.

➤ La Théorie De La Fonctionnelle De La Densité DFT

Au cours des trente dernières années, la Théorie de la Fonctionnelle de la Densité DFT (Density Functional Theory), venant compléter les méthodes HF et Post HF, s'est révélée comme particulièrement performante. Cette approche présente plusieurs avantages qui facilitent son utilisation dans différents domaines, surtout en chimie et sciences des matériaux.

- D'abord, elle inclut dans son formalisme une grande part de la corrélation électronique.
- L'aspect mono-électronique permet de donner une interprétation chimique de la fonction d'onde.

Dans le cas des méthodes Hartree-Fock, il s'agissait d'exprimer l'énergie du système en fonction de la fonction d'onde Ψ , alors que dans le cas de la DFT, l'énergie sera une fonctionnelle de la densité électronique ρ du système [23].

II.4.2. Différents champs de force en mécanique moléculaire :

Différents champs de force sont proposés dans la littérature, ils se distinguent les uns des autres par les termes dans le développement de l'expression de l'énergie de la molécule.

Chacun a un domaine d'application spécifique de sorte que le choix d'un champ de force dépend des propriétés et de l'application du système que l'on veut étudier. Les champs de forces en MM peuvent être groupés en trois classes principales [24].

- Champs de force contenant uniquement les termes harmoniques.
- Champs de force utilisant les termes d'ordre supérieur (cubique, quadratique,...).
- Champs de force suggérés par Allinger et col. [25] ne considérant pas que les termes de la mécanique moléculaire classique mais aussi les effets chimiques comme l'électronégativité.

MM2: Il a été développé par Allinger en 1976 et c'est le champ de force le plus utilisé par la communauté des chimistes organiciens [26]. Il a été conçu au début pour les molécules simples (alcane, alcènes et alcynes non conjugués, les composés carbonyles, les sulfures, les amines...),

mais ses version améliorées MM3 (1989) [27] et MM4 (1996) [28] lui permet de traiter des molécules organiques de plus en plus complexes.

MM+: Est une extension du champ de force MM2, avec l'ajout de quelques paramètres additionnels [29]. MM+ est un champ de force robuste, il a l'aptitude de prendre en considération les paramètres négligés dans d'autres champs de force et peut donc s'appliquer pour des molécules plus complexes tels que les composés inorganiques [30]

AMBER: (Assisted Model Building with Energy Refinement), ce champ de force a été paramétrisé pour les protéines et les acides nucléiques. Il a été utilisé pour les polymères et pour d'autres petites molécules [31].

OPLS: (Optimised Potentials for Liquid Simulation), comme l'indique son nom, est conçu pour optimiser le potentiel qui permet la description des propriétés de solvation [32-33].

GROMOS: (Groningen Molecular Simulation Program Package), est conçu spécialement pour les biomolécules en milieu aqueux en vue de l'étude des interactions entre les molécules d'eau et les groupements polaires des protéines [34].

CHARMM (Bio+): (Chemistry Harvard Macromolecular Mechanic), Son concept est semblable à celui d'AMBER [35].

MMFF94: C'est l'un des champs de force de mécanique moléculaire les plus récents, il a été développé par Tom Halgren [36-37]. Les termes d'énergie d'élongation, de valence et de torsion sont les mêmes que ceux de MM3 (expansion de MM2). MMFF94 comprend des termes croisés de liaison-valence [38].

II.4.3. Dynamique moléculaire :

La dynamique moléculaire (MD) est une technique de calcul dans laquelle des configurations successives d'un système sont générées en intégrant la loi du mouvement de Newton. Le résultat est une trajectoire qui décrit comment les positions et les vitesses des particules dans le système varient avec le temps [39] Les simulations MD de macromolécules biologiques fournissent des détails atomiques sur les mouvements internes de ces systèmes. Des améliorations constantes de la méthodologie et de la puissance de calcul ont étendu l'utilisation des études de dynamique moléculaire à des systèmes plus grands, y compris, par exemple, un environnement explicite de solvant et/ou de membrane, des changements de conformation plus importants et des échelles de temps plus longues. [40] Les applications de conception de médicaments des simulations de dynamique moléculaire comprennent l'estimation des énergies libres de liaison, la prédiction de la sélectivité de la cible, la génération de plusieurs conformations pour un amarrage flexible, le développement de pharmacophores dynamiques à base de protéines, la prédiction du métabolisme et le raffinement des modèles d'homologie des protéines [41].

II.5. Les relations quantitatives structure-activité/propriété (QSAR)

II.5.1. Définition :

Depuis le milieu du 19^{ème} siècle et à partir de la fin de 20^{ème} siècle, une appréciation que l'activité biologique est liée à la structure chimique surtout aux propriétés physicochimiques est développée[42]. Une Relation Quantitative Structure à Activité (en anglais : Quantitative StructureActivity Relations hipou QSAR) est le procédé par lequel une structure chimique est corrélée avec un effet bien déterminé comme l'activité biologique ou la réactivité chimique. Ainsi, l'activité biologique peut être exprimée de manière quantitative, comme pour la concentration de substance nécessaire pour obtenir une certaine réponse biologique. De plus lorsque les propriétés ou structures physicochimiques sont exprimées par des chiffres, on peut proposer une relation mathématique, ou Relation Quantitative Structure à Activité, entre les deux. L'expression mathématique obtenue peut alors être utilisée comme moyen prédictif de la réponse biologique pour des structures similaires [43].

La QSAR la plus commune est de la forme : activité = f (propriétés physico chimiques et/ou structurales)

Les relations quantitatives structure-activité/propriété (QSAR/QSPR) sont de plus en plus utilisées, du fait de la croissance des moyens de calculs. Très récemment, la mise en place du nouveau règlement européen REACH, qui recommande leur utilisation pour limiter le recours à l'expérience, donne un nouvel essor au développement de tels modèles prédictifs.

En fait, les premiers développements dans le sens de telles méthodologies sont plutôt anciens. Dès 1868, Crum-Brown et Fraser ont postulé l'existence de relations entre les activités physiologiques et les structures chimiques [41] en reliant les changements d'activité biologique à des modifications structurales simples, ne disposant alors pas de moyen pour caractériser les structures chimiques en termes quantitatifs.

Une avancée importante vers les modèles QSPR proprement dits a été réalisée grâce au développement des équations de Hammett [45-47] dans lesquelles les constantes du même nom σ caractérisent de manière quantitative les vitesses de réactions pour les composés organiques .

II.5.2. Principe et Méthodologie des méthodes QSAR :

Le principe des méthodes QSAR est de mettre en place une relation mathématique reliant de manière quantitative des propriétés moléculaires aussi bien électroniques que géométriques, appelées descripteurs,

avec une activité biologique/toxicité, pour une série de composés chimiques similaires à l'aide de méthodes d'analyses de données en l'occurrence les régressions linéaires simple et multiple (SLR, MLR), régressions aux moindres carrés partiels (PLS) [48] méthode d'analyse en composantes principales (PCA) [49].

Pour aboutir à un modèle QSAR explicatif et prédictif on doit :

➤ Elaboration du modèle QSAR :

Le développement d'un modèle commence par la collecte de données expérimentales fiables et en nombre important (Base de données); Le choix de la base de données expérimentale initiale est une étape critique pour le développement des modèles QSAR .

Les modèles QSPR sont généralement plus efficaces pour un jeu de molécules similaires. Pareillement, il est important de considérer une base de données dont les molécules se comportent de manière similaire vis-à-vis de la propriété ciblée. De même, afin de réduire la probabilité de rencontrer des molécules très différentes de la majorité des autres, qui peuvent peser significativement sur le modèle et accroître l'incertitude sur les prédictions (Outliers) dans le jeu de données, plus la diversité structurale d'une base de données est importante et plus la taille de la base de données doit être importante [50].

Ensuite il est nécessaire de générer un nombre de descripteurs, caractérisant les structures moléculaires, Un descripteur moléculaire est toute propriété (une valeur numérique) propre à une structure chimique donnée. Ces valeurs peuvent être obtenues expérimentalement ou calculées à partir de la structure de la molécule.

L'activité biologique ne peut pas être reliée directement à la structure moléculaire ; cette dernière est codée par des grandeurs qui représentent d'une manière quantitative les informations contenues dans la structure moléculaire telles que les caractéristiques physicochimiques et structurales. Ces grandeurs sont les descripteurs. Une fois ces derniers disponibles, il est possible d'établir des relations entre ceux-ci et une activité à l'aide des outils de la modélisation (Méthodes statistiques) [51].

Dans ce qui suit, nous allons présenter que les descripteurs moléculaires qui ont été utilisés dans l'ensemble de nos travaux [52]:

Le poids moléculaire : noté MW (appelé aussi le poids de formule), mesuré en daltons (Da). C'est la somme des poids atomiques des différents atomes constituant la molécule. Il est utilisé dans l'étude de

CHAPITRE II : Représentation Des Méthodes De Calculs

transport dont la diffusion et le mode de fonctionnement. Les composés avec des poids plus élevés sont moins susceptibles d'être absorbés et donc ne peuvent pas atteindre le site d'action. Ainsi, essayer de garder des poids moléculaires aussi bas que possible devrait être l'objectif pour établir un médicament. Pour les médicaments délivrés par voie orale le poids moléculaire doit être inférieur ou égal à 500 daltons (optimum autour de 300 daltons) .

Le volume moléculaire : noté MV, en cm³, est défini par la formule suivante :

$$MV = \frac{MW}{d}$$

Avec : MW est le poids moléculaire et d la densité.

Le coefficient de partage Octanol/Eau: Le transport, le passage à travers les membranes et l'activité pharmacologique d'une molécule peuvent être conditionnés par son partage entre une phase lipidique et une phase aqueuse, c'est-à-dire son caractère hydrophile.

Celui-ci peut être quantifié par le coefficient de partage Octanol-Eau, noté (log P), qui mesure la solubilité différentielle d'un soluté dans ces deux solvants non miscibles. C'est une mesure importante pour l'identification de la similarité médicamenteuse, selon la règle de Lipinski, les médicaments délivrés par voie orale doivent avoir des valeurs de log P supérieures ou égales à -2 et inférieures ou égales à 5) . Il est défini par la formule suivante:

$$\log P = \frac{[Octanol]}{[H_2O]}$$

[Octanol] et [H₂O] sont les concentrations du soluté dans l'Octanol et l'eau. Les composés qui ont les valeurs de log P > 0 sont dites lipophiles, et les composés qui ont les valeurs de log P < 0 sont dites hydrophiles. Si le Log P est positif et très élevé, cela exprime le fait que la molécule est plus soluble dans l'Octanol que dans l'eau, ce qui reflète son caractère lipophile, et inversement, si le Log P est négatif cela signifie que la molécule est hydrophile.

Un Log P nul signifie que la molécule est aussi soluble dans un solvant que dans l'autre.

La réfractivité moléculaire : notée (MR), en m³/mol, est le volume de la substance absorbée par mole de cette substance. Elle est définie par Lorentz-Lorenz par la formule suivante:

$$MR = \frac{n^2 - 1}{n^2 + 2} \frac{MW}{d} = \frac{n^2 - 1}{n^2 + 2} MV$$

Où : MW est le poids moléculaire ; d est la densité ; n est l'indice de réfraction ; MV est le volume molaire.

La réfractivité moléculaire est également proportionnelle à la polarisabilité α_e , par la relation suivante :

$$MR = \frac{4}{3} \pi N A \alpha_e \quad N A = 6.022 \cdot 10^{23}$$

Où : NA est le nombre d'Avogadro qui est, le nombre de molécules dans une mole de substance.

La polarisabilité : notée (α_e), en (m³), est l'aptitude à la déformation du nuage électronique de la molécule sous l'influence d'un champ électrique uniforme. C'est l'un des paramètres qui traduisent les propriétés moléculaires liées à l'hydrophobie et par conséquent aux activités biologiques. Elle est calculée à partir de la réfractivité molaire ou du volume molaire comme suit:

$$\alpha_e = 0,3964308 \times MR = 0,3964308 \times \frac{n^2 - 1}{n^2 + 2} MV$$

L'énergie totale : Pour une molécule isolée à l'état fondamental, l'énergie totale calculée, notée Et, , peut être utilisée comme descripteur moléculaire quantique. Cette énergie approximative a été calculée pour une conformation optimisée de la géométrie la plus stable dont la structure d'énergie est minimale.

Le moment dipolaire: noté μ , mesuré en debye (D), mesure la polarité nette moléculaire, et décrit la séparation de charge dans une molécule où la densité d'électrons est partagée inégalement entre les atomes. L'existence d'un moment dipolaire dans une molécule a son origine dans la différence d'électronégativité entre les atomes. La densité électronique est plus élevée au voisinage de l'atome le plus électronégatif. Ceci entraîne une dissymétrie dans la répartition des électrons de liaison. Ainsi, plus le moment dipolaire d'une molécule est élevé, plus la dissymétrie dans la molécule est importante.

L'énergie HOMO: notée EHOMO, est le niveau d'énergie le plus élevé dans la molécule qui contient des électrons, il est directement lié au potentiel d'ionisation.

Lorsqu'une molécule agit comme une base de Lewis (un doublet d'électrons donneur) dans la formation d'une liaison, les électrons sont alimentés à partir de cette orbite. Il mesure la nucléophilie d'une molécule et caractérise la susceptibilité de la molécule à l'attaque par des électrophiles.

CHAPITRE II : Représentation Des Méthodes De Calculs

L'énergie LUMO: notée ELUMO, est le niveau d'énergie le plus bas dans la molécule qui ne contient pas d'électrons, il est directement lié à l'affinité d'électron. Lorsqu'une molécule agit comme un acide de Lewis (un doublet d'électrons accepteur) dans la formation de liaisons, des doublets d'électrons entrants sont reçus dans cette orbite. Il mesure l'électrophilicité d'une molécule et caractérise la susceptibilité de la molécule à l'attaque par les nucléophiles.

Le Gap énergétique : ou l'écart HOMO-LUMO, noté E_{gap} , traduit l'énergie comprise entre l'orbitale moléculaire la plus haute occupée et la plus basse vacante, C'est un indice de stabilité important. Cette différence d'énergie sert de mesure de l'excitabilité d'une molécule. Ainsi, plus l'intervalle d'énergie sera faible et plus la molécule pourra interagir avec l'environnement.

Un grand écart HOMO-LUMO implique une grande stabilité pour la molécule dans le sens de sa faible réactivité dans les réactions chimiques, et de même, un faible écart implique une grande réactivité de la molécule. L'écart HOMO-LUMO a également été utilisé comme une approximation de la plus faible énergie d'excitation de la molécule.

Descripteurs de charge : noté q_i , Ce sont des descripteurs électroniques définis en termes de charges atomiques utilisées pour décrire les aspects électroniques aussi bien de la molécule entière que de régions particulières, comme les atomes, les liaisons, les fragments moléculaires et les orbitales. Les charges renseignent sur l'importance de la localisation de la densité électronique dans une molécule: des valeurs q_i négatives signifient la localisation d'un excès de charge électronique au centre i , alors que des valeurs positives indiquent que le centre i est électroniquement déficitaire.

Surface moléculaire : (SAG) C'est une enveloppe entourant les atomes localisés à la périphérie, qui explique la surface de contact Ligand-Récepteur c'est-à-dire l'affinité de liaison L-R [52-54].

L'HyperChem permet de calculer la surface grille (SAG), c'est une méthode de grille où une méthode plus rapide plus approximative; quel que soit la zone accessible au solvant ou la surface de van der Waals [55]. Dans cette théorie, chaque atome de la molécule est représenté par une sphère. La surface extérieure de toutes les sphères atomiques définit la surface de Van der Waals [56].

Energie d'hydratation : (HE) L'hydratation est La formation d'une solution implique l'interaction du soluté avec des molécules de solvant, différents liquides peuvent être utilisés comme solvants, mais l'eau est le solvant le plus couramment utilisé. Lorsque l'eau est utilisée comme solvant, le processus de dissolution est appelée hydratation .

Pour élaborer un modèle QSAR il faut une méthode d'analyse des données. Il existe plusieurs méthodes pour élaborer un modèle et analyser les données statistiques de ce

CHAPITRE II : Représentation Des Méthodes De Calculs

dernier, on peut obtenir ces méthodes dans des logiciels tels que, Excel, OriginMicrocal, Minitab, Statistica mais en notre cas SPSS 19 .

➤ Interprétation et validation d'un modèle QSAR :

les méthodes de validation appliquer dans cette thèse sont :

✓ Interne:

La technique la plus employée pour déterminer la stabilité du modèle prédictif est de LOO (leave-one-out), Quel que soit le but ultime du modèle QSAR, il doit être obligatoirement validé avant d'être interprété ou utilisé à fins prédictives. Il existe différentes solutions pour s'assurer de validité d'un modèle. La validation croisée *LOO (leave-one-out)* est un processus qui permet de tester la précision prédictive d'un modèle, cette méthode basée sur le calcul de quelques paramètres statistiques tels que :

$$PRESS = \sum (Y_{obs} - Y_{cal})^2$$

$$TSS = \sum (Y_{obs} - \bar{Y})^2$$

$$R_{adj}^2 = (1 - r^2) \left(\frac{n-1}{n-p-1} \right)$$

$$R_{cv}^2 = 1 - \frac{PRESS}{TSS}$$

$$S_{PRESS} = \sqrt{\frac{PRESS}{n}}$$

$$F = \frac{\frac{\sum (\hat{y}_i - \bar{y}_i)^2}{p}}{\frac{\sum (y_i - \hat{y}_i)^2}{n-p-1}}$$

$$SE = \frac{\sum (y_i - \hat{y}_i)^2}{n-p-1}$$

$$S = \sqrt{\frac{RSS}{n-p-1}}$$

$$PE = 0.6745 (1 - r^2) / \sqrt{n}$$

Avec: **n** : nombre de molécules de la série étudiée.

PRESS: Somme des carrés des écarts entre les valeurs observées et les valeurs prévues à partir du modèle.

SSY: Somme des carrés des distances des valeurs observées pour une variable par rapport à la moyenne de cette variable, la somme des carrés permet de mesurer la variation totale dans une variable.

R²adj : Estimation de la qualité d'ajustement du modèle à la population.

R²cv : Est la variation de la statistique du R² obtenue en ajoutant ou en enlevant une variable indépendante.

SPRESS : La capacité prédictive des modèles est évaluée par l'erreur de racine-carrée moyenne.

F : L'indice de Fisher *F* est également couramment employé afin de mesurer le niveau de signifiante statistique du modèle, c'est-à-dire la qualité du choix du jeu de paramètres .

SE : Il est l'Error Standard .

S:L'écart type (s) est un autre paramètre habituellement rapporté; il indique dans quelle mesure la fonction de régression prédite les données observées. et *p* est le nombre de variables indépendants.

PE:L'erreur de prédiction du coefficient de corrélation est utilisée pour déterminer le pouvoir prédictif du modèle QSAR.

pour Le choix du meilleur modèle on vérifié :

- ❖ F: quand la valeur de F est importante, l'hypothèse nulle peut être rejetée .
- ❖ SE: est une mesure de variation de la valeur d'un test statistique, d'un échantillon à l'autre
- ❖ R: est un coefficient de corrélation entre les valeurs observées et les prévisions des variables dépendantes, sa valeur variée entre 0 et 1 ($R^2 > 0.6$).
- ❖ modèle QSAR raisonnable (statistiquement significatif) = le rapport PRESS / SSY doit être inférieur à 0,4.
- ❖ Un bon pouvoir prédictif des modèles QSAR = $R > 6PE$.

✓ Externe :

La validation externe consiste à estimer l'activité d'une série de composés dite série de test qui ne sont pas inclus dans la série d'apprentissage. Le principe de cette technique est basée sur le calcul du coefficient de corrélation R^2 (test) et vérifier les critères de Tropsha (Tropshacriteria) [57].

$$R_{pred}^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

$$R^{o2} = 1 - \frac{\sum_{i=1}^{ntest} (\hat{y}_i - y_i^{r^o})^2}{\sum_{i=1}^{ntest} (\hat{y}_i - \bar{y})^2}, y_i^{r^o} = K \hat{y}_i$$

$$R'^{o2} = 1 - \frac{\sum_{i=1}^{ntest} (y_i - \hat{y}_i^{r^o})^2}{\sum_{i=1}^{ntest} (y_i - \bar{y})^2}, \hat{y}_i^{r^o} = K' y_i$$

$$K = \frac{\sum_{i=1}^{ntest} y_i \hat{y}_i}{\sum_{i=1}^{ntest} \hat{y}_i^2}$$

$$K' = \frac{\sum_{i=1}^{ntest} y_i \hat{y}_i}{\sum_{i=1}^{ntest} y_i^2}$$

les critères de Tropsha :

R^2_{pred} : coefficient de corrélation pour les molécules de la série de test. $R^2_{pred} > 0.7$.

R^{o2} : coefficient de corrélation entre les valeurs prédites et expérimentales pour la série de test.

R'^{o2} : coefficient de corrélation entre les valeurs expérimentales et prédites pour la série de test.

K : constante de la droite (à l' origine) de corrélation (valeurs prédites en fonction des valeurs expérimentales)

K' : constante de la droite (à l' origine) de corrélation (valeurs expérimentales en fonction des valeurs prédites).

pour Le choix du meilleur modèle on vérifié :

- ❖ $R^2 > 0.7$
- ❖ $R^2 - R^2_0 / R^2 < 0.1$ et $0.85 \leq k \leq 1.15$
- ❖ $R^2 - R'^2_0 / R^2 < 0.1$ et $0.85 \leq k' \leq 1.15$
- ❖ $|R^2 - R^2_0| \leq 0.3$

➤ Domaine d'applicabilité :

La méthode de « leverage » est la méthode la plus utilisée pour la détermination du domaine d'applicabilité d'un modèle QSAR [58-59] ou l'intervalle dans lequel nos composés pourront être prédits avec confiance. Cette technique est basée sur la variation des résiduels standardisés de la variable dépendance avec « leverage » (la distance entre les valeurs des descripteurs et leurs moyennes). Le principe est que le leverage d'un composé doit être inférieur à la valeur de seuil $h^* = 3p/n$ (ou p est le nombre de descripteurs plus 1 et n le nombre de points). Dans le cas contraire le composé sera considéré en dehors du domaine d'applicabilité du modèle obtenu .

II.6. Les Régressions linéaires

L'objectif de l'analyse statistique est justement de « démêler » ces descripteurs et d'identifier ceux qui sont corrélés à la variable cible, qui produisent du signal, de ceux qui ne le sont pas, qui produisent du bruit. L'analyse statistique permet également d'identifier les descripteurs qui sont corrélés entre eux pour ne garder que les principaux et réduire ainsi la redondance d'informations [60].

L'analyse statistique détermine et quantifie les corrélations entre les descripteurs et la variable cible. Elle indique également la contribution relative de chaque descripteur dans l'explication globale de l'activité. Le modèle statistique est une équation donnant la valeur de la variable cible en fonction de la somme des valeurs pondérées des descripteurs. Les principaux outils statistiques pour obtenir un modèle sont :

- la régression linéaire multi variée (*Multivariate Linear Regression* - MLR).
- la régression en composante principale (*Principale Component Regression* - PCR).
- la régression des moindres carrés partiels (*Partial Least Squares* - PLS) [61].
- les réseaux de neurones artificiels (*Artificial Neural Network* - ANN).

II.6.1. Régression linéaire multiple:

Une analyse de régression où la variable dépendante Y dépend linéairement de plusieurs variables indépendantes X_1, X_2, \dots, X_j est appelée régression linéaire multiple.

L'équation de régression linéaire multiple est de la forme [60].:

$$Y = f(X_1, X_2, \dots, X_j) \text{ où } f(X_1, X_2, \dots, X_j) \text{ est une fonction linéaire de } X_1, X_2, \dots, X_j$$

Le modèle mathématique peut se présenter sous la forme :

$$Y^i = a + \sum_{j=1}^{j=p} b_j x_j^i$$

où \mathbf{b} sont les coefficients de régression et \mathbf{a} la constante de régression. Considérons un modèle à N individus, p variables \mathbf{X}_j , où \mathbf{Y}_{obs} et \mathbf{Y}_{calc} sont respectivement les valeurs observées et calculées de la propriété et \bar{Y} la valeur moyenne de la variable à expliquer.

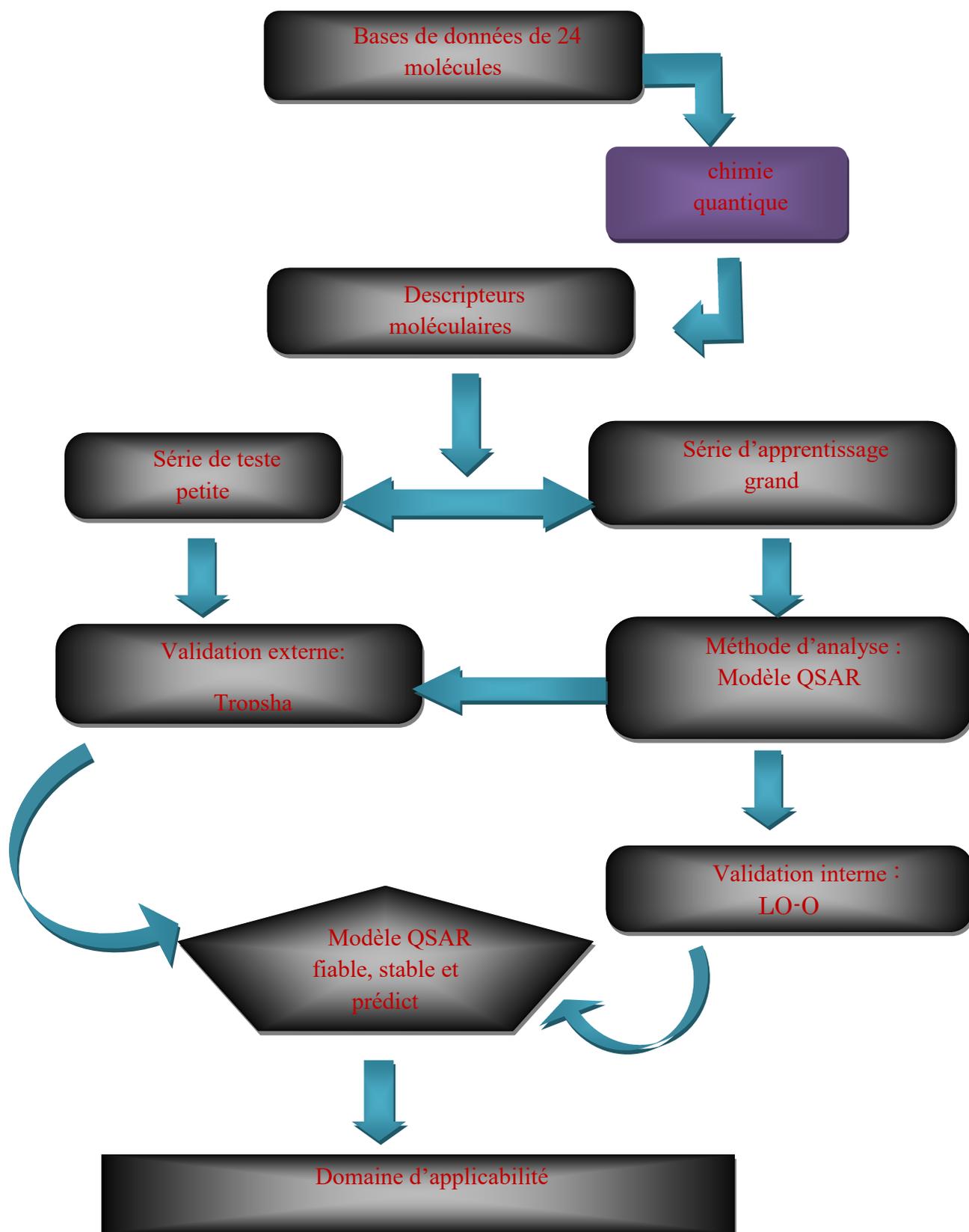


Figure.1:Méthodologie QSAR

II.7. Références Bibliographiques

1. G. Ohanessian, *Applications de la Chimie Quantique en Chimie-Cours*, École Polytechnique, Orsay, 2005.
2. R. Pollet, *Méthodes de la Chimie Quantique*, Cours de Master de Chimie–2ème année, Université Paris-Sud 11 et ENS Cachan, 2006.
3. Van Gunsteren, W. ; Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* 1990, 29, 992.
4. L. Pauling, E.B.J. Wilson, *Introduction to Quantum Mechanics with Applications to Chemistry*, McGraw-Hill Book Company Inc., New York, **1935**.
5. K.S. Pitzer, *Quantum Chemistry*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, **1953**.
6. Szabo, N.S. Ostlund, *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory*, Macmillan Publishing Co. Inc., New York, **1982**.
7. R.G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New-York, **1989**.
8. J.-L. Rivail, *Elements de chimie quantique à l'usage des chimistes*, CNRS Editions, Paris, **1999**.
9. D.C. Young, *Computational Chemistry: A Practical Guide for Applying Techniques to RealWorld Problems*, John Wiley & Sons Inc., New York, **2001**.
10. C.J. Cramer, *Essentials of Computational Chemistry - Theories and Models*, Wiley, Chichester, U.K., **2004**.
11. Guillaume Fayet. Développement de modèles QSPR pour la prédiction des propriétés d'explosibilité des composés nitroaromatiques. *Chimie. Chimie ParisTech*, 2010. Français. .
12. H. Dugas, *Principes de base en modélisation moléculaire, Aspects théorique et pratiques*, Libraire de L'Université de Montréal, 1996.
13. Hagler, A. T., Lifson, S., & Dauber, P. (1979). Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 2. A benchmark for the objective comparison of alternative force fields. *Journal of the American Chemical Society*, 101(18), 5122-5130.
14. J. Pople, D. Beveridge, P. Dobosh, Approximate self-consistent molecular-orbital theory. V. Intermediate neglect of differential overlap, *The Journal of Chemical Physics*, 47 (1967) 2026-2033.
15. J.A. Pople, G.A. Segal, Approximate self-consistent molecular orbital theory. III. CNDO results for AB₂ and AB₃ systems, *The Journal of Chemical Physics*, 44 (1966) 3289-3296.

CHAPITRE II : Représentation Des Méthodes De Calculs

16. L. Ordronneau, V. Aubert, R. Métivier, E. Ishow, J. Boixel, K. Nakatani, F. Ibersiene, D. Hammoutène, A. Boucekkine, H. Le Bozec, Tunable double photochromism of a family of bis-DTE bipyridine ligands and their dipolar Zn complexes, *Physical Chemistry Chemical Physics*, 14 (2012) 2599-2605.
17. J.A. Pople, D.P. Santry, G.A. Segal, Approximate self-consistent molecular orbital theory. I. Invariant procedures, *The Journal of Chemical Physics*, 43 (1965) S129-S135.
18. R.C. Bingham, M.J. Dewar, D.H. Lo, Ground states of molecules. XXV. MINDO/3. Improved version of the MINDO semiempirical SCF-MO method, *Journal of the American Chemical Society*, 97 (1975) 1285-1293.
19. N.C. Baird, M.J. Dewar, Ground States of σ -Bonded Molecules. IV. The MINDO Method and Its Application to Hydrocarbons, *The Journal of Chemical Physics*, 50 (1969) 1262-1274.
20. M.J. Dewar, W. Thiel, Ground states of molecules. 38. The MNDO method. Approximations and parameters, *Journal of the American Chemical Society*, 99 (1977) 4899-4907.
21. M.J. Dewar, E.F. Healy, A.J. Holder, Y.C. Yuan, Comments on a comparison of AM1 with the recently developed PM3 method, *Journal of computational chemistry*, 11 (1990) 541-542.
22. J.J. Stewart, Optimization of parameters for semiempirical methods II. Applications, *Journal of computational chemistry*, 10 (1989) 221-264.
23. N. Merabet, Thèse de Doctorat: Université 8 mai 1945 Guelma, 2017.
24. T. Shimanouchi, Force constants of small molecules, *Pure Appl. Chem*, 7, 131-146, 1963.
25. N.L. Allinger, Q. C. P. E, 3, 32, 1983.
26. N.L. Allinger, Y.H. Yuh, J-H. Lii, Molecular mechanics. The MM3 force field for hydrocarbons, *J. Am. Chem. Soc*, 111, 8551-8566, 1989.
27. N.L. Allinger, K. Chen, J-H. Lii, An improved force field (MM4) for saturated hydrocarbons, *J. Comp. Chem*, 17, 642-668, 1996.
28. G.B. Jones, B.J. Chapman, π stacking effects in asymmetric synthesis, *J. Syn. Org. Chem*, 5, 475-497, 1995.
29. Hocquet, M. Langgard, An evaluation of the MM+ force field, *J. Mol. Model*, 4, 94-112, 1998.
30. S.J. Weiner, P.A. Kollman, T. Nguyent, D.A. Case, An all atom force field for simulations of proteins and nucleic acids, *J. Comput. Chem*, 7, 230-252, 1986.
31. W.L. Jorgensen, J.T. Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc*, 110, 1657-1666, 1988.

CHAPITRE II : Représentation Des Méthodes De Calculs

32. J.C. Smith & M. Karplus, Empirical force field study of geometries and conformational transitions of some organic molecules, *J. Am. Chem. Soc.*, 114, 801-812, 1992.
33. W.F. Van Gusten, M. Karplus, Effect of constraints on the dynamics of macromolecules, *Macromolecules*, 15, 1528-1544, 1982.
34. B.R. Brooks, R.E. Bruccoleri, B.D. Ofalson, D.J. States, S. Swaminathan, M. Karplus, *Macromolecular Energy, Minimization, and Dynamics Calculations*, *J. Comp. Chem.*, 4, 187-217, 2004.
35. F. Tristram, V. Durier, G. Vergoton, The structures and vibrational frequencies of a series of aliphatic alcohols as obtained using the spectroscopic potential SPASIBA, *J. Mol. Struct.*, 378, 249-256, 1996.
36. T.A. Halgren, Merck molecular force field. IV. conformational energies and geometries for MMFF94, *J. Comp. Chem.*, 17, 490, 1996.
37. T.A. Halgren, R.B. Nachbar, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J. Comp. Chem.*, 17, 525-552, 1996.
38. M. Baaden, Outils pour étudier la structure et la dynamique des peptides et des protéines, Institut de Biologie Physico-chimique, PARIS, CNRS UPR9080, 2003.
39. Hansson, T.; Oostenbrink, C.; van Gunsteren, W. F. (2002) Molecular dynamics simulations. *Curr. Opin. Struct. Biol.*, 12, 190-196.
40. Huo, S. H.; Wang, J. M.; Cieplak, P.; Kollman, P. A.; Kuntz, I. D. (2002). Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: Insight into structure-based ligand design. *J. Med. Chem.*, 45, 1412- 1419.
41. Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, 9, 71.
42. T.W. Schultz, M.T. Cronin, J.D. Walker, A.O. Aptula, Quantitative structure–activity relationships (QSARs) in toxicology: a historical perspective, *Journal of Molecular structure: THEOCHEM*, 622 (2003) 1-22.
43. R. Hmamouchi, M. Bouachrine, T. Lakhlifi, Tentative Pratique du Relation Quantitatives Structure-Activité/Propriété (QSAR/QSPR), *Revue Interdisciplinaire*, 1.
44. Crum Brown, T.R. Fraser, *Trans. Roy.Soc. Edinburgh*, 1868, 25, 151-203.
45. L.P. Hammett, *J. Am. Chem. Soc.*, 1937, 59, 96-103.
46. C. Hansch, *Acc. Chem. Res.*, 1969, 2, 232-239.
47. C. Hansch, A. Leo, R.W. Taft, *Chem. Rev.*, 1991, 91, 165-195
48. R. Leardi, Genetic algorithms in chemometrics and chemistry: a review, *Journal of Chemometrics: A Journal of the Chemometrics Society*, 15 (2001) 559-569.

CHAPITRE II : Représentation Des Méthodes De Calculs

49. K. BELLIFA, Etude des relations quantitatives structure–toxicité des composés chimiques à l'aide des descripteurs moléculaires.«Modélisation QSAR.
50. Goulon-Sigwalt-Abram, 2008.
51. J. Wu, Y. Wang, Y. Shen, Chemomet. Intel. Lab. Syst. 131, 2014, 22-30.
52. Belouerghi Souhila Azri Souad ; Etude qualitative et quantitative des relations structureactivité d'une série de composés phénoliques, Mémoire de Master; Uniersité Mohamed khider de Biskra ; 2020 .
53. Hugo Kubinyi, Gerd Folkers, Yvonne C. Martin, *3D QSAR in Drug Design, Springer, 1998.*
54. Abhilash Thakur. Mamta Thakur, Total QSAR study on Phenyl Acridine Derivatives in Reference to DNA Binding Affinity, 3rd International conference on Medical, Biological and Pharmaceutical Sciences, **2013.**
55. HyperChem help
56. Ovidiu Ivanciuc, 3D QSAR MODELS, Department of Organic-chemistry, Faculty of Chemical Technology University "Politehnica" of Bucharest, Romania.
57. S.K. Balasubramanian, J.A. Karrh, H. Patwardhan, Audience response to product placements: An integrative framework and future research agenda, Journal of advertising, 35 (2006) 115-141.
58. Tetko, I. Sushko, A.K. Pandey, H. Zhu, A. Tropsha, E. Papa, T. Oberg, R. Todeschini, D. Fourches, A. Varnek, Critical assessment of QSAR models of environmental toxicity against Tetrahymena pyriformis: focusing on applicability domain and overfitting by variable selection, Journal of chemical information and modeling, 48 (2008) 1733-1746.
59. T.I. Netzeva, A.P. Worth, T. Aldenberg, R. Benigni, M.T. Cronin, P. Gramatica, J.S. Jaworska, S. Kahn, G. Klopman, C.A. Marchant, Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships, ATLA, 33 (2005).Fortuné, Thèse de doctorat, Techniques de Modélisation Moléculaire Appliquées à l'Etude et à l'Optimisation de Molécules Immunogènes et de Modulateurs de la Chimiorésistance, Université Joseph Fourier – Grenoble I, **2006.**
60. P. Khadikar, M. Jaiswal, M. Gupta, D. Mandloi and R. S. Sisodia, *Bioorg. Med. Chem. Lett.*, **15**, 1249 **2005**

CHAPITRE 3 :
Etude QSAR d'une série de dérivés
de 7-AZAINDOLE

III.1. Introduction

Le modèle QSAR (quantitative structure-activité Relationship) étudie la corrélation basée sur un calcul statistique, de paramètres physico-chimiques avec l'activité biologique [1]

La comparaison de l'activité biologique de certaines molécules et de leur structure a permis d'établir dans de nombreux cas des corrélations entre les paramètres structuraux et les propriétés d'une molécule. L'association de la variation de l'activité aux paramètres permet d'obtenir un système d'équations qui donne, pour une série chimique donnée et pour une activité définie, une équation de corrélation.

Dans ce travail, nous nous intéressons à l'étude des propriétés physico-chimiques et électroniques de 7-azaindole et ses dérivés qui permettent de prévoir ces paramètres qui sont influents sur l'activité biologique, afin de prédire l'activité biologique de nouvelles molécules. Dans ce chapitre, le modèle final qui lie l'efficacité des dérivés du 7-azaindole comme C-met kinase inhibiteurs et la capacité prédictive de ce modèle a été discutée.

III.2. MATÉRIELS ET MÉTHODES

Pour réaliser cette étude nous disposons des logiciels suivants :

Tout d'abord, les Trent deux molécules du 7-azaindole ont été pré-optimisées au moyen de la mécanique moléculaire, champ de force (MM+) puis les structures obtenues ont été optimisées avec la méthode semi-empirique PM3, tout cela en utilisant le logiciel HyperChem 8.0 [2].

Ensuite, ces dérivées ont été ré-optimisées en utilisant le logiciel Gaussian 09 [3], au niveau de la théorie de la densité fonctionnelle DFT en utilisant le Lee-Yang-Parr à trois paramètres de Becke (B3LYP), avec de base 6 - 311G (d, p). A la fin des calculs de logiciel, des propriétés électroniques seront obtenues sont: le Gap énergétique (ΔE), L'énergie HOMO, L'énergie LUMO, l'énergie totale (ET), le moment dipolaire (MD) et les charges de quelques molécules (Q).

Le module « propriétés QSAR » de l'HyperChem 8.03 a été utilisé pour calculer les paramètres suivants: La polarisabilité (Pol), la réfractivité molaire (MR), le coefficient de partage octanol/eau ($\log P$), le volume molaire (MV), la surface moléculaire (SAG), le poids moléculaire (MW) et l'énergie d'hydratation (HE).

La méthode de régression multilinéaire (MLR) a été utilisée pour développer un modèle QSAR en utilisant le logiciel SPSS 19 [4].

III.3. Elaboration Du Modèle QSAR

III.3.1. Une Base De Données :

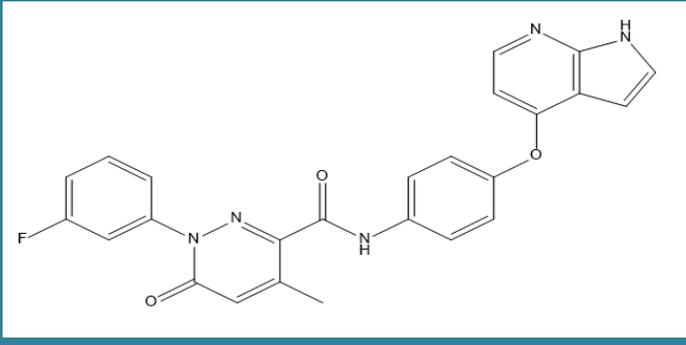
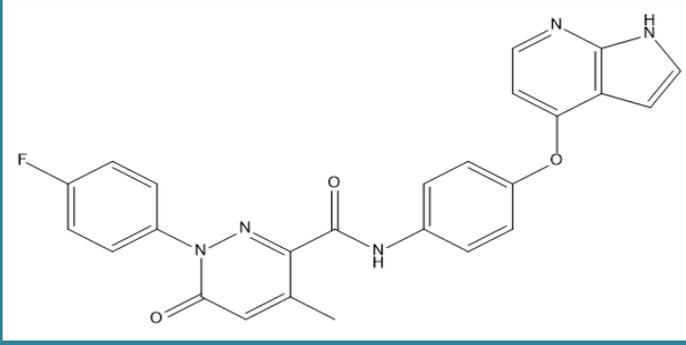
Plusieurs dérivés de 7-azaindole présentent des propriétés pharmacologiques importantes. Ils sont utilisés dans le traitement et la prévention d'un Plusieurs maladies.

Le choix de la base de données expérimentale initiale est une étape critique pour le développement des modèles QSAR. Elle doit être composée de données expérimentales fiables.

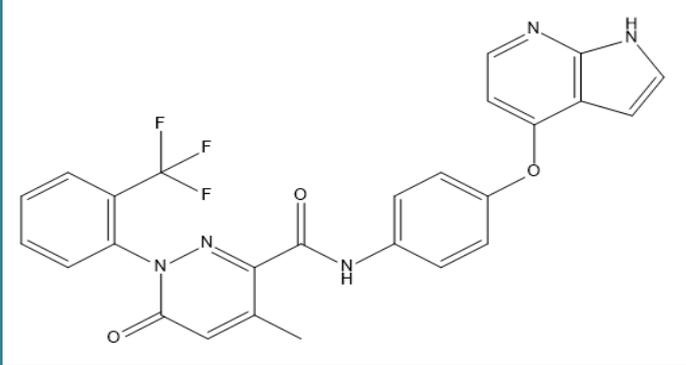
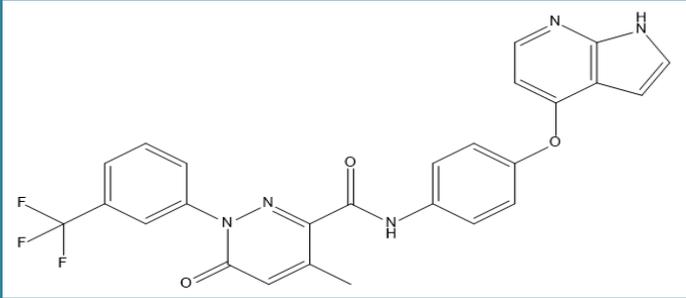
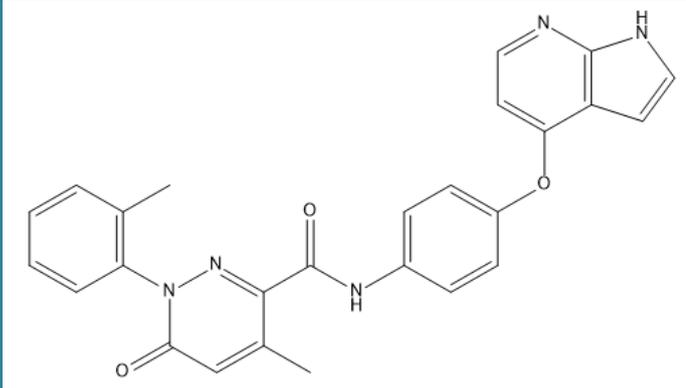
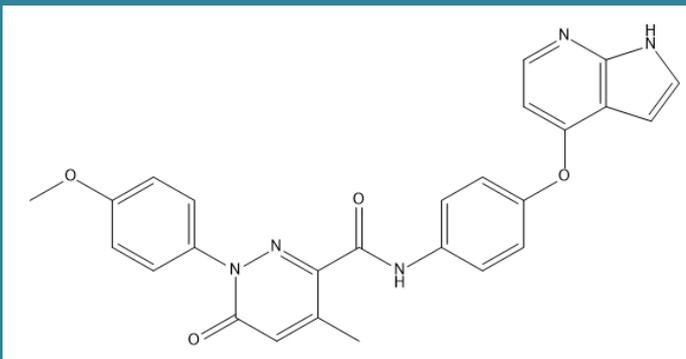
Pour cette étude nous avons utilisé L'activité P-IC50 des dérivés de 7-azaindole .

Les structures et l'activité biologique antiprolifératif expérimentale des dérivés de 7-azaindole sont rapportées dans le tableau 1:

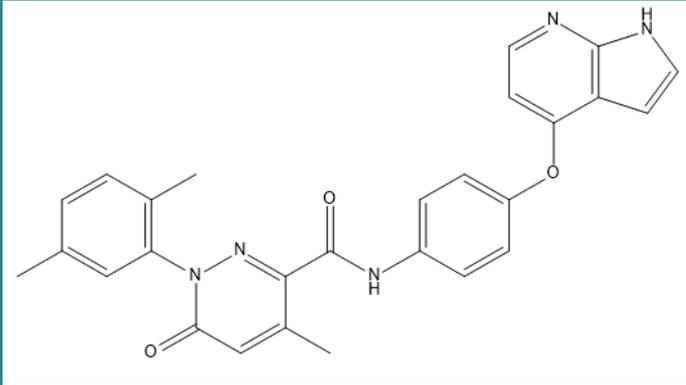
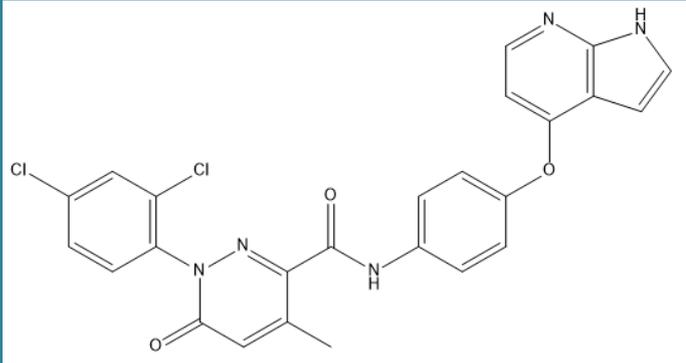
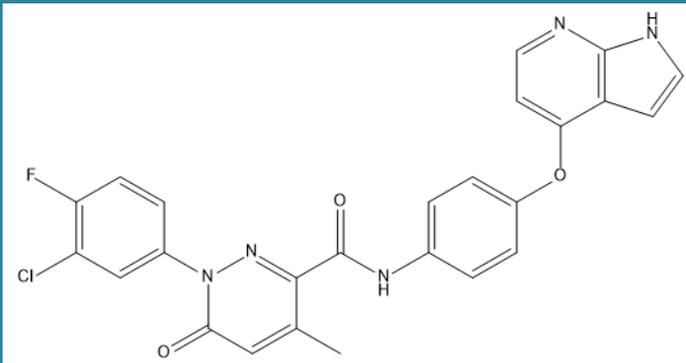
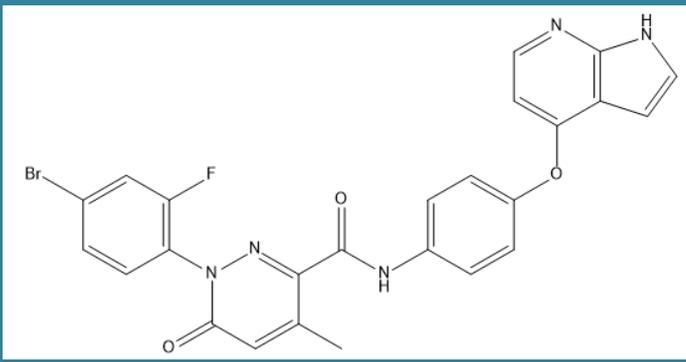
Tableaux.1 :les structures chimiques et L'activité biologique des dérivés de 7-azaindole .

<u>N°</u>	<u>X</u>	<u>R1</u>	<u>R2</u>	<u>Molécule</u>	<u>Ic50</u>	<u>P</u> <u>(Ic50)</u>
01	H	H	3-F		11.63±1.05	-1.065
02	H	H	4-F		8.42±0.58	-0.925

Chapitre III: Etude QSAR D'une Série De Derivés De 7- Azaindole

03	H	H	2-CF ₃		39.52±2.18	-1.598
04	H	H	3-CF ₃		35.29±1.72	-1.548
05	H	H	2-CH ₃		19.14±1.62	-1.281
06	H	H	4-OCH ₃		46.91±2.29	-1.671

Chapitre III: Etude QSAR D'une Série De Derivés De 7- Azaindole

07	H	H	2,5-(CH ₃) ₂		3.351±1.48	-1.525
08	H	H	2,4-(Cl) ₂		56.25±2.06	-1.750
09	H	H	3-Cl,4-F		42.15±1.39	-1.624
10	H	H	2-F,4-Br		41.23±1.86	-1.615

Chapitre III: Etude QSAR D'une Série De Derivés De 7- Azaindole

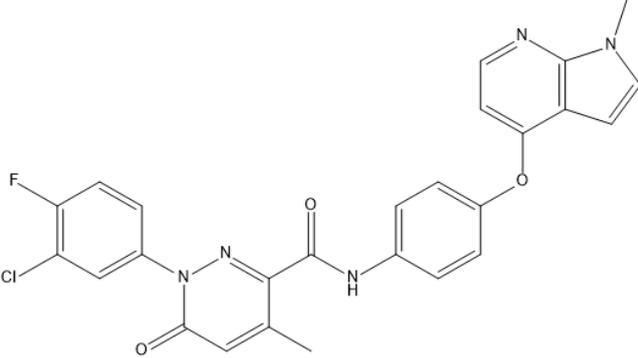
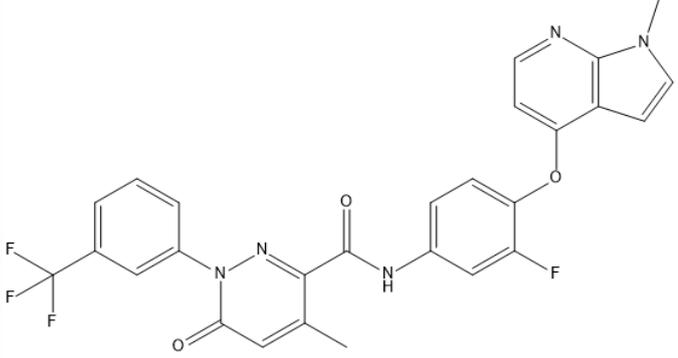
11	H	F	3-F		5.61 ± 0.63	-0.749
12	H	F	4-F		1.06 ± 0.12	-0.025
13	H	F	2-CF ₃		11.23 ± 1.58	-1.050
14	H	F	4-OCH ₃		18.25 ± 1.35	-1.261

Chapitre III: Etude QSAR D'une Série De Derivés De 7- Azaindole

15	H	F	2,4-(Cl) ₂		22.39±1.36	-1.350
16	H	F	2-F,4-Br		18.24±1.03	-1.261
17	CH ₃	H	H		19.54±2.06	-1.291
18	CH ₃	H	4-F		9.21±2.36	-0.964

Chapitre III: Etude QSAR D'une Série De Derivés De 7- Azaindole

19	CH_3 H $2-CF_3$		46.27 ± 1.75	-1.665
20	CH_3 H $3-CF_3$		32.15 ± 1.80	-1.507
21	CH_3 H $2,5-CH_3$		41.25 ± 1.43	-1.615
22	CH_3 H $2,4-(Cl)_2$		49.16 ± 2.39	-1.691

23	CH ₃	H	3-Cl,4-F		45.31±1.50	-1.656
24	CH ₃	F	3-CF ₃		29.35±1.71	-1.468

III.3.2. Calcul Des Descripteurs

Le choix des descripteurs dépend des outils dont on dispose, de la nature des composés décrits et de la propriété ciblée. L'information codée d'un descripteur moléculaire dépend du type de représentation moléculaire employée et de l'algorithme défini pour son calcul. Dans cette partie nous avons étudié sept propriétés physico-chimiques d'une série de dérivés de 7-azaindole à l'aide du logiciel HyperChem 8.0 [2]. sont représentés dans le tableau suivant :

Le volume moléculaire; Le poids moléculaire(masse); Le coefficient de partage Octanol/Eau (logP); La réfractivité moléculaire (refr); La polarisabilité (pol); Surface moléculaire (SAG); Energie d'hydratation (HE).

Chapitre III: Etude QSAR D'une Série De Dérivés De 7- Azaindole

Tableaux 2 : les descripteurs obtenu par Hyper-chem 8.0

<i>MOLICULE</i>	<i>Pol</i> (Å ³)	<i>Refr</i> (Å ³)	<i>Log P</i>	<i>Masse</i> (amu)	<i>Volume</i> (Å ³)	<i>SAG</i> (Å ²)	<i>HE</i> Kcal/mol
01	47.498	135.633	-1.354	455.448	1236.410	738.782	-13.367
02	47.498	135.633	-1.354	455.448	1236.870	741.745	-13.393
03	49.151	140.720	-0.184	505.456	1280.290	742.015	-12.768
04	49.151	140.720	-0.184	505.456	1305.039	770.053	-13.190
05	49.424	139.787	-0.600	451.484	1269.168	744.160	-12.865
06	50.061	141.880	-1.746	467.484	1304.838	771.614	-15.465
07	51.259	144.069	-0.446	465.511	1320.759	770.930	-11.760
08	51.445	144.938	-1.198	506.348	1303.754	768.152	-12.938
09	49.552	125.177	4.076	489.893	1277.867	765.812	-13.059
10	50.124	143.167	-1.303	534.344	1294.099	767.907	-12.964
11	47.407	135.761	-1.955	473.438	1245.027	739.843	-13.041
12	47.407	135.761	-1.955	473.438	1245.476	739.288	-13.042
13	49.060	140.848	-0.785	523.446	1285.839	745.602	-12.278
14	49.970	142.008	-2.347	485.474	1314.709	781.716	-14.860
15	51.354	145.066	-1.799	524.338	1312.014	773.941	-12.468
16	50.033	143.295	-1.904	552.334	1303.849	773.258	-12.636
17	49.424	140.402	-0.506	451.484	1282.418	757.781	-10.065
18	49.333	140.530	-1.108	469.475	1291.589	769.295	-9.644
19	51.112	135.233	4.437	519.483	1292.606	713.670	-8.993
20	50.986	145.617	0.063	519.483	1324.600	755.789	-9.571
21	53.094	148.966	-0.200	479.538	1374.821	803.787	-7.881
22	53.280	149.835	-0.952	520.374	1359.453	793.814	-9.190
23	51.261	145.246	-1.330	503.920	1333.680	791.034	-9.336
24	50.895	145.744	-0.539	537.473	1368.634	805.695	-8.977

Et nous avons étudié les propriétés quantique de cette série de dérivés de 7-azaindole à l'aide du logiciel Gaussian 09 [3]., sont représentés dans le tableau suivant :

Le moment dipolaire; L'énergie HOMO; L'énergie LUMO; Le gap énergétique ; L'énergie totale(E).... tableaux 3

Et les charges des noyaux N3,N27,C9,C20,C21,C22tableaux 4

Tableaux 3 : les descripteurs obtenu par gaution 09. 01

N°	Md	Humo	Lumo	Gap	E
01	2.2852	-0.21800	-0.10664	0.11136	-1.224271
02	1.5132	-0.21913	-0.10538	0.11375	-1.027343
03	3.3235	-0.21583	-0.10221	0.11362	-7.334231
04	4.2749	-0.22037	-0.11108	0.10929	-8.749020
05	2.2948	-0.21746	-0.09706	0.1204	-1.241239
06	3.1028	-0.21684	-0.09706	0.11978	-5.773664
07	2.4479	-0.21741	-0.09574	0.12167	-1.815746
08	0.6467	-0.21832	-0.10525	0.11307	-1.538861
09	2.8360	-0.21911	-0.11170	0.10741	-1.128243
10	0.6931	-0.21758	-0.10490	0.11268	-2.430664
11	2.4338	-0.22051	-0.10995	0.11056	-1.459989
12	1.4099	-0.20754	-0.09545	0.11209	-5.412013
13	5.2262	-0.21630	-0.10581	0.11049	-6.567986
14	4.4652	-0.21839	-0.10078	0.11761	-1.154973
15	2.0213	-0.21890	-0.10896	0.10994	-3.797723
16	1.5463	-0.21875	-0.10816	0.11059	-1.261640
17	2.2145	-0.21296	-0.09940	0.11356	-1.463176
18	1.5162	-0.21444	-0.10501	0.10943	-6.841814
19	2.1946	-0.21425	-0.08271	0.13154	-3.168827
20	7.8031	-0.22287	-0.09984	0.12303	-3.384541
21	2.4335	-0.21277	-0.09539	0.11738	-3.847107
22	1.0542	-0.21365	-0.10488	0.10877	-8.918095
23	2.5225	-0.21446	-0.11132	0.10314	-1.349400
24	4.1850	-0.21681	-0.11395	0.10286	-5.935826

Tableaux 4 : les descripteurs obtenu par gaution 09. 02

N°	Q-N3	Q-N27	Q-C9	Q-C20	Q-C21	Q-C22
01	-0.48533	-0.54076	-0.21303	-0.27007	0.40836	-0.25666
02	-0.48518	-0.54066	-0.21277	-0.18621	-0.25688	0.40841
03	-0.48667	-0.54118	-0.21484	-0.16669	-0.15760	-0.18410
04	-0.48442	-0.54050	-0.21612	-0.16004	-0.15367	-0.16000
05	-0.48592	-0.54090	-0.21905	0.00990	-0.19626	-0.18154
06	-0.48599	-0.54099	-0.21783	-0.18846	-0.22913	0.31889
07	-0.48603	-0.54095	-0.21987	0.00167	-0.18719	-0.18950
08	-0.48552	-0.54079	-0.21368	0.02270	-0.25687	0.02048
09	-0.48472	-0.54057	-0.21144	-0.22098	-0.05762	0.37660
10	-0.48566	-0.54089	-0.21615	0.41950	-0.28193	-0.07402
11	-0.48217	-0.54032	0.38351	-0.25255	0.41145	-0.25455
12	-0.46856	-0.57070	0.33141	-0.22928	-0.30668	0.36552
13	-0.48380	-0.54093	0.38374	-0.16710	-0.15742	-0.18361
14	-0.48307	-0.54067	0.38498	-0.17104	-0.22487	0.31961
15	-0.48271	-0.54056	0.38541	0.02216	-0.25687	0.02036
16	-0.48281	-0.54056	0.38310	0.41939	-0.28119	-0.07394
17	-0.49766	-0.38338	-0.21694	-0.20333	-0.19378	-0.19119
18	-0.49751	-0.38321	-0.21331	-0.18614	-0.25690	0.40837
19	-0.49285	-0.38325	-0.22107	-0.15418	-0.15558	-0.18229
20	-0.49013	-0.38195	-0.23485	-0.16692	-0.15606	-0.16448
21	-0.49837	-0.38348	-0.21388	0.00167	-0.18721	-0.18952
22	-0.49783	-0.38332	-0.21384	0.02271	-0.25689	0.02047
23	-0.49699	-0.38311	-0.21129	-0.22093	-0.05759	0.37654
24	-0.49394	-0.38279	0.38274	0.38274	-0.15307	-0.15923

III.4. Interprétation D'un Modèle QSAR

Dans cette étude nous avons essayé de développer le meilleur modèle QSAR pour expliquer les corrélations entre les paramètres physico-chimiques et électroniques et des valeurs PIC50 de l'activité biologique de dérivés 7-azaindole

Une série de 24 dérivés de 7-azaindole ont été utilisés pour la production du modèle de régression multilinéaire. Différents descripteurs physico-chimiques ont été utilisés comme variables indépendantes et ont été corrélés avec l'activité biologique en utilisant le logiciel SPSS 19 [4].

06 modèles de régression linéaire multiple (MLR) avec Ils sont définis dans le tableau ci-dessous :

Tableaux 5 : Modèles sélectionnés et paramètres statistiques des corrélations entre les propriétés moléculaires et l'activité biologique antiprolifératif

Modèle	Equation	R	R ²	F	SE	N
1	P-IC50= 25.861-0.044 Log P-0.02 Vol + 0.021 SAG + 0.169 HE + 0.47 MD +27.09 LUMO - 0.34 E + 24.804 Q-N3 +0.549 Q-C9	0.981	0.981	80.984	0.068584	24
2	P-IC50= 30.054 -0.069 Réfr - 0.247 Log P +0.183 HE + 7.452 LUMO + 38.774 Q-N3	0.975	0.950	68.485	0.09902	24
3	P-IC50= 27.769 -0.097 Réfr - 0.314 Log p +0.004 SAG + 0.213 HE +16.893 Gap - 0.027 E + 37.568 Q-N3	0.985	0.970	74.401	0.081136	24
4	P-IC50= 18.152 - 0.135 Pol -0.004 Masse -0.108 MD + 47.943 HUMO + 0.893 Q-C9	0.967	0.936	52.485	0.112653	24

5	P-IC50= 32.135 - 0.075 Pol -0.041 Réfr - 0.174 Log p +0.17 HE + 18.418 HUMO + 37.094 Q-N3 - 0.136 Q-C20	0.988	0.976	94.922	0.073965	24
6	P-IC50= 13.191- 0.058 Réfr -0.242 Log P +0.086 HE - 0.067 MD + 15.894 LUMO -0.049 E +0.575 Q-C9	0.969	0.939	35.245	0116861	24

Parmi 06 équations de QSAR le meilleur modèle QSAR a été sélectionné sur la base de divers paramètres statistiques tels que :

- Nous avons une forte relation si $R \in [0,8, 1]$ ou $R \in [-1, -0,8]$;
- relation modérée si $R \in [0,5, 0,8]$ ou $R \in [-0,8, -0,5]$;
- faible relation si $R \in [-0,5, 0,5]$.
- Le coefficient de corrélation au carré ($R^2 > 0,6$), qui est une mesure relative de la qualité de l'ajustement, l'erreur standard de l'estimation ($SE < 0,3$) représentant mesure absolue de la qualité de l'ajustement, La valeur de Fischer (F), F est le rapport de Fisher, reflète le rapport de la variance expliquée par le modèle et la variance due à l'erreur dans la régression. Les valeurs élevées du test F indique que le modèle est statistiquement significatif. [5]

Après une analyse de régression linéaire multiple sur le logiciel SPSS 19 [4]. les meilleures corrélations statistiquement significatives générées ainsi que les paramètres statistiques pertinentes sont donnés ci-dessous (**modèle 1**) :

$$P-IC50 = 25.861 - 0.044 \text{ Log P} - 0.02 \text{ Vol} + 0.021 \text{ SAG} + 0.169 \text{ HE} + 0.47 \text{ MD} \\ + 27.09 \text{ LUMO} - 0.34 \text{ E} + 24.804 \text{ O-N3} + 0.549 \text{ O-C9}$$

$$R^2=0.981, R^2_{adj}=0.969, F=80.984, SE=0.068584, N=24, P=9$$

- ✓ Les valeurs $R=0.991$, $R^2 = 0,981$ nous permis d'indiquer fermement la corrélation entre les différents descripteurs et l'activité biologique étudiée (P-IC50).

Chapitre III: Etude QSAR D'une Série De Dérivés De 7- Azaindole

- ✓ La valeur de F calculée est importante 80.984 et indique donc que la relation de régression était statistiquement significative [5].
- ✓ SE: est une mesure de variation de la valeur d'un test statistique, d'un échantillon à l'autre il doit être très petits pour considérer le modèle significative.

Dans les équations de modèle 1, le coefficient négatif de volume explique que toute augmentation de volume moléculaire des dérivés de 7-azaindole entraîne une diminution de l'activité biologique antiprolifératif.

Le coefficient positif de l'énergie d'hydratation et le coefficient négative de Log P indique que les dérivés hydrophiles donnent une bonne activité biologique.

Dans le modèle on remarque toute augmentation de la surface Moléculaire , Energie d'hydratation (HE), le Moment dipolaire , L'énergie LUMO et les charges des noyaux N3,C9 provoque à une augmentation de l'activité biologique antiprolifératif .

Dans le modèle on remarque toute augmentation de Le coefficient de partage Octanol/Eau (logP) et L'énergie totale(E) provoque à une diminution de l'activité biologique antiprolifératif .

III.5.Validation De Modèle

Afin de tester la validité de pouvoir prédictif de modèle QSAR sélectionné on utilise la technique de validation croisée LOO (leave-one-out) , présenter les paramètres statistiques calculé pour la validation modèle :

Tableau 6 :Paramètres de validation croisée.

PRESS	SSY	PRESS/SSY	SPRESS	R ² cv	R ² adj	6PE
0.06778712	3.59686996	0.018846141	0.05314568	0.98115386	0.969	0.015695717

- PRESS (carrés de la somme résiduelle prédite)Sa valeur étant inférieure à SSY souligne que le modèle prédit mieux que le hasard et peut être considéré comme statistiquement important, SSY (somme des carrés de la valeur de réponse)En outre, pour un modèle raisonnable de QSAR, le rapport PRESS/SSY doit être inférieur à 0,4[6]. Les résultats présentés dans le tableau indiquent que ce rapport est de 0.018846.

Chapitre III: Etude QSAR D'une Série De Derivés De 7- Azaindole

- Le modèle QSAR développé présente une faible valeur de Spres (<0.200) ce qui indique que le modèle présente une petite valeur résiduelle entre l'activité biologique observée et prédite. Les résultats présentés dans le tableau indiquent que ce rapport est de 0.053.
- les valeurs élevées du R^2CV et R^2adj (dépassent 0,5) sont des critères essentiels pour une meilleure qualification des modèles QSAR [7]. Les résultats obtenus de R^2CV et R^2adj pour ce modèle QSAR est 0.981 et 0.961 .
- PE : L'erreur de prédiction du coefficient de corrélation est utilisée pour déterminer le pouvoir prédictif du modèle QSAR, Un bon pouvoir prédictif des modèles QSAR = $R > 6PE$. Les résultats présentés dans le tableau indiquent que ce rapport est de 0.015695717 inférieur a R qui est égale 0.991 .

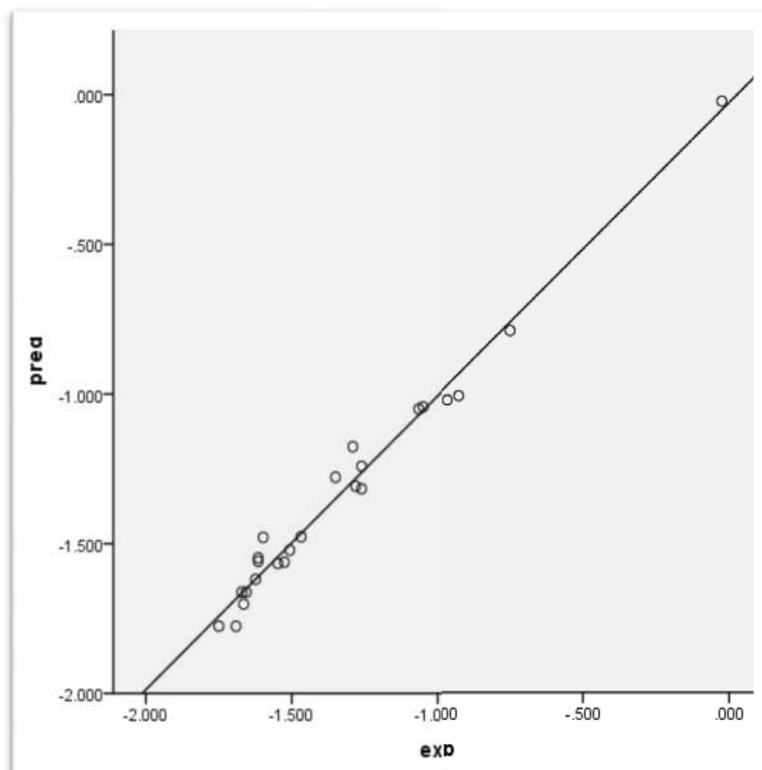
L'activité expérimentale, prédite et résiduelle des dérivés de 7-azaindole sont représentés dans le tableau 7 ont été déduit par logiciel SPSS [4] . Le résidu est dû d'une part à l'incertitude de la mesure expérimentale et d'autre part à l'imperfection du modèle, donc si l'activité prédite est plus proche de l'activité expérimentale, le modèle peut être applicable [8].

Tableau 7 : Valeurs Expérimentales , Prédites Et Résiduelles De P-IC50 Des Dérivées De 7-Azaindole

N°	P-IC50	Prévision	Résidu	N°	P-IC50	Prévision	Résidu
1	-1.065	-1.04961	-.015389	13	-1.050	-1.04273	-.007271
2	-.925	-1.00699	.081987	14	-1.261	-1.31640	.055402
3	-1.598	-1.47866	-.119340	15	-1.350	-1.27764	-.072360
4	-1.548	-1.56601	.018005	16	-1.261	-1.24114	-.019856
5	-1.281	-1.30790	.026901	17	-1.291	-1.17592	-.115077
6	-1.671	-1.66180	-.009202	18	-.964	-1.02136	.057362
7	-1.525	-1.56159	.036593	19	-1.665	-1.70158	.036581
8	-1.750	-1.77539	.025386	20	-1.507	-1.52153	.014535
9	-1.624	-1.61902	-.004983	21	-1.615	-1.55884	-.056163
10	-1.615	-1.54739	-.067612	22	-1.691	-1.77557	.084572
11	-.749	-.78904	.040035	23	-1.656	-1.66183	.005835
12	-.025	-.02042	-.004579	24	-1.468	-1.47664	.008638

La figure ci-dessous (figure. 1) présente la corrélation entre les valeurs prédites et expérimentales pour l'activité biologique étudiée des dérivés 7-azaindole. Du point de vue des performances statistiques, ce modèle présente de coefficient de corrélation R^2 significatif de valeur. La valeur de R^2 est égale à 0.991, pour ce modèle. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité étudiée.

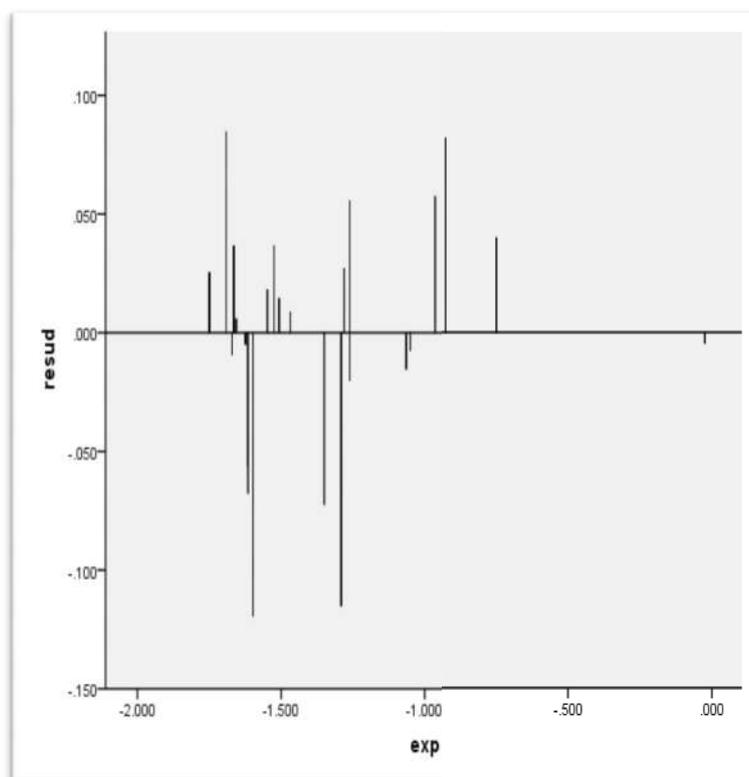
Figure. 1 : Corrélation Entre L'activité Biologique Expérimentale Et Prédite



Pour rechercher la présence d'une erreur systématique dans l'élaboration des modèles QSAR, les valeurs résiduelles de l'activité biologique ont été tracées en fonction des valeurs expérimentales, comme le montre la figure 2.

La répartition des résidus dans les deux côtés de l'axe des abscisses indique l'absence de l'erreur systématique [9]. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité biologique antiprolifératif pour l'ensemble des dérivés de 7-azaindoles utilisés dans le développement de modèle QSAR.

Figure 2 Corrélation Entre L'activité Biologique Expérimentale Et Résiduell



III.6. Références Bibliographiques

1. V. N. Viswanadhan, M.R.Reddy, R.J.Bacquet et M.D.Erion, journal of computational Chemistry, 14, 1993, 1019.
2. HyperChem v8. Molecular Modelling System, Hypercube Inc., 1115 NW 4th Street, Gainesville, FL 32601, USA, 2009.
3. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09, Gaussian, Inc., Wallingford CT, 2010.
4. SPSS 19 for Windows, SPSS Incorporation, <http://www.sp>
5. K.J. Sanmati, J. Rahul, S. Lokesh, K. Y. Arvind, *Journal of Chemical and Pharmaceutical Research*, 4, 3215- 3223, 201.
6. Reenu, & Vikas. (2015). Exploring the role of quantum chemical descriptors in modeling acute toxicity of diverse chemicals to *Daphnia magna*. *Journal of Molecular Graphics and Modelling*, 61, 89-101.
7. S.O. Podunavac-Kuzmanović, D.D. Cvetković, D.J. Barna, QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against *Pseudomonas aeruginosa*, *Int. J. Mol. Sci*, 10, 1670-1682, 2009.
8. H. Gonzalez-Diaz, E. Olazabal, L. Santana, E. Uriarte, Y. Gonzalez-Diaz, N.Castanedo, QSAR study of anticoccidial activity for diverse chemical compounds: Prediction and experimental assay of trans-2-(2-nitrovinyl) furan, *Bioorg. Med. Chem*, 15, 962-968, 2007.
9. M. Jalali-Heravi, A. Kyani, Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds, *J. Chem. Inf. Comput. Sci*, 44, 1328-1335, 2004.

Conclusion générale

Conclusion Générale

Dans ce travail, nous avons appliqué les méthodes de la chimie informatique sur une série de 24 dérivés de 7-azaindole. Différentes méthodes de modélisation moléculaire ont été utilisées dans notre travail.

La mécanique et la dynamique moléculaire ont été utilisées dans l'étude des propriétés physicochimiques des dérivés de 7-azaindole. Les méthodes de la mécanique quantique; PM3 et DFT (B3LYP/6-311 G(d,p)) ont été utilisées pour déterminer les paramètres structurales, électroniques et énergétiques associés aux molécules étudiées.

Des modèles QSAR ont été établis en utilisant la méthode de régression multilinéaire (MLR). Les modèles obtenus peuvent être utilisés pour prédire l'activité spécifique de 7-azaindole et la production de l'activité des autres inhibiteurs de C-MET tyrosine kinase .

Le pouvoir prédictif des modèles obtenus a été confirmé par la méthode de validation croisée LOO. Une forte corrélation a été observée entre les valeurs expérimentales et prédites des activités biologiques P-IC50, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Dans le cadre de travaux de recherche, Le présent travail comporte une étude fondamentale et originale sur les 7azaindole pour prédire l'activité biologique des composés étudiés .

Une étude quantitative de relation structure-Activité (QSAR) a été effectuée sur vingt quatre molécules de 7-azaindole. la régression linéaire multiple(MLR) a été utilisé pour quantifier la relation entre les descripteurs moléculaire et l'activité antiprolifératif de 7-azaindole .le pouvoir prédictif du modele obtenu a été confirmé par $R = 0,991$, $R^2 = 0,981$ et $F=80.984$. La méthode de validation croisée LOO nous montre une forte corrélation a été observée entre les valeurs expérimentales et prédites des activités biologiques, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Mot-Clés: antiprolifératif, 7-AZAINDOLE, C-Met Tyrosine Kinase, QSAR,

MLR. LOO.

Résumé :

Dans le cadre de travaux de recherche, Le présent travail comporte une étude fondamentale et originale sur les 7azaïndole pour prédire l'activité biologique des composés étudiés .

Une étude quantitative de relation structure-Activité (QSAR) a été effectuée sur vingt quatre molécules de 7-azaïndole. la régression linéaire multiple (MLR) a été utilisée pour quantifier la relation entre les descripteurs moléculaire et l'activité antiprolifératif de 7-azaïndole .le pouvoir prédictif du modèle obtenu a été confirmé par $R = 0,991$, $R^2 = 0,981$ et $F=80.984$. La méthode de validation croisée LOO nous montre une forte corrélation a été observée entre les valeurs expérimentales et prédites des activités biologiques, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Mot-Clés: antiprolifératif, 7-AZAINDOLE, C-Met Tyrosine Kinase, QSAR, MLR. LOO.

Abstract:

As part of research works, This work involves a fundamental and original research on the 7- azaïndoles to predict biological activity of the studied compounds.

Quantitative structure-activity relationship « QSAR » study was performed on twenty four molecules of 7-azaïndol. The multiple linear regression method (MLR) was used to quantify the relationship between molecular descriptors and the antiproliferative activity of 7-azaïndol compounds. The predictive power of the model obtained was confirmed by $R = 0,991$, $R^2 = 0,981$ and $F=80.984$. the LOO cross-validation method shows us that a strong correlation was observed between the experimental and predicted values of the biological activities , wich indicates the validity and quality of the QSAR model obtained

Keywords: antiproliferative, 7-AZAINDOLE, C-Met Tyrosine Kinase, QSAR, MLR. LOO.