

Ecole Nationale Supérieure Agronomique El-Harrach, Alger
Ecole Doctorale de Biotechnologies Végétales
Mémoire présenté en vue de l'obtention du diplôme de Magister en Biotechnologies Végétales
Option: Biotechnologies Végétales

***Etude in silico de l'effet d'un ensemble
de flavonoïdes et de coumarines sur le
chaperon moléculaire Hsp90***
Screening virtuel et Docking moléculaire

Par:

Sabki Abdellah

Directeur de Thèse: Kameli A. Prof ENS

Soutenu à l'ENSA, le : 19/06/ 2013

devant le jury composé de : Président : Ounane S.M. Prof ENSA Examineurs : Boudjeniba M. Prof ENS
Redouane N. MCA ENS

Table des matières

Remerciements . . .	5
Résumé . . .	6
ص خ ل م . . .	7
Introduction Générale . . .	8
Chapitre I : Les flavonoïdes . . .	12
1 Introduction . . .	12
2 Classification, nomenclature et sources alimentaires . . .	12
2.1 Les flavonoïdes . . .	13
2.2 Les isoflavonoïdes . . .	16
2.3 Les néoflavonoïdes . . .	17
2.4 Les flavonoïdes mineurs . . .	18
2.5 Les coumarines . . .	18
3 Absorption et métabolisme . . .	19
4 Les flavonoïdes et le cancer . . .	21
4.1 Les études épidémiologiques . . .	21
4.2. Etudes in vitro et in vivo et mécanismes d'action moléculaires . . .	22
Chapitre II : Les Protéines de Choc Thermique Hsp90 . . .	25
1 Introduction . . .	25
2 La famille des chaperons Hsp90 . . .	25
2.1 Gp96, Trap1, et HspN . . .	26
2.2 Hsp90 . . .	27
Chapitre III : Screening Virtuel et Docking Moléculaire . . .	41
1 Introduction . . .	41
2 La stratégie du criblage virtuel . . .	43
1.1 Préparation d'une chimiothèque . . .	43
1.2 Préfiltrage de la chimiothèque ADME/T . . .	44
1.3 Obtention de la structure de la cible . . .	46
1.4 Criblage virtuel . . .	47
2 Le docking moléculaire . . .	51
2.1 Principe . . .	52
Chapitre IV : Matériels et Méthodes . . .	58
1 Matériels . . .	58
1.1 Les programmes et applications Schrödinger . . .	58
1.2 Les Solutions Schrödinger . . .	60
2 Méthodes . . .	61
2.1 Choix des cibles . . .	62
2.2 Préparation des cibles . . .	63
2.3 Test de la fiabilité du programme de docking utilisé, Glide 5.7 . . .	64
2.4 Criblage virtuel d'une petite chimiothèque de flavonoïdes . . .	67
2.5 Criblage virtuel d'une chimiothèque relativement grande de flavonoïdes et de coumarines . . .	69

2.6 Evaluation des résultats . . .	70
Chapitre V : Résultats et Discussion . .	73
1 Résultats du test de la fiabilité du programme Glide 5.7 . .	73
Discussion : . . .	74
2 Résultats du criblage virtuel de la petite chimiothèque de flavonoïdes . .	75
Discussion . . .	76
3 Résultats du criblage de la grande chimiothèque de flavonoïdes et coumarines . .	84
Discussion . . .	84
Bibliographie . .	93
Annexes . .	106
Annexe01 . . .	106

Remerciements

Avant tout, je remercie et je loue Mon Seigneur ##pour qui, quoi que je fasse, je ne saurai décrire l'amour, la reconnaissance, et la gratitude que j'éprouve.

Je remercie aussi du plus profond de mon cœur mon père -qu 'ALLAH illumine sa tombe- et ma mère -qu'ALLAH la protège de tout mal- qui ont tout sacrifié pour me voir épanouir.

Je remercie mon directeur de thèse le professeur Kameli A. qui m'a initié au monde aussi intéressant que fascinant qu'est la bioinformatique que je ne connaissais guère, et qui m'a appris à être référentiel dans toutes les démarches de ma thèse.

J'exprime également ma reconnaissance au jury qui a accepté de juger mon travail : le professeur Ounane SM. en qualité de président du jury, et les professeurs Boudjeniba M. et Redouane N. en qualité d'examineurs.

Grand merci à ma chère femme (je te dois beaucoup), merci à Baali S. qui a proposé le sujet de la thèse, merci à Mme Tigrine S. pour toute son aide précieuse, merci à Ounissi M. notre pionnier en bioinformatique, merci à mon ami Fouzi et mon beau-frère Hamza pour m'avoir prêté leur PC, merci à tous mes enseignants et à tous ceux qui m'ont aidé de près ou de loin, merci tout le monde.

Résumé

L'objectif de cette étude a été d'étudier in silico l'effet inhibiteur potentiel des flavonoïdes du site ATPasique N-terminal du chaperon moléculaire Hsp90 comme mécanisme moléculaire anticancéreux. Au début nous avons élaboré un protocole pertinent de criblage virtuel basé sur la structure de la cible (structure-based virtual screening). Ce protocole de criblage a été appliqué en premier lieu d'une manière poussée sur 25 flavonoïdes des plus abondants dans l'alimentation, et en deuxième lieu, sur une chimiothèque de flavonoïdes et de coumarines de plus de 1000 molécules préfiltrée à partir de plus 5700 structures selon des critères pharmaceutiques, afin d'y découvrir des molécules à effet potentiel contre la Hsp90.

Suite à une évaluation multicritères fine des résultats du premier et deuxième criblages, 16 molécules de différentes sous-classes de flavonoïdes se sont révélées plus ou moins actives contre la Hsp90 in silico, dont 5 sont très prometteuses, à savoir: 1'(-)-épigallocatechine gallate, 1'(-)-épicatéchine gallate, 1'épicatéchine 3-0-(4-0-méthylgallate), la 3,4,2',4'-tetrahydroxy-2-geranyldihydrochalcone, et l'artonine E.

Autant que nous sachions, honnis 1'(-)-épigallocatechine gallate dont l'effet inhibiteur sur le site N-terminal ATPasique de la Hsp90 est controversé, les molécules jugées actives in silico dans cette étude n'ont jamais été évoquées dans la littérature comme inhibiteurs du site ATPasique N-terminal de la Hsp90, et nous serions les premiers à les découvrir, et si elles s'avèrent tels in vitro, ce serait là un nouveau mécanisme moléculaire découvert de l'effet anticancéreux des flavonoïdes.

Motsclés: flavonoïdes, Hsp90, criblage virtuel, docking, anticancéreux.

ص خلم

يهدف هذا العمل إلى دراسة إمكانية تثبيط بروتينات الصدمة الحرارية Hsp 90 بمجموعة من الفلافونيدات والكومارينات كآلية جزيئية مضادة للسرطان باستعمال التقنية البيومعلوماتية : الخريشة الافتراضية باعتماد بنية البروتين المستهدف (*structure-based virtual screening*).

كخطوة أولية، قمنا بإعداد بروتوكول خريشة افتراضية مؤسس على دراسة معمقة حول Hsp90. ثم قمنا بتطبيق البروتوكول المعد سابقاً أولاً: على 25 فلافونيد الأكثر انتشاراً في غذاء الإنسان، وثانياً: على بنك معطيات لأكثر من 1000 فلافونيد وكومارين منتخب حسب معايير صيدلانية من أكثر من 5700 بنية، بحثاً عن جزيئات ذات فعالية كامنة مضادة لـHsp90.

بعد تقييم متعدد المعايير ودقيق لنتائج الخريشة الافتراضية الأولى والثانية، تم الكشف عن 16 مركباً ذات فعالية مثبوتة ضد Hsp90، معظمها يتم الكشف عنه لأول مرة حسب اطلاعنا منها 5 مركبات واعدة جداً هي: 1'(-)-épigallocatechine gallate, 1'(-)-épicatéchine gallate, la 3,4,2',4'-tetrahydroxy-2-geranyldihydrochalcone, l'artoinine E. 3-0-(4-0-méthylgallate), 1'(-)-épicatéchine

ناج الكلمات: فلافونيدات, Hsp90, خريشة افتراضية, إرساء جزيئي, مضاد للسرطان.

Introduction Générale

Depuis l'antiquité, les produits naturels, notamment ceux d'origine végétale ont toujours été une source importante d'agents thérapeutiques. Actuellement, environ 25-30% de tous les médicaments disponibles sont des dérivés de produits naturels.

Les flavonoïdes et leurs conjugués forment un groupe très vaste de ces produits naturels. Ce sont des métabolites secondaires qui s'accumulent en quantités relativement élevées dans les plantes, où ils ont une multitude de fonctions : de structure, de défense, d'attractifs pour les pollinisateurs, d'agents allélopathiques, de protecteurs contre les UV, et de molécules de signalisation au sein de la plante, et dans ses interactions avec l'environnement.

Le rôle des flavonoïdes, comme constituants alimentaires protecteurs, est devenu un domaine, de plus en plus, important de la recherche en nutrition humaine. Contrairement aux vitamines traditionnelles, ils ne sont pas essentiels à court terme à l'organisme, mais il y a un nombre de preuves, grandissant de jour en jour, que des apports faibles peuvent avoir, à long terme, des effets potentiels sur le métabolisme humain d'une manière favorable pour la prévention ou la réduction du risque de maladies dégénératives, comme les maladies cardiovasculaires, le diabète, l'obésité et le cancer.

Le cancer est devenu un danger de santé publique majeur. Selon une enquête de l'OMS, il y a eu 7,4 millions décès à cause du cancer (13% de tous les décès) dans le monde en 2004. Ce nombre ne cesse d'augmenter, malgré les efforts considérables et les dépenses colossales fournis dans le cadre de la lutte contre le cancer, et il est estimé qu'il atteindrait 12 millions environ en 2030.

Le processus cancéreux est caractérisé notamment par des anomalies des voies de signalisation qui se concrétisent par la surexpression de certaines protéines oncogéniques (mutées ou non) et/ou par l'inactivation de produits de gènes suppresseurs de tumeur.

Un mécanisme participant au maintien de ces anomalies d'expression protéique implique la protéine de choc thermique (Hsp90), un chaperon moléculaire. Dans des conditions normales, la Hsp90 ne semble pas participer directement au repliement des protéines de novo synthétisées, mais permet, avec l'aide de co-chaperons moléculaires, à certaines protéines cibles d'acquérir leur conformation active optimum. En condition de stress, la Hsp90 est surexprimée suggérant un rôle protecteur au niveau cellulaire. Cependant, parmi les protéines clientes de la Hsp90, beaucoup d'entre elles sont impliquées dans les pathologies cancéreuses. Il s'agit de kinases, de facteurs de transcription (comme les récepteurs aux androgènes et aux œstrogènes), d'oncogènes et de protéines du cytosquelette (tubuline, actine). Certaines de ces protéines sont impliquées dans le contrôle du cycle cellulaire et de la signalisation. La Hsp90 et les complexes qu'elle forme dans la cellule cancéreuse représentent des cibles d'intérêt dans la stratégie antitumorale.

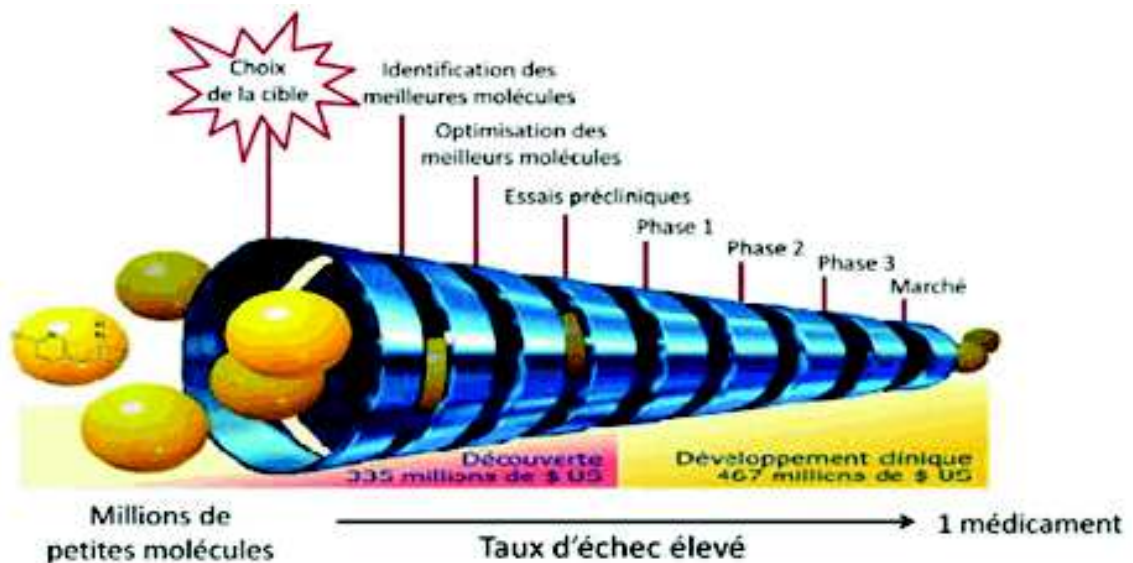
Si la Hsp90 est indispensable au bon fonctionnement des cellules saines, une telle activité représente un inconvénient dans la lutte contre le cancer, car la protection de protéines mutées conduit à la survie de cellules cancéreuses. Inhiber la Hsp90 provoque la dégradation des complexes protéiques formés de protéines-clients et de la Hsp90 par le protéasome, ce qui élimine, entre autres, les protéines oncogènes indispensables à la

croissance de la tumeur. En plus de stopper le phénomène de la signalisation proliférative et de favoriser l'apoptose, l'inhibition de la Hsp90 cible les autres caractères phénotypiques majeurs du cancer (angiogénèse, invasion cellulaire, migration...etc.).

Plusieurs inhibiteurs de la Hsp90, utilisés seuls ou en combinaison, sont actuellement évalués dans le cadre d'essais cliniques de phases II/III. Ces molécules exercent majoritairement leur activité via l'inhibition du site N-terminal de liaison à l'ATP, bien caractérisé. Leur utilisation en clinique présente cependant quelques inconvénients liés pour certains à leur faible solubilité en milieu aqueux et pour d'autres aux effets toxiques secondaires qu'ils induisent. Aujourd'hui, si les mécanismes de l'inhibition de la Hsp90 sont mieux compris, les conséquences de cette inhibition et le mécanisme de l'activité antitumorale des inhibiteurs sont très complexes et ne restent que partiellement élucidés.

La pression est grandissante pour les industries pharmaceutiques qui doivent trouver de nouveaux médicaments. Avec un coût de l'ordre de 800 millions de dollars répartis sur plus d'une dizaine d'années pour qu'une molécule parvienne sur le marché, celles-ci s'interrogent sur le taux élevé des échecs dans les phases de développement clinique, un moment où les sommes d'argent avancées sont les plus grandes.

Lors de ces phases, le candidat-médicament est d'abord testé sur des volontaires sains pour étudier son comportement dans l'organisme (pharmacocinétique, biodisponibilité) et estimer sa tolérabilité (phase I), puis il est testé sur des patients pour évaluer son efficacité en parallèle de sa toxicité (phase II) et doit enfin prouver qu'il est plus efficace et moins toxique que le traitement actuel (phase III) pour justifier son entrée sur le marché.



Une fois la cible pharmaceutique identifiée et validée, plusieurs options sont possibles, dont celle du criblage haut-débit robotisé, pouvant être employée comme point de départ d'un programme de recherche de nouveaux médicaments. Cette approche permet l'évaluation, dans un temps très court, des propriétés pharmacologiques de l'ensemble des molécules contenues dans une chimiothèque sur une ou des cible(s) d'intérêt pharmacologique donnée(s).

Cette technique permet la recherche à grande échelle de substances actives sur une cible caractérisée (quelques milliers à quelques millions de composés testés en parallèle sur celle-ci). Initialement développé dans l'industrie pharmaceutique, le criblage haut-débit

devient progressivement accessible aux laboratoires académiques grâce à l'émergence de plateformes techniques, actuellement encore en nombre limité.

L'obtention de résultats positifs et leur confirmation conduisent à l'identification de touches (ou "hits"). Il s'agit de composés interagissant significativement plus que la moyenne des autres composés testés sur la cible visée. Toutes les touches identifiées ne présentent pas obligatoirement les critères pour devenir un candidat-médicament (p. ex. : compétitif vis-à-vis d'une molécule de référence, actif in vivo...).

Aperçu d'une plateforme de criblage haut-débit robotisé et d'une plaque à 96 puits.



Le criblage à haut débit, par le coût de ses tests (estimé à ~1 € par molécule), ne permet pas de tester des millions de composés sans disposer de moyens financiers conséquents. Pour cette raison, les compagnies pharmaceutiques peuvent avoir recours aux techniques de criblage virtuel en complément du criblage robotisé. En milieu académique, pour cause de moyens plus limités, cette approche est davantage privilégiée.

Ces techniques in silico sont basées sur les connaissances accumulées à propos du système étudié et qui peuvent être : la structure de ligands de référence ("ligand-based drug design") ou la structure tridimensionnelle de la cible ("structure-based drug design"). Ces techniques se révèlent être très utiles pour réduire les temps de recherche d'identification des nouvelles touches et, par conséquent, les coûts qui y sont associés. En effet, elles permettent d'explorer de façon exhaustive l'espace chimique et de proposer au final une liste raisonnable (financièrement parlant) de molécules à tester expérimentalement et/ou à synthétiser. Cette rationalisation de l'identification de nouvelles touches diffère de l'approche aléatoire du criblage haut-débit exécuté "à l'aveugle", mais permet ainsi, en la complétant, de l'étendre aux cibles de plus en plus complexes de la recherche pharmaceutique.

L'avantage du criblage virtuel est de pouvoir tester in silico aussi bien des composés déjà synthétisés que d'autres qui ne le sont pas encore et qui peuvent avoir été générés par un programme informatique. La diversité moléculaire des chimiothèques criblées, expérimentalement ou virtuellement, est un critère important pour mieux couvrir l'espace chimique et donc pour augmenter les chances d'identifier de nouvelles touches. Cette diversité peut, entre autres, être obtenue par les techniques de synthèse combinatoire ou par l'accès aux substances naturelles issues de la biodiversité. Les molécules doivent également avoir les propriétés physico-chimiques caractérisant un médicament. Pour cela, des règles élémentaires permettent de filtrer rapidement hors des bibliothèques criblées celles ne répondant pas à ces règles ; les plus connues sont celles de Lipinski.

L'effet anticancéreux des flavonoïdes a été étudié et évalué dans un grand nombre d'études épidémiologiques, in vitro et in vivo, et plusieurs mécanismes moléculaires de cet

effet ont été élucidés et d'autres ont été proposés ; l'objectif de cette étude est d'investiguer in silico l'effet potentiel des flavonoïdes comme inhibiteurs du site ATPasique N- terminal de la Hsp90, qui n'a été guère étudié, autant que nous sachions.

Pour ce faire, nous avons élaboré un protocole pertinent de criblage virtuel basé sur la structure de la cible (structure-based virtual screening), qui consiste à tester les molécules d'intérêt par le docking moléculaire contre la cible, avec plusieurs enrichissements basés sur l'étude approfondie de la Hsp90. Ce protocole de criblage a été réalisé en premier lieu d'une manière poussée sur 25 flavonoïdes des plus abondants dans l'alimentation, et en deuxième lieu, sur une chimiothèque de flavonoïdes et de coumarines de plus de 1000 molécules préfiltrée à partir de plus 5700 structures selon des critères pharmaceutiques, afin d'y découvrir des molécules à effet potentiel contre la Hsp90.

Chapitre I : Les flavonoïdes

1 Introduction

Les flavonoïdes et leurs conjugués forment un groupe très vaste de produits naturels. C'est, qualitativement et quantitativement, l'un des plus grands groupes de produits naturels connu.

Ce sont des métabolites secondaires qui s'accumulent en quantités relativement élevées dans les plantes, où ils ont une multitude de fonctions : de structure, de défense, d'attractifs pour les pollinisateurs, d'agents allélopathiques, de protecteurs contre les UV, et de molécules de signalisation au sein de la plante, et dans ses interactions avec l'environnement.

Ils ont été utilisés comme colorants, fibres, colles, huiles, cires, agents aromatisants, remèdes, et parfums, et sont considérés comme des sources potentielles de nouveaux médicaments, d'antibiotiques naturels, d'insecticides, et d'herbicides (Croteau, Kutchan and Lewis, 2000; Dewick, 2002).

Le rôle des flavonoïdes, comme constituants alimentaires protecteurs, est devenu un domaine, de plus en plus, important de la recherche en nutrition humaine. Contrairement aux vitamines traditionnelles, ils ne sont pas essentiels à court terme à l'organisme, mais il y a un nombre de preuves, grandissant de jour en jour, que des apports faibles peuvent avoir, à long terme, des effets potentiels sur le métabolisme humain d'une manière favorable pour la prévention ou la réduction du risque de maladies dégénératives, comme les maladies cardiovasculaires, le diabète, l'obésité et le cancer (Anderson et al., 1999)

2 Classification, nomenclature et sources alimentaires

Les flavonoïdes sont des composés polyphénoliques comprenant un squelette de base en C6-C3-C6, avec 2 cycles aromatiques, A et B, reliés par un hétérocycle C. En fonction de la position de l'attachement du noyau B à l'hétérocycle C, ce groupe de produits naturels peut être divisé en trois classes (Marais et al., 2006) :

- Les flavonoïdes 1,
- Les isoflavonoïdes 2,
- Les néoflavonoïdes 3.

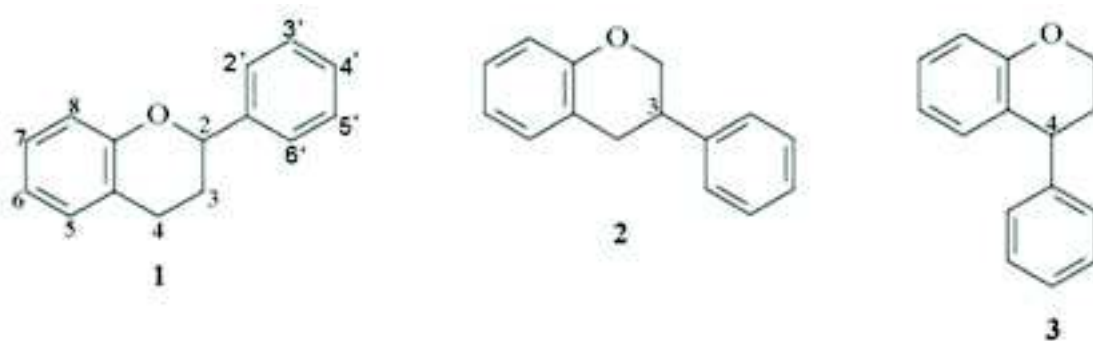


Figure 1 : classes des flavonoïdes. 1 flavonoïdes, 2 isoflavonoïdes, 3 néoflavonoïdes.

Les flavonoïdes peuvent être modifiés par hydroxylation, méthylation, ou par O-glycosylation des groupes hydroxyles, ainsi que par C-glycosylation directement à un atome de carbone du squelette de base. En outre, des groupes alkyles (souvent des prényles) peuvent s'y attacher de manière covalente, et parfois d'autres cycles sont condensés au squelette de base. Les glycosides de flavonoïdes sont souvent acylés avec des molécules d'acides aliphatiques ou aromatiques (Kachlicki and Stobiecki,

2006). Ces caractéristiques peuvent à elles seules expliquer l'existence d'environ 3×10^5 flavonoïdes. Le nombre réel des flavonoïdes qui ont été trouvés à ce jour, et dont la structure a été complètement élucidée est grand, mais, probablement, ne dépasse pas 1 % du nombre théorique des variantes possibles (Havsteen, 2002).

En dehors de quelques exceptions, les flavonoïdes sont presque tous liés à des sucres de différentes natures. La glycosylation, et par la suite l'acylation des sucres (liaison d'un ou plusieurs acides, notamment l'acide p-coumarique et l'acide caféique), ainsi que la méthylation, modifient les propriétés biologiques, mais aussi physico-chimiques des flavonoïdes (Kong et al., 2003; Aron and Kennedy, 2007).

2.1 Les flavonoïdes

En fonction du degré de l'oxydation et du type de la saturation au niveau de l'hétérocycle C, les flavonoïdes typiques peuvent être divisés en les sous-classes suivantes (Marais et al., 2006) :

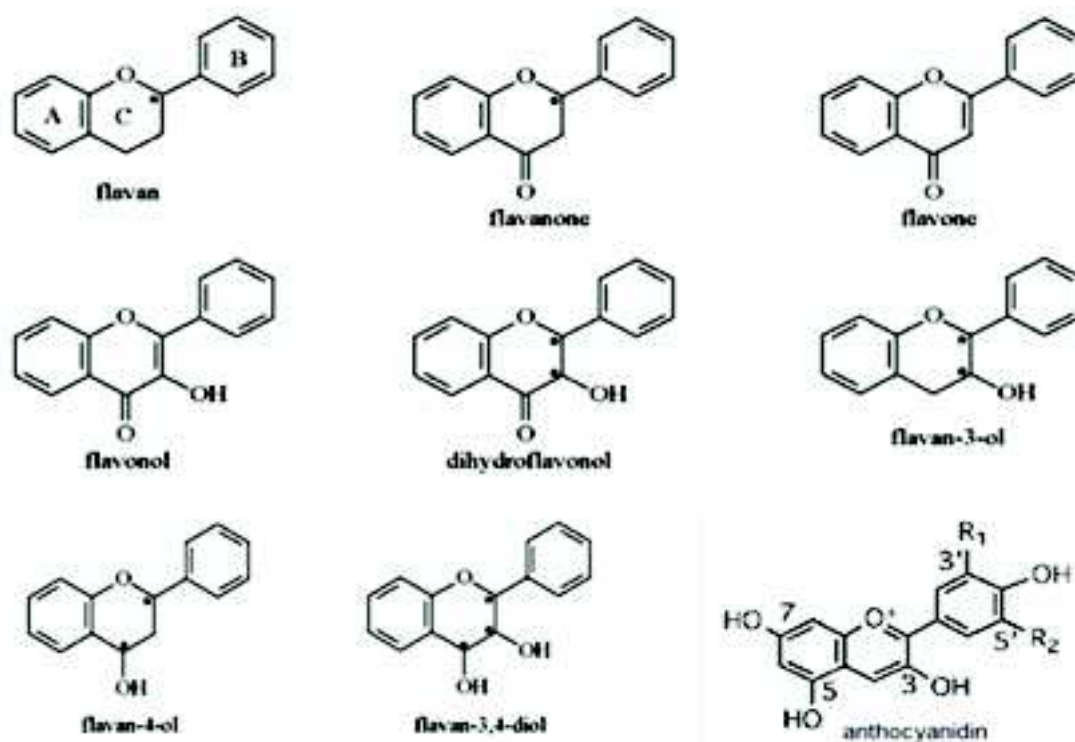


Figure 2 : sous-classes des flavonoïdes.

2.1.1 Les flavonols

Les flavonols sont les plus répandus des flavonoïdes dans la nature et dans l'alimentation. Ils sont présents dans une grande variété de fruits et de légumes (Ross and Kasum, 2002).

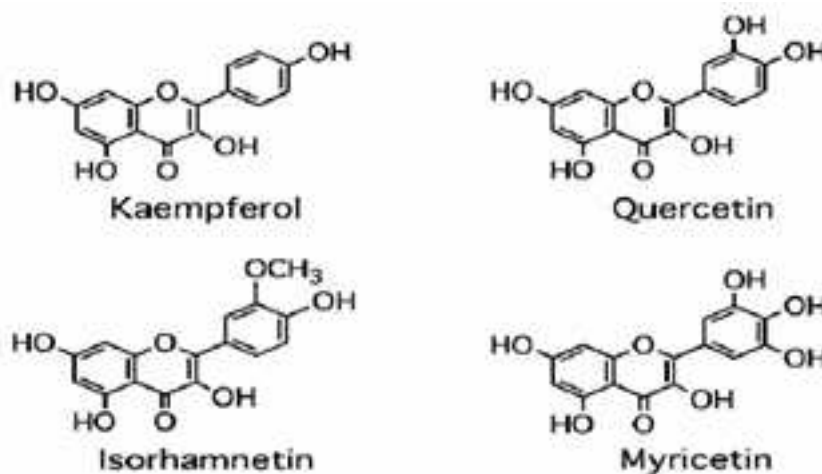


Figure 3 : flavonols aglycones communs.

Ils sont représentés principalement par la quercétine, le kaempférol, et la myricétine ; tandis que le dérivé méthylé isorhamnétine est également assez courant.

2.1.2 Les flavones

Contrairement aux flavonols, les flavones ne sont pas largement distribuées dans le monde végétal, avec des concentrations importantes signalées dans seulement : le céleri, le persil, et l'artichaut (Crozier, Ashihara and Clifford, 2006). Les flavones sont principalement représentées dans l'alimentation par l'apigénine, la lutéoline et la chrysin.

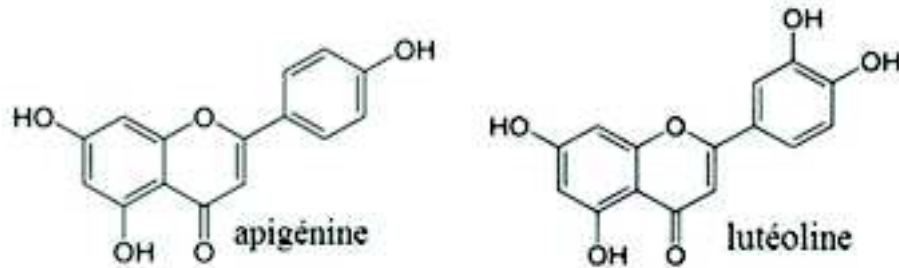


Figure 4 : flavones aglycones communes.

2.1.3 Les flavan-3-ols

Flavan-3-ols sont structurellement une sous-classe complexe de flavonoïdes, allant des simples monomères de (+)-catéchine et de son isomère (-)-épicatéchine aux oligomères et polymères proanthocyanidines, qui sont également connues sous le nom de tannins condensés (Crozier, Jaganath and Clifford, 2006).

Les flavan-3-ols se trouvent en abondance dans les fruits comme les abricots, les cerises acides, les raisins et les mûres (Gu et al., 2004). Ils peuvent subir une estérification avec l'acide gallique pour former des catéchines gallates, et des hydroxylations pour former des gallocatéchines. La (+)-gallocatéchine, l'(-)-épigallocatéchine, l'(-)-épigallocatéchine gallate, et l'(-)-épicatéchine gallate sont abondantes dans les infusions de thé vert (Stewart, Mullen and Crozier, 2005).

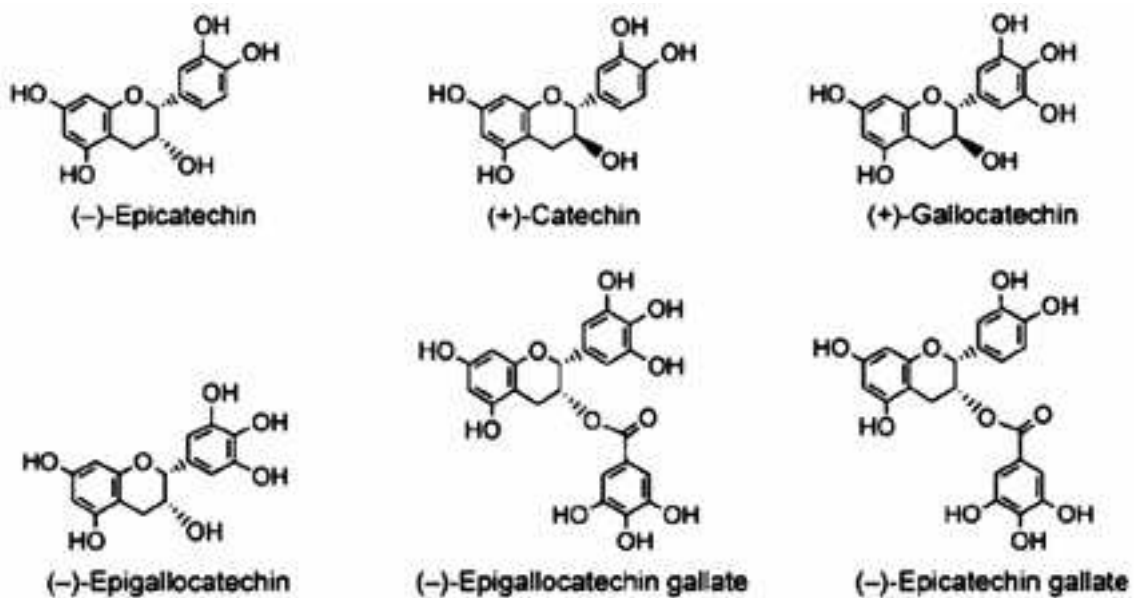


Figure5: flavan-3-olscommuns.

2.1.4 Lesflavanones

Les flavanones sont représentées dans l'alimentation principalement par la naringénine, l'hespérétine et l'ériodictyol.

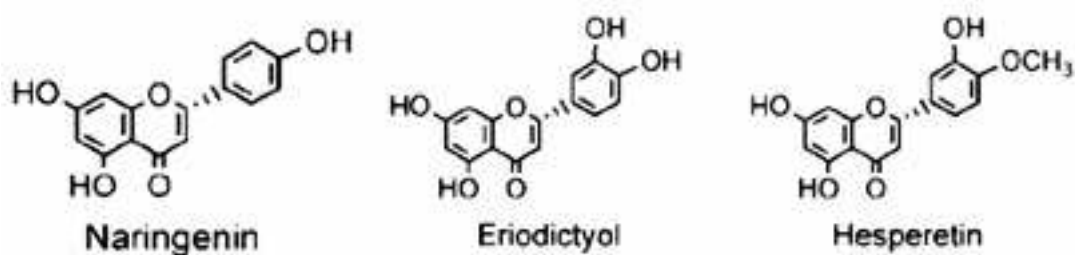


Figure 6 : flavanones aglycones communes

2.1.5 Les anthocyanidines et les anthocyanes

Les anthocyanidines sont des pigments végétaux. Elles se trouvent principalement sous forme de glycosides hydrosolubles, les anthocyanes, dans une variété de fruits (Wu et al., 2006). Il y a environ 17 anthocyanidines trouvées dans la nature, mais seulement 6 (cyanidine, delphinidine, pétunidine, péonidine, pélargonidine, malvidine) sont ubiquitaires et d'une importance alimentaire. Contrairement aux autres sous-groupes des flavonoïdes avec le même squelette C6-C3-C6, les anthocyanes ont une charge positive à pH acide (Prior and Wu, 2006).

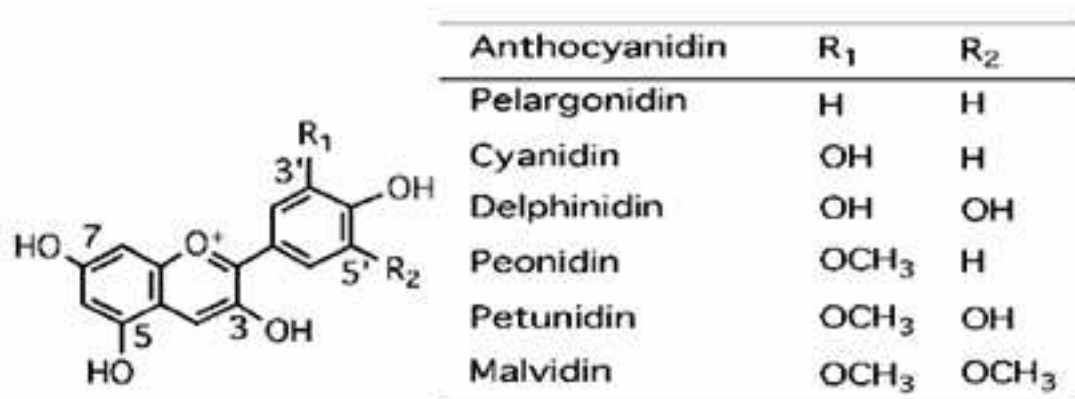


Figure 7 : anthocyanidines communes.

2.2 Les isoflavonoïdes

Contrairement à la plupart des autres flavonoïdes, les isoflavones sont caractérisés par le cycle B fixé au C3 de l'hétérocycle central C, plutôt qu'au C2.

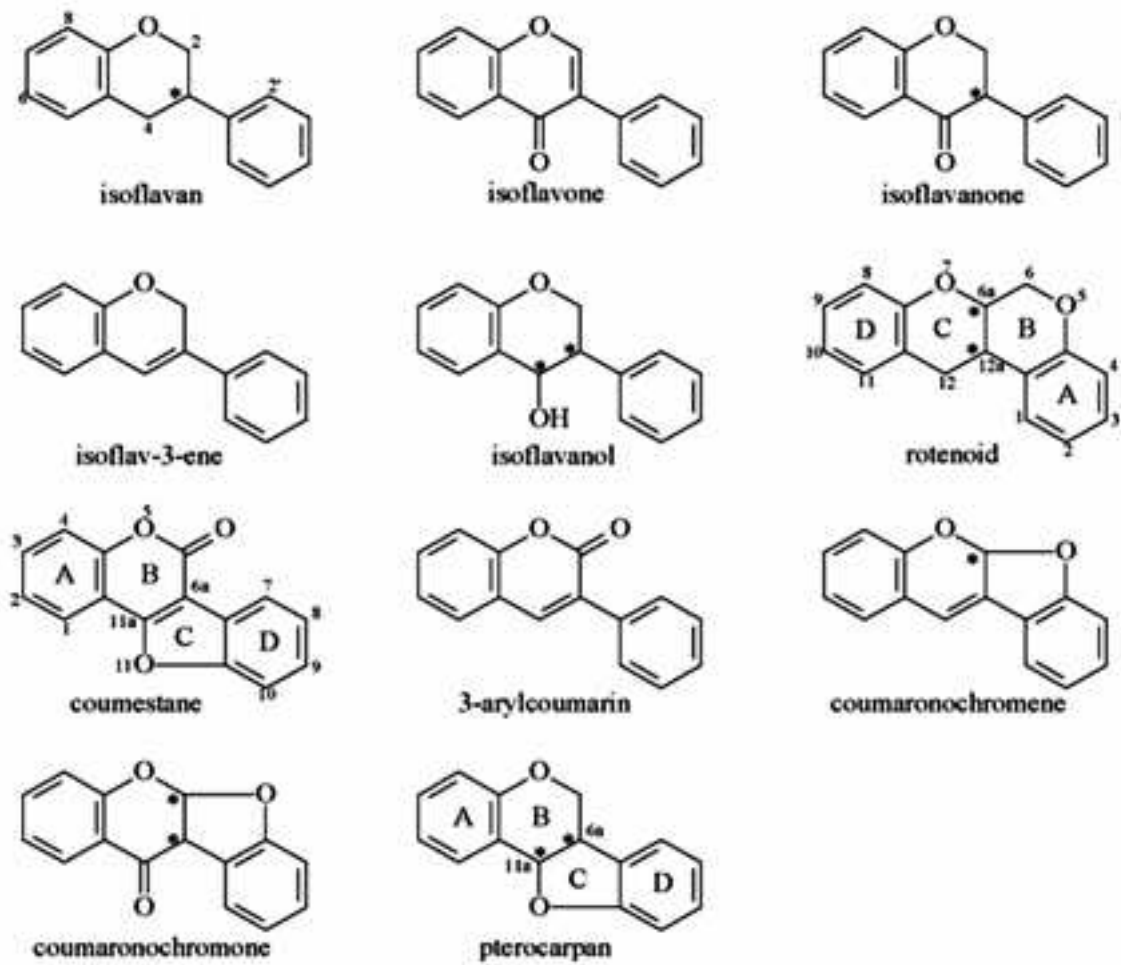


Figure 8 : sous-classes des isoflavonoïdes.

Ils ont une distribution très limitée dans le règne végétal, avec des quantités importantes dans les espèces légumineuses (Graham, 1991; Dixon and Steele, 1999).

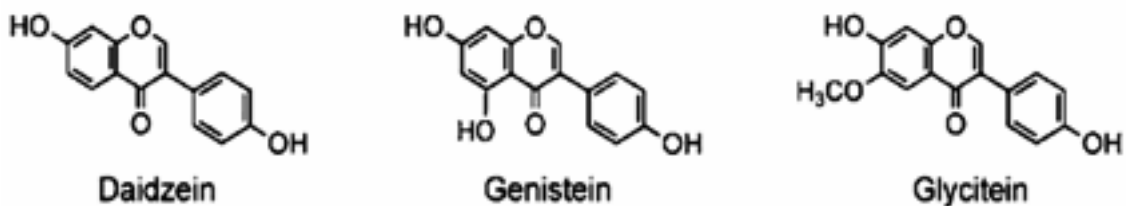


Figure 9 : isoflavonoïdes communs.

Ils sont représentés dans l'alimentation principalement par la daidzéine, la génistéine, et la glycitéine.

2.3 Les néoflavonoïdes

Les néoflavonoïdes sont très proches dans la structure et l'origine biosynthétique aux flavonoïdes et les isoflavonoïdes ; ils comprennent les 4-arylcoumarines, les 3,4-dihydro-4-arylcoumarines, et les néoflavènes (Marais et al., 2006).



Figure 10 : sous-classes des néoflavonoïdes.

2.4 Les flavonoïdes mineurs

Les composés naturels comme les chalcones et les aurones, comprennent aussi un squelette en C6-C3-C6, et sont considérés comme étant des flavonoïdes mineurs, dont les 2'-hydroxychalcones, les 2'-OH- dihydrochalcones, les 2'-OH-retro-chalcones, les aurones, et les auronols (Marais et al., 2006).

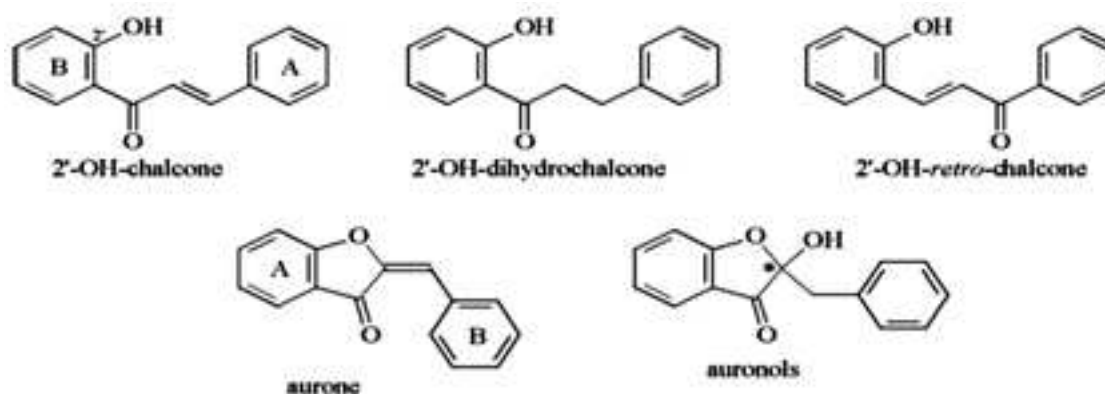


Figure 11 : structures de quelques flavonoïdes mineurs.

2.5 Lescoumarines

Les coumarines sont des hétérocycles oxygénés ayant comme structure de base le benzo-2-pyrone, C6-C3.

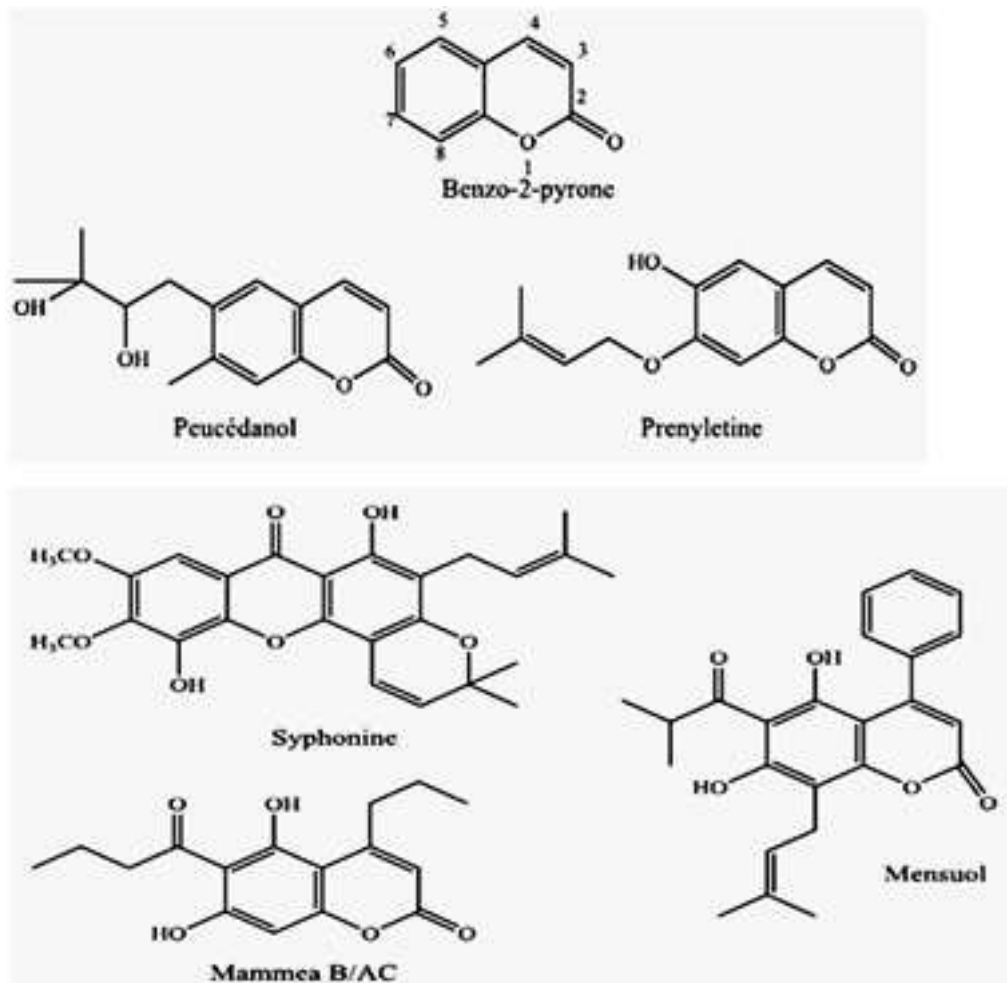


Figure 12 : structures de quelques coumarines.

Elles ont été isolées pour la première fois par Vogel en 1820 dans le *Coumarouna odorata*. Près de 1000 composés coumariniques ont été isolés dans plus de 800 espèces de plantes et dans les microorganismes. Dans les plantes, on les rencontre chez les Apiacées, les Astéracées, les Fabacées, les Rosacées, les Rubiacées, les Rutacées et les Solanacées (Smyth, Ramachandran and Smyth, 2009).

3 Absorption et métabolisme

Bien que les études sur la biodisponibilité et le métabolisme des flavonoïdes en aient dévoilé beaucoup d'aspects, l'information en la question reste, à la fois, très fragmentée et contradictoire, et malgré les variations quantitatives considérables entre différentes études, même pour le même composé, quant aux proportions absorbées, excrétées, dégradées ou métabolisées par différentes voies, certaines généralisations peuvent être faites.

Il a été estimé que l'apport alimentaire en flavonoïdes est d'environ 100-

200 mg/jour ; environ 5 % de cet apport est absorbé par l'intestin, et environ 95 % passe dans le côlon.

Il est établi que la nette absorption intestinale de quelques flavonoïdes individuels peut atteindre jusqu'à 10-20% de l'apport dans des conditions expérimentales, mais dans des conditions normales, elle ne dépasse pas probablement 1-2%, et certains composés peuvent avoir encore des taux plus bas.

Les flavonoïdes plus hydrophobes, consommés sous forme de glucosides, semblent être les mieux absorbés, généralement via déglycosylation enzymatique, suivie de la diffusion de l'aglycone à travers les membranes des cellules intestinales. Des composés plus polaires, tels que les catéchines, sont relativement mal absorbés, et les flavonoïdes glycosylés, autres que les glucosides, sont résistants à la déglycosylation, et sont lentement absorbés par diffusion, ou éventuellement par les transporteurs des sucres. Basé sur des mesures de l'excrétion urinaire, les anthocyanines semblent être à peine absorbées, mais les données récentes suggèrent que leur absorption peut être, à la fois, relativement importante et rapide, mais contrebalancée par une dégradation très rapide, et par conséquent, une très courte demi-vie plasmatique.

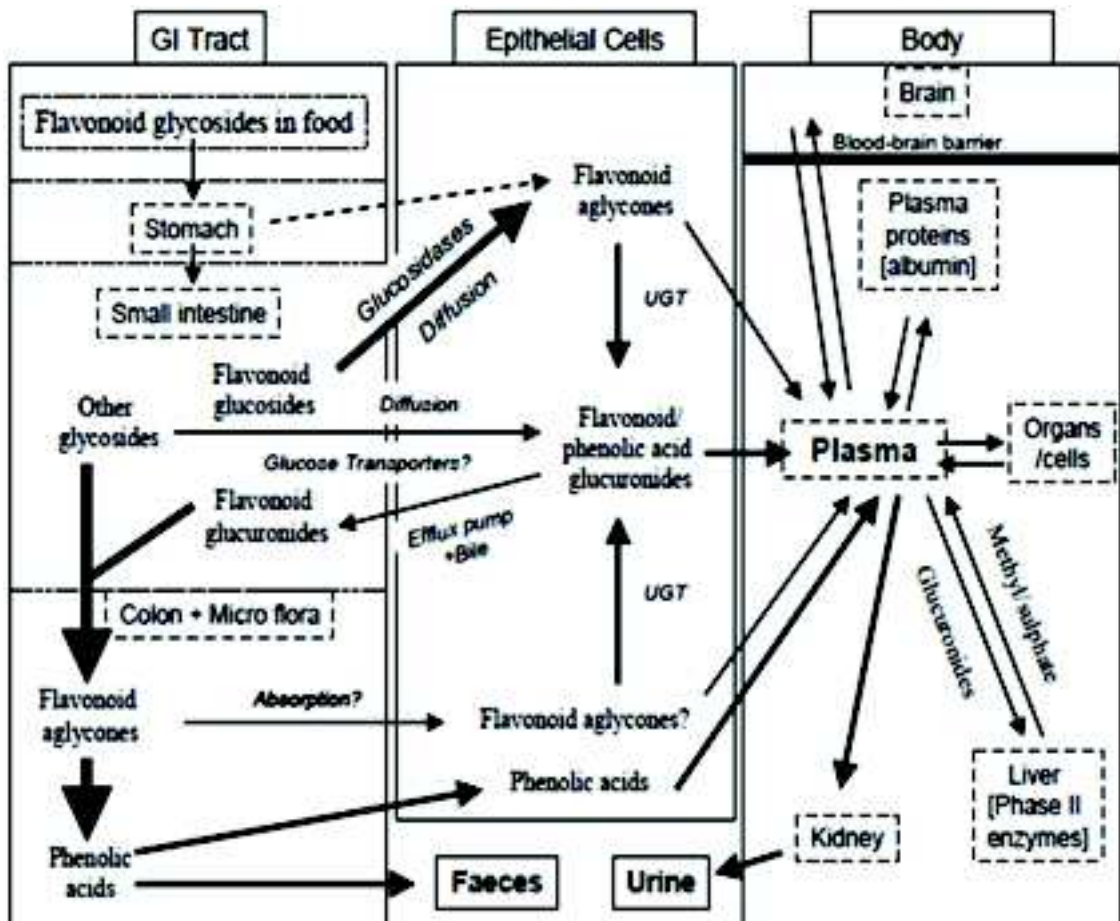


Figure 13 : Résumé de la grande complexité du métabolisme des flavonoïdes, tel qu'il est actuellement compris. L'épaisseur de flèches est une indication de l'importance quantitative relative des flux. UGT = Uridine diphosphoglucuronyl transférase

Il semble que le plasma ne contient, généralement, que des traces de la plupart des flavonoïdes aglycones ou de leurs formes glycosidiques originales. Un très grand, mais assez variable, pourcentage est sous la forme des conjugués métaboliques polaires, les glucuronides et les sulfates. Les flavonoïdes contenus dans le plasma semblent être

principalement liés à l'albumine sérique et à d'autres possibles protéines, et ce fait, avec leur efflux par les pompes cellulaires MDR (multi-drug resistance), limite leur biodisponibilité dans les organes, notamment le cerveau. La seule exception apparente est le foie, qui peut accumuler activement des flavonoïdes glucuronidiques à des concentrations relativement élevées, et semble modifier leur conjugaison. Il semble que l'intestin, principalement, glucuronide les flavonoïdes, alors que le foie peut les déglucuronider, et/ou les sulfater, et/ou les méthyler.

Il semble qu'environ 90 % de l'apport alimentaire de la plupart des flavonoïdes passe dans le côlon, soit directement, après leur absorption et efflux par les pompes MDR des cellules intestinales, ou par excrétion biliaire.

La microflore colique peut déglycosyler la plupart des flavonoïdes glycosidiques, et, par conséquent, potentiellement augmenter leur biodisponibilité. Mais elle peut aussi dégrader ensuite leurs aglycones en acides phénoliques, et en composés plus simples encore, tels que l'acide benzoïque, l'acide acétique et l'anhydride carbonique.

Les acides phénoliques semblent être généralement très bien absorbés, et peuvent avoir une importante contribution indirecte aux effets bénéfiques sur la santé attribués aux flavonoïdes dont ils ont été le résultat de dégradation (Stevenson, Scheepens and Hurst, 2009).

4 Les flavonoïdes et le cancer

Les effets anticancéreux des flavonoïdes ont fait l'objet d'innombrables études épidémiologiques, in vitro, in vivo et même cliniques.

4.1 Les études épidémiologiques

Un grand nombre d'études épidémiologiques a été mené pour prouver l'effet protecteur des flavonoïdes contre le cancer.

La consommation d'isoflavones a été associée à une diminution du risque des cancers liés à l'œstrogène.

Des apports élevés en quercétine et en kaempférol ont été associés à une réduction du risque du cancer du poumon, et à une réduction de 40 à 50 % du risque du cancer de l'estomac.

Une inverse association a été rapportée entre des apports élevés en flavonoïdes (avec l'apport total en bêta-carotène et en vitamine E aussi) et le risque de différents cancers (cavité buccale, du pharynx, du larynx, de l'œsophage, des poumons, de l'estomac) (Batra and Sharma, 2013).

Une inverse association a été rapportée entre des apports élevés en flavonoïdes totaux (0,56), en flavanones (0,51), en flavonols (0,62), et en flavanols (0,64) et le risque du cancer de la bouche ; en flavonoïdes totaux (0,60), en flavanones (0,60), et en flavonols (0,32) et le risque du cancer du larynx ; en flavanones (0,38) et le risque du cancer de l'œsophage ; en anthocyanidines (0,67), en flavonols (0,64), en flavones (0,78), en isoflavones (0,76), et en proanthocyanidines et le risque du cancer colorectal ; en flavones (0,81) et en flavonols (0,80) et le risque du cancer du sein ; en flavonols (0,63) et en isoflavones (0,51) et le cancer

de l'ovaire ; en flavonols (0,69) et en flavones (0,68) et le risque du cancer du rein (Rossi et al., 2010) ; en flavonols et des flavones et le carcinome des cellules rénales (Bosetti et al., 2007).

4.2. Etudes in vitro et in vivo et mécanismes d'action moléculaires

D'innombrables études ont évalué les effets anticancéreux des flavonoïdes in vitro et in vivo, et ont essayé d'élucider les mécanismes moléculaires derrière ces effets (Batra and Sharma, 2013; Ren et al., 2003; Kanadaswami et al., 2005; Le Marchand, 2002; Birt, Hendrich and Wang, 2001).

In vitro, plusieurs flavonoïdes et extraits flavonoïdiques de plantes ont montré de puissants effets inhibiteurs sur différents types de facteurs de croissance, de facteurs de transcription, de récepteurs, de protéines oncogènes, d'hormones... etc. qui ont un rôle clé dans processus de la carcinogénèse.

<i>Cancer</i>	<i>Cell</i>	<i>Flavonoïd</i>
Human oral cancer	HSC-2, HSG, SCC-25	Flavanones, isoflavans, EGC, chalcones, EGCG, curcumin, genistein, ECG, quercetin, cisplatin
Human breast cancer	MCF-7	Flavanones, daidzein, genistein, quercetin, luteolin
Human thyroid cancer	ARQ, NPA, WRO	Genistein, apigenin, kaempferol, chrysin, luteolin, biochanin A
Human lung cancer	SK-LU1, SW90Q, H441, H661, haGo-K-1, A549	Flavone, quercetin
Human prostate cancer	LNCaP, PC3, DU145	Catechin, epicatechin, quercetin, kaempferol, luteolin, genistein, apigenin, myricetin, silymarin
Human colon cancer	Caco-2, HT-29, IEC-6, HCT-15	Flavone, quercetin, genistein, anthocyanin
Human leukaemia	HL-60, K562, Jurkat	Apigenin, quercetin, myricetin, chalcones
B16 mouse melanoma	4A5	Chalcones

Table 1: effets antiprolifératifs des flavonoïdes contre des lignées cellulaires cancéreuses.

Ils ont également montré de puissants effets antiprolifératifs contre un grand nombre de lignées cellulaires cancéreuses.

In vivo, des études sur plusieurs modèles animaux de cancers, ainsi des xénogreffes de cancers sur des animaux ont montré l'effet inhibiteur de plusieurs flavonoïdes et extraits flavonoïdiques de plantes sur différents stages du processus de la carcinogénèse, notamment : l'initiation, la promotion, et la progression des tumeurs.

Les principaux mécanismes moléculaires de ces effets anticancéreux sont :

4.2.1 Inhibition des PK (protein kinases)

La phosphorylation des protéines est un mécanisme important de transduction du signal intracellulaire impliqué dans diverses réponses cellulaires, comme la régulation de la croissance cellulaire et la prolifération.

Un mécanisme possible pour les effets anticancéreux potentiels des flavonoïdes pourrait être l'inhibition de diverses PK, telles que les PKC (protein kinase C), les EGFR (Epidermal Growth Factor) tyrosine kinases, et les FAK (Focal Adhesion kinase), bloquant ainsi la transduction du signal de la prolifération cellulaire.

4.2.2 Inhibition des enzymes prooxydantes et propriétés antioxydantes

La formation d'espèces réactives de l'oxygène (ROS) est une étape importante dans la promotion des tumeurs et leur progression. Les ROS agissent comme messagers secondaires dans plusieurs voies qui conduisent à l'hyperprolifération cellulaire, à la résistance à l'apoptose, et à l'activation de proto-oncogènes.

Les flavonoïdes pourraient exercer une partie de leurs propriétés anticarcinogéniques et antiprolifératives par l'intermédiaire de l'inhibition directe d'enzymes prooxydantes, telles que les lipoxygénases (LOX), les cyclooxygénases (COX), et les xanthine oxydases (XO), dont le cycle catalytique implique des ROS.

Outre l'inhibition des enzymes prooxydantes, les flavonoïdes sont de puissants antioxydants naturels (piégeage direct de radicaux libres) *in vitro*, et pourraient exercer une partie de leurs effets anticancéreux potentiels à travers leur activité antioxydante.

Cette activité pourrait être modulée par la liaison des flavonoïdes aux protéines. Les complexes albumines-flavonoïdes, avec une affinité pour les LDL, peuvent agir en tant que vrais antioxydants plasmatiques. En outre, la complexation des flavonoïdes à des protéines pourrait leur assurer une protection contre la dégradation oxydative.

4.2.3 Modulation du métabolisme des carcinogènes

L'activation des procarcinogènes en cancérogènes est une étape importante dans la carcinogénèse qui peut être modulée par des flavonoïdes.

Les flavonoïdes peuvent exercer leur effet possible par deux mécanismes. Tout d'abord, en interagissant avec les enzymes de la phase I du métabolisme, telles que les isozymes du cytochrome P450 (CYP 1A1, 1A2, 1B1, 3A4, 3A5, etc.), qui sont impliqués dans l'activation du métabolisme d'un grand nombre de procarcinogènes endogènes et xénobiotiques en cancérogènes.

Le deuxième mode d'action peut être la désintoxication et l'élimination des substances cancérigènes par l'induction des enzymes de la phase II du métabolisme, comme l'UDP-glucuronyl transférase, la quinone réductase et la glutathion-S-transférase.

4.2.4 Inhibition de la multirésistance (multidrug resistance)

Les cellules cancéreuses surexpriment typiquement la P-glycoprotéine (Pgp) ou protéine associée à la multirésistance (MRP), qui est un transporteur ABC (ATP-binding cassette) transmembranaire capable d'expulser une grande variété de médicaments utilisés dans le traitement du cancer, au prix de l'hydrolyse de l'ATP. Ce phénomène est connu sous le nom de multirésistance (multidrug resistance, MDR), et l'inhibition de la MDR, pour empêcher

l'efflux des médicaments, représente une application clinique potentielle pour le traitement du cancer.

Les flavonoïdes modulent la multirésistance cellulaire médiée par la Pgp en inhibant sa surexpression, en l'inhibant directement (liaison avec grande affinité à son site ATPasique), et en empêchant son interaction avec les médicaments.

4.2.5 Antiangiogénèse

L'angiogénèse, la formation de nouveaux vaisseaux sanguins, est un processus important qui est régulé par des facteurs angiogéniques et angiostatiques endogènes. Toute altération de ce processus étroitement régulé peut conduire à une croissance persistante et incontrôlée, et à la métastase des tumeurs.

Les flavonoïdes ont été signalés comme de puissants inhibiteurs de l'angiogénèse. Ils peuvent entraîner un manque de diffusion des nutriments et de l'oxygène dans les cellules cancéreuses à forte croissance, et par conséquent, conduire à la mort cellulaire.

Ils peuvent interférer avec les différentes étapes de l'angiogénèse, telles que la prolifération et la migration des cellules endothéliales et la formation de la lumière par différents mécanismes.

4.2.6 Induction de l'apoptose et l'arrêt du cycle cellulaire

L'apoptose est une mort cellulaire programmée pour éliminer les cellules endommagées ou indésirables. Elle est régie par une série de gènes qui peuvent, favoriser l'apoptose, ou la survie des cellules en réponse à des stimuli internes ou externes. La dérégulation de l'apoptose joue un rôle crucial dans l'oncogénèse. Parmi ces régulateurs, le suppresseur de tumeur p53 joue un rôle essentiel dans le contrôle du cycle cellulaire, de l'apoptose, de l'intégrité du génome, et de la réparation de l'ADN. Plusieurs flavonoïdes induisent l'apoptose et l'arrêt du cycle cellulaire en :

- activant principalement p53 et d'autres facteurs relatifs.
- en bloquant une ou plusieurs étapes de la voie de signalisation du NF- κ B (Nuclear factor-kappa B), importante pour la croissance et la transformation cellulaire, la suppression de l'apoptose, l'invasion, la métastase, la chimiorésistance, la radiorésistance et l'inflammation.
- en réprimant AP-1 (activateur protéin-1) et en modulant l'activation des gènes cibles d'AP-1, importants modulateurs de l'invasion et la métastase, de l'angiogénèse, de la prolifération, de la différenciation, et de la persistance.
- en interagissant avec plusieurs autres facteurs et signaux régulant l'apoptose et le cycle cellulaire.

Chapitre II : Les Protéines de Choc Thermique Hsp90

1 Introduction

Les chaperons moléculaires ont été découverts à l'origine comme étant des protéines exprimées sous des conditions de stress, entre autres, les chocs thermiques, d'où elles ont été appelées : protéines de choc thermique, ou Heat Shock Proteins, Hsp (Tissieres, Mitchell and Tracy, 1974).

La réponse au choc thermique correspond à l'activation de facteurs de transcription appelés Heat Shock Factors (HSFs). Ces protéines se lient à l'ADN au niveau de séquences spécifiques situées dans le promoteur des gènes des Hsp : les HSE (Heat Shock Elements), et induisent l'expression des différentes Hsp. Cette réponse a été décrite à l'origine comme survenant après un choc thermique, mais de nombreux autres stimuli sont également capables d'induire l'activation de ces facteurs de transcription (Sorger and Pelham, 1988; Nakai, 1999; Nover et al., 1996; Wu, 1995).

La plupart des inducteurs de stress altèrent les structures des protéines cellulaires, perturbant ainsi le métabolisme cellulaire et entraînant une toxicité. Les Hsp sont alors capables de s'associer avec ces protéines altérées et d'empêcher leur agrégation. Elles peuvent également redonner aux protéines nouvellement formées ou anormales leur conformation tridimensionnelle adéquate (repliement des protéines) (Burel et al., 1992).

Cependant, beaucoup de ces protéines sont abondantes dans la cellule et fonctionnellement importantes, même dans des conditions normales. Par exemple, la Hsp90 représente 1-2% du total des protéines cytosolubles dans les cellules non stressées, ce qui en fait l'une des protéines les plus abondantes (Lai et al., 1984)

Le repliement des protéines médié par ces chaperons se croise avec d'autres domaines du métabolisme des protéines, tels que leur translocation dans les organites, et la régulation de leur conformation, de leur fonction, ainsi que de leur turnover intracellulaire (Albanese et al., 2006).

Chez les mammifères, les Hsp sont regroupées en six grandes familles en fonction de leur poids moléculaire : une de faible poids moléculaire (small Hsps ou sHsps), la famille Hsp47, Hsp60, Hsp70, Hsp90 et Hsp110 (Kampinga et al., 2009).

Bien que les Hsp partagent des propriétés communes, chaque famille présente des particularités concernant, par exemple, leur localisation intracellulaire, leur dépendance vis-à-vis de l'ATP, leur expression basale, ou leur induction par un stress.

2 La famille des chaperons Hsp90

La séquence de la protéine Hsp90 est fortement conservée entre les espèces ; la Hsp90 humaine a un pourcentage d'homologie de 60 % avec la Hsp90 de la levure boulangère, et de 40 % avec celle de *Escherichia coli*, la HtpG (Bardwell and Craig, 1987). La Hsp90 a été trouvée essentielle dans tous les eucaryotes qui ont été testés, comme, à titre d'exemple, la levure boulangère (Aligue, Akhavan-Niak and Russell, 1994), la drosophile (Yue et al., 1999) et la souris (Voss, Thomas and Gruss, 2000).

Cette famille est composée de cinq isoformes principaux : Hsp90 α , Hsp90 β , GP96, Trap1, et HspN.

2.1 Gp96, Trap1, et HspN

La Gp96 (Glycoprotein 96), aussi appelée Grp94, est présente chez les vertébrés et quelques organismes invertébrés, mais n'existe pas chez les bactéries, les levures ou la drosophile (Arrigo et al., 2005). Cette protéine de 96 kDa est abondante au sein du réticulum endoplasmique, grâce à une séquence de rétention C-terminale (Maki, Old and Srivastava, 1990). On la retrouve également exprimée à la surface cellulaire (Robert, Menoret and Cohen, 1999). Sa distribution est ubiquitaire, mais son expression peut être augmentée suite à un stress réticulaire (Bush, Goldberg and Nigam, 1997). Elle est également exprimée au sein de diverses tumeurs (Reddy et al., 2002), dans lesquelles elle présente un motif de glycosylation différent de celui observé dans des tissus sains (Suriano et al., 2005).

La Gp96 est un chaperon moléculaire impliqué dans la maturation des protéines (Yang and Li, 2005). Dernièrement, de nombreuses études tendent à lui attribuer un rôle dans le système immunitaire. En effet, la Gp96 semble participer à la présentation des antigènes, et est capable d'activer les réponses immunitaires innée et adaptative médiées notamment par les cellules présentatrices de l'antigène (Binder et al., 2007; Yang et al., 2007).

La Trap1 (Tumor Necrosis Factor Receptor-Associated Protein 1) est une protéine dont l'activité dépend de l'ATP, située dans la mitochondrie et dont la fonction reste méconnue. Elle semble avoir des fonctions antioxydante et antiapoptotique (Felts et al., 2000; Montesano Gesualdi et al.,

2007). Il a été montré qu'elle a un rôle important dans la régulation de la l'homéostasie mitochondriale des cellules tumorales (Kang et al., 2007). Néanmoins, il apparaît que la TRAP1 n'interagit pas directement avec des protéines clientes comme sa consœur Hsp90, ni avec les co-chaperons classiques de cette dernière (Felts et al., 2000).

La Hsp90N, est une variante Hsp90 principalement associée à la membrane, à cause du domaine hydrophobe unique de son extrémité N- terminale; la fonction cellulaire de cette variante reste inconnue (Grammatikakis et al., 2002).

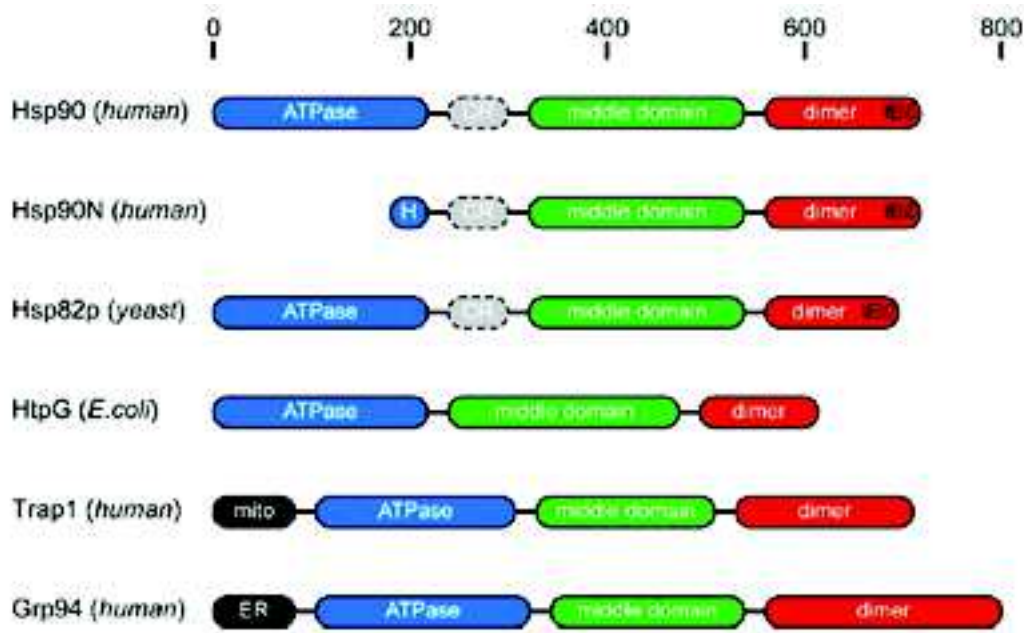


Figure 14 : Comparaison de la structure des domaines des membres de la famille Hsp90. Bleu -domaine ATPase (pour Hsp90N: région hydrophobe (H); gris foncé - séquence signal de localisation dans la mitochondrie (mito) et dans le réticulum endoplasmique (ER); gris clair - région chargée (CR), vert – domaine médian ; rouge - domaine de dimérisation, la présence du motif C-terminal MEEVD est également indiquée. (Wegele, Muller and Buchner, 2004).

2.2 Hsp90

Cette dénomination recouvre les deux isoformes Hsp90 α (Hsp86, Hsp90A) et Hsp90 β (Hsp84, Hsp90B), respectivement codées par les gènes Hsp90AA1 et Hsp90AB1 (Gupta, 1995). Ces deux protéines présentent 86 % d'homologie entre elles (Hickey et al., 1986).

Pendant longtemps, le terme Hsp90 a fait référence indistinctement aux deux isoformes (Chen et al., 2005). La Hsp90 α est thermo-inductible, alors que la Hsp90 β est constitutive. Les plantes semblent avoir plusieurs gènes des Hsp90 (Krishna and Gloor, 2001).

Les protéines Hsp90 sont situées principalement dans le cytoplasme. Cependant, elles ont également été détectées dans le noyau, avec une légère augmentation de leur accumulation nucléaire lors d'un choc hypoxique (Katschinski et al., 2002). Il y est également des preuves que la Hsp90 pourrait être présente à la surface cellulaire (Ferrarini et al., 1992). La sécrétion de la Hsp90 humaine dans la matrice extracellulaire environnant les cellules tumorales a été signalée, et elle semble contribuer au pouvoir invasif des cellules tumorales (Eustace et al., 2004).

La dimérisation de la Hsp90 est requise pour ses différentes fonctions (Minami et al., 1994). La Hsp90 est active principalement sous la forme d'homodimères α/α ou β/β , mais elle est détectée également sous forme hétérodimérique, voire monomérique (essentiellement la forme β).

2.2.1 Structure

Récemment, la Hsp90 de la levure boulangère a été cocrystallisée dans sa totalité avec un analogue non hydrolysable de l'ATP (AMP-PNP) et le co-chaperon p23/Sba1 (Ali et al., 2006).

Depuis, la cocrystallisation de la Hsp90 avec différents inhibiteurs a été réalisée, et a permis de mieux cerner la structure de la protéine.

La Hsp90 est bien un homodimère dont les principaux contacts inter- monomères sont assurés par les 190 résidus C-terminaux. Chaque monomère de Hsp90 est constitué de trois domaines : un domaine N- terminal hautement conservé dans les espèces, d'environ 25 kDa, qui est connecté à un domaine médian conservé de 35 kDa. Puis vient le domaine C-terminal, très conservé de 12 kDa nécessaire à la dimérisation (Meyer et al., 2003). Une séquence chargée relie le domaine N-terminal au domaine médian.

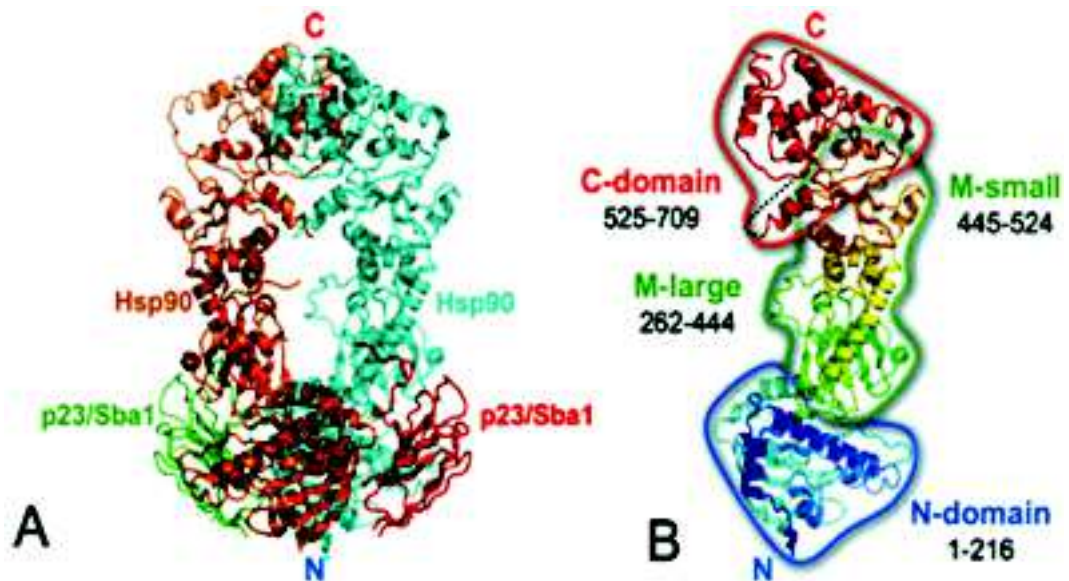


Figure 15 : (A) (Structure cristalline du complexe Hsp90-p23/Sba1 avec un analogue non hydrolysable d'ATP (p [NH] APP)).

Chacune des deux molécules Hsp90 (orange, bleu) s'enroule autour de l'autre, interagissant via une interface constitutive du dimère à l'extrémité C-terminale et une interface transitoire dépendant de l'ATP à l'extrémité N-terminale. P23/Sba1 le co-chaperon, qui interagit spécifiquement avec la conformation de Hsp90 liée à l'ATP, se lie au niveau d'une dépression à la jonction des deux domaines N-terminaux. (B) Schéma montrant l'architecture globale du monomère Hsp90. (Pearl, Prodromou and Workman, 2008)

2.2.1.1 Domaine N-terminal (ND)

Ce domaine est le plus étudié et le mieux connu. Les structures des domaines N-terminaux isolés des Hsp90 humaine et de levure (Prodromou et al., 1997; Prodromou et al., 2000) ont été déterminées par cristallographie, ainsi que celles des complexes du domaine N-terminal lié à des nucléotides (ADP) ou des inhibiteurs (geldanamycine, radicicol, NVP-AUY9, PU3 ...etc.). Le domaine N-terminal contient un site de liaison pour l'ATP (dont l'hydrolyse produit l'ADP, fournissant ainsi l'énergie nécessaire au fonctionnement de la Hsp90). Ce site ATPasique présente une homologie structurale avec celui des protéines appartenant à la superfamille d'ATPases connues sous l'acronyme de GHKL (ADN Gyrase, Hsp90, Histidine Kinase, MutL) (Prodromou et al., 1997). Les résidus impliqués dans la

liaison et l'hydrolyse de l'ATP forment une poche de liaison inhabituelle (différente des poches de liaison d'autres ATPases), connue sous le nom de « Bergerat fold » et sont très conservés dans cette superfamille (Bergerat et al., 1997).

Les études structurales montrent que l'ADP adopte une conformation pliée (folded) totalement inhabituelle au niveau du groupement phosphate β , et qui n'est pas observée dans d'autres ATPases, telles que les kinases ou la Hsp70 (Chene, 2002).

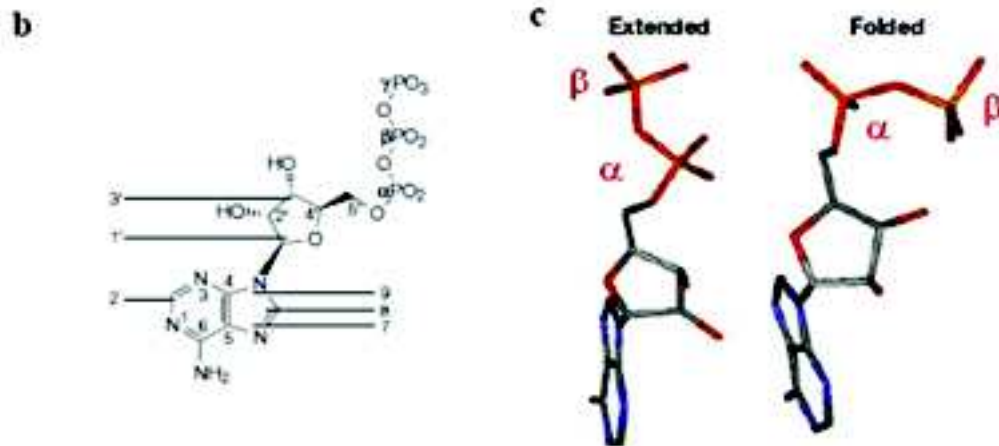


Figure 16 : b. structure de l'ATP et numérotation chimique ; c. conformations étendue (extended) ou pliée (folded) du groupement phosphate β de l'ADP (Chene, 2002).

2.2.1.2 Domaine médian (MD)

Chez les eucaryotes, une flexible séquence de liaison hautement chargée relie le domaine N-terminal au domaine médian de la Hsp90 ; cette séquence de liaison est dispensable pour la viabilité des levures (Louvion, Warth and Picard, 1996).

La structure du domaine médian indique qu'il a un rôle important dans la modulation de l'hydrolyse de l'ATP par l'interaction avec le phosphate λ des molécules d'ATP qui sont liées à la poche N-terminale (Ali et al., 2006; Meyer et al., 2003; Soti, Racz and Csermely, 2002).

Une analogie avec la gyrase B ainsi que des études de mutagenèse (Meyer et al., 2003; Pearl and Prodromou, 2006; Sato, Fujita and Tsuruo, 2000) indiquent que le segment médian est un site essentiel pour les interactions avec les protéines-clients via des résidus hydrophobes.

La séquence de liaison chargée joue, d'une part un rôle dans la fixation de l'ATP au niveau du site N-terminal, et d'autre part augmente l'affinité du site N-terminal pour l'ATP en conférant à ce site une conformation particulière. En effet, après fixation de l'ATP au niveau du site N-terminal, la région chargée modifie la conformation de ce site et empêche ainsi l'approche d'une seconde molécule d'ATP. Cette région chargée joue donc un rôle essentiel dans la régulation de la fonction chaperon de la Hsp90.

2.2.1.3 Domaine C-terminal (CD)

Une autre flexible séquence de liaison relie le domaine médian au domaine C-terminal ; la structure de ce dernier est inconnue pour la Hsp90 humaine. En revanche, celle de son homologue HtpG d'E. Coli a été déterminée (Harris, Shiao and Agard, 2004). Il comporte le domaine de dimérisation des monomères de Hsp90 (Minami et al., 1994). De par leur

homologie structurelle et fonctionnelle, le domaine C-terminal de la Hsp90 humaine est supposé être similaire de celui de la HtpG d'E. Coli.

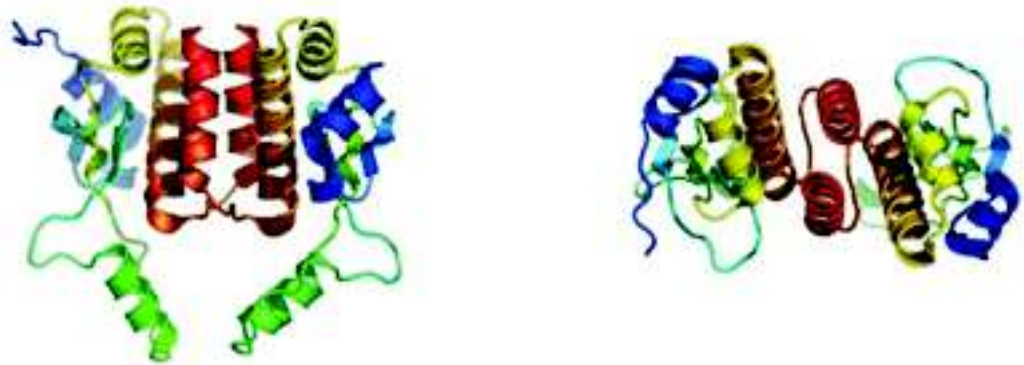


Figure 17 : Structure du domaine C-terminal de l'homologue de la Hsp90 d'E. coli (HtpG) sous deux vues différentes (Pearl and Prodromou, 2006).

La suppression de ce domaine C-terminal porte atteinte de façon drastique à l'activité ATPasique de la Hsp90, ce qui montre clairement le rôle de très coopératives interactions intermoléculaires et intramoléculaires dans la régulation de l'hydrolyse de l'ATP par ce chaperon.

L'existence d'un second site de fixation de l'ATP au niveau du domaine C-terminal de la Hsp90 est hypothétique et très controversée, mais est avancée par plusieurs groupes (Marcu et al., 2000; Allan et al., 2006). Cependant, du fait qu'aucune structure cristalline avec des inhibiteurs de ce site ATPasique n'ait été publiée, la connaissance de ce site reste limitée.

Enfin, le domaine C-terminal de la Hsp90 porte un motif conservé MEEVD qui est responsable du recrutement des différents co-chaperons contenant des domaines TPR (Tetratricopeptide repeat), tels que les immunophilines, HOP (Hsp70/Hsp90 Organizing Protein), et la protéine hosphatase 5 (Pearl and Prodromou, 2006; Pearl, Prodromou and Workman, 2008). Le motif MEEVD est absent dans la Trap1 et la Gpr94.

2.2.2 Activité ATPasique de la Hsp90

L'activité ATPasique de la Hsp90 est un processus essentiel à la maturation des protéines-clientes, et de nombreux changements de conformation y sont associés (Wandinger, Richter and Buchner, 2008).

La fixation de l'ATP au site N-terminal entraîne le passage (ou "switch") d'une conformation ouverte (open) à une fermée (closed) connue sous le nom de pince moléculaire "clamp". Hessling et al. ont suggéré une succession de cinq états différents qui précèdent l'hydrolyse de l'ATP (Hessling, Richter and Buchner, 2009).

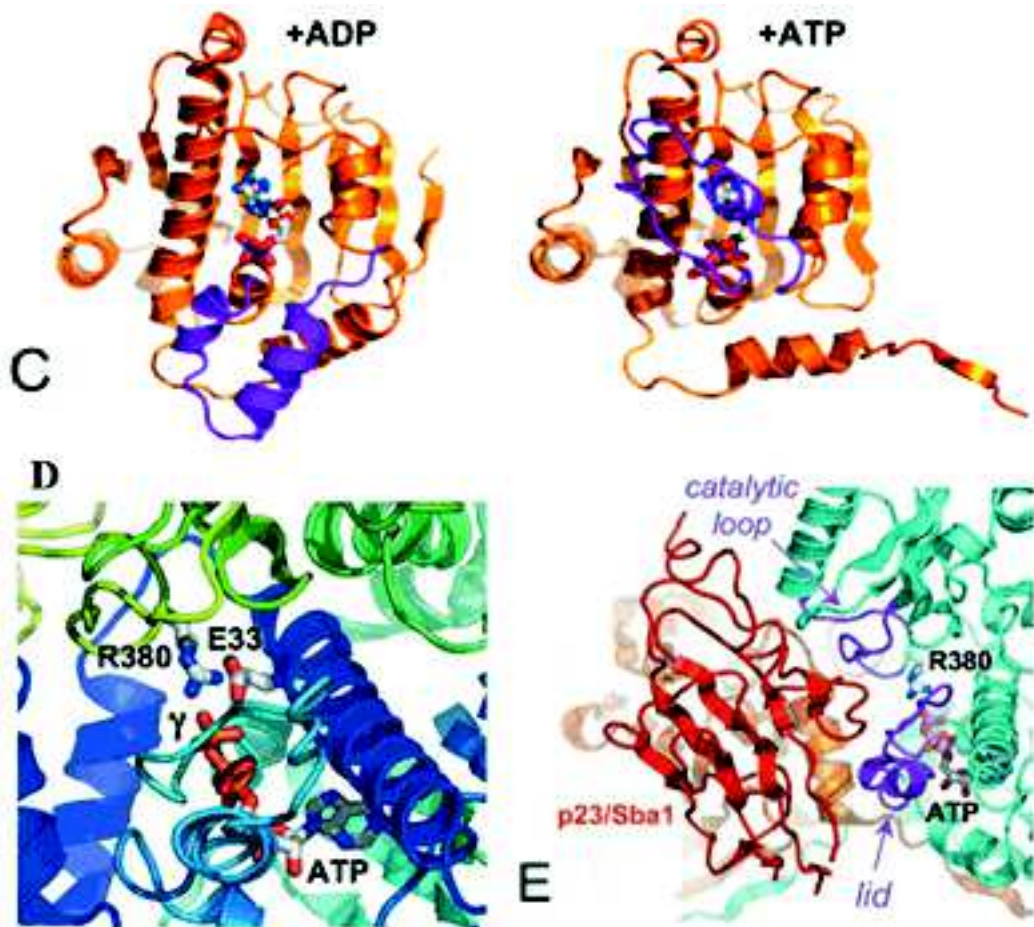


Figure 18 : (C) conformation du couvercle «lid» (magenta) dans l'apostructure (liée à l'ADP) (à gauche) et dans la structure dépendante de l'ATP du domaine N-terminal.

(D) Close-up de l'appareil catalytique ATPasique de la Hsp90. Glu33 à partir du domaine N-terminal polarise une molécule d'eau pour l'attaque nucléophile sur la liaison β - γ phosphodiester de l'ATP, qui est polarisée par la chaîne latérale basique de l'Arg380 du domaine médian du chaperon. (E) p23/Sba1 interagit avec les deux principaux «switches» qui régissent l'activité ATPasique de Hsp90, la boucle catalytique et le couvercle (Pearl, Prodromou and Workman, 2008).

Dans ce modèle, en absence d'ATP, le dimère de la Hsp90 est à l'état relaxé (open, étape 1). L'extrémité N-terminale forme une sorte de couvercle (lid), au niveau du domaine liant le nucléotide ; ouvert, le site de liaison de l'ATP est accessible. Les extrémités N-terminales ne sont pas alors contraintes. Puis, la liaison de molécules d'ATP (étape 2) entraîne la fermeture de ce couvercle (étape 3), suivie par le rapprochement des deux sites N-terminaux et leur dimérisation (étape 4). L'interaction des domaines N-terminaux avec les domaines médians conduit à la conformation complètement fermée (closed, étape 5), qui correspond à la configuration permettant l'activité ATPasique. En effet, cette conformation contrainte permet au résidu Arg380 (numérotation correspondant à la levure, Arg400 pour l'Homme), situé dans le domaine médian, de s'insérer dans le site catalytique du domaine N-terminal et d'interagir avec le phosphate γ de l'ATP, afin de permettre son hydrolyse. Enfin, les produits (ADP + Pi) sont relâchés, et le chaperon revient à l'état ouvert (open).

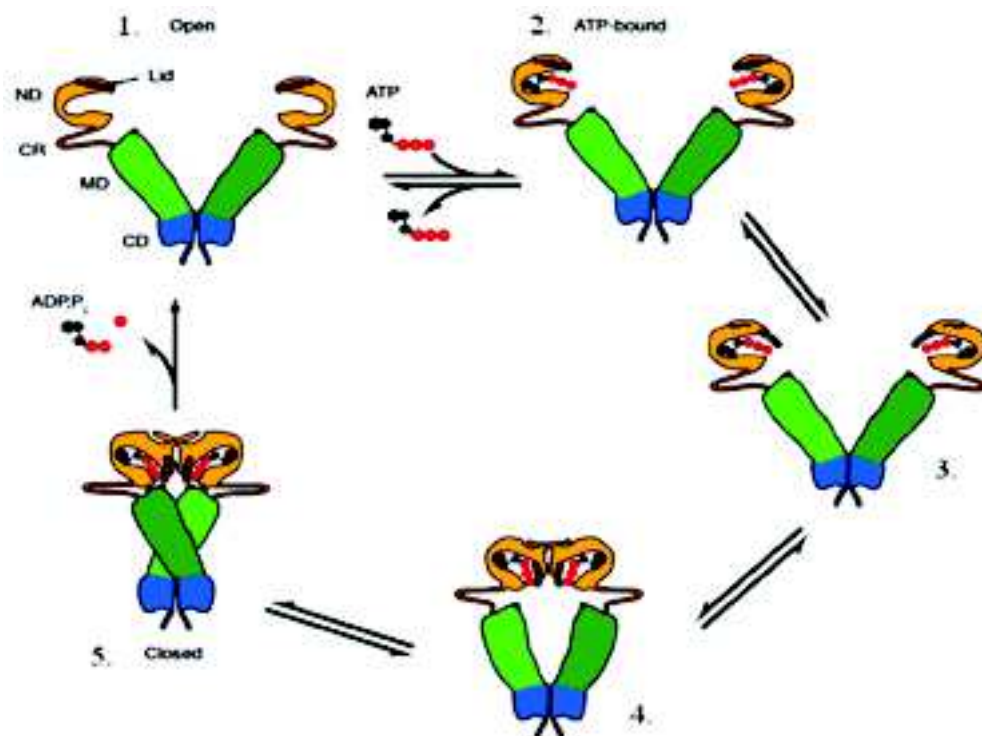


Figure 19 : Hsp90 : modèle du cycle ATPasique (Hessling, Richter and Buchner, 2009).

Les protéines-clients et les co-chaperons ne sont pas représentés. Légende : Lid (couvercle, en rouge), ND (domaine N-terminal, en jaune), CR (région/séquence chargée, en orange), MD (domaine médian, en vert), CD (domaine C-terminal, en bleu).

Ce mécanisme suggère que la fixation de l'ATP à la Hsp90 induit tout d'abord un changement conformationnel qui permet au chaperon de former une liaison stable avec une protéine-cliente. Puis, l'hydrolyse de l'ATP permet la libération de la protéine transformée lors de la phase finale du processus.

L'ATP joue donc un rôle primordial pour le fonctionnement de la Hsp90, car il permet le recrutement des co-chaperons, et provoque la liaison de la Hsp90 aux protéines-clients. Ceci entraîne, après les modifications structurales assurées par le complexe macromoléculaire, la stabilité conformationnelle des clientes qui leur permettra d'exercer leur fonction biologique.

La nature des co-chaperons recrutés dépend de celle de la protéine-cliente prise en charge par la Hsp90. Tout d'abord, la protéine-cliente (substrat) s'associe à la Hsp70, à la Hsp40 et au co-chaperon Hip (Hsp70 interacting protein). Ce complexe interagit avec la Hsp90 (qui n'est pas liée à l'ATP) grâce à la protéine Hop (Hsp70/Hsp90 organizing protein), bloquant temporairement l'accès de l'ATP à son site de fixation, pour former un complexe intermédiaire.

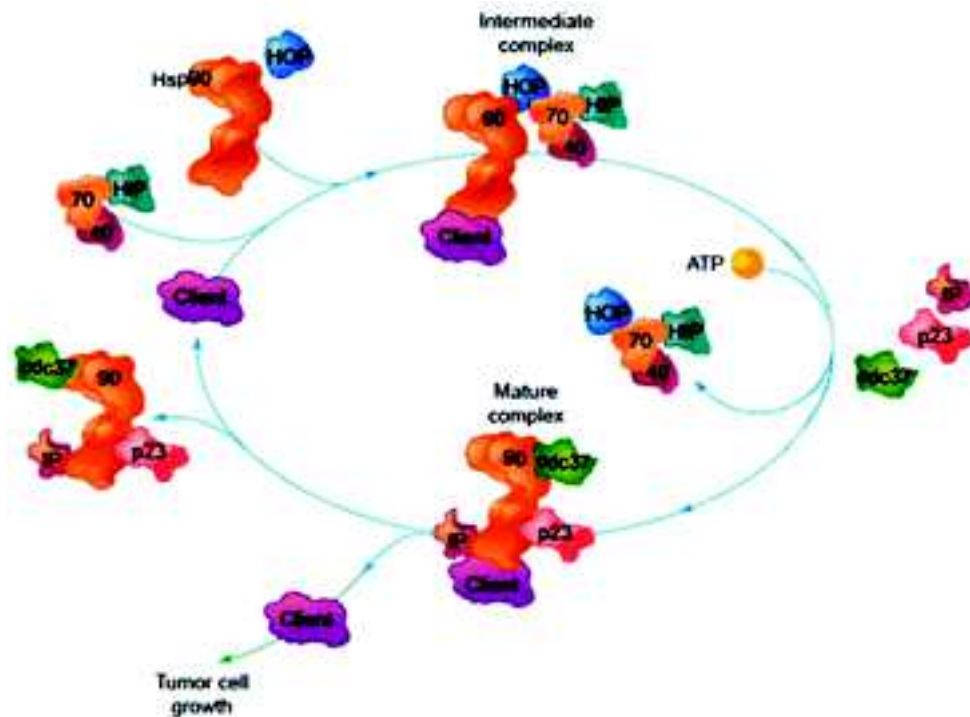


Figure 20 : cycle ATPasique hypothétique de la Hsp90 (Kamal, Boehm and Burrows, 2004). HIP (Hsp70 Interacting Protein), HOP (Hsp70/Hsp90 Organizing Protein), p23 (co-chaperon), cdc37 (co-chaperon), IP (immunophiline).

La fixation de l'ATP entraîne un changement de conformation de la Hsp90, le déplacement des premiers co-chaperons (Hsp40, Hsp70, Hip et Hop) et la formation d'un complexe mature, après le recrutement d'autres co-chaperons, notamment p23 (co-chaperon essentiel à l'activité de la Hsp90), cdc37 et les immunophilines (IP). Ce complexe macromoléculaire catalyse la maturation conformationnelle de la protéine cliente, qui, après hydrolyse de l'ATP, est libérée dans le cytoplasme (Prodromou, 2012; Wandinger, Richter and Buchner, 2008).

Bien qu'au cours des dernières années de nombreux progrès aient été réalisés dans la compréhension du mécanisme d'action de la Hsp90 et de son cycle catalytique, plusieurs aspects ne sont toujours pas élucidés et des contradictions persistent.

2.2.3 Activité de chaperon, protéines-clientes, et co-chaperons

En conditions physiologiques, la Hsp90 peut lier de nombreuses protéines, et former avec elles des complexes stables. Ces partenaires d'interaction de la Hsp90 sont dénommés « protéines-clientes », on en dénombre à l'heure actuelle plus de 200, identifiées chez les mammifères (pour une liste mise à jour, voir [http:// www.picard.ch/downloads/ Hsp90facts.pdf](http://www.picard.ch/downloads/Hsp90facts.pdf)). Certaines de ces protéines sont des acteurs essentiels de la vie cellulaire, ce qui confère à la Hsp90 un rôle fondamental dans la maintenance protéique. En formant des complexes avec elles, la Hsp90 leur donne la conformation tridimensionnelle fonctionnelle, prévient leur agrégation et contrôle aussi leur activité par séquestration. Néanmoins, la Hsp90 semble plus particulièrement impliquée dans le chaperonnage de deux grandes classes de protéines (Prodromou, 2012) :

- Les facteurs de transcription. Elle intervient dans la maturation et la régulation de l'activité de plusieurs facteurs de transcription, notamment les récepteurs aux hormones stéroïdiennes.

Les protéines kinases. Elle participe à la stabilisation et à l'activation des protéines kinases, telles que Akt/PKB (Protéine Kinase B), v-Src kinase ou Raf Kinase.

La Hsp90 est également impliquée dans la stabilisation du cytosquelette en s'associant à la fois à l'actine et la tubuline. L'interaction avec les microfilaments permet de préserver l'intégrité cellulaire après un stress et le transport intracytoplasmique des protéines (Czar, Welsh and Pratt, 1996; Pratt, 1992).

Co-Chaperones		Function	
TTR co-chaperones			
Mammals	Gene name (human)	Yeast	Fruit
Hsp	STB1	S81	Hsp
Hsp65	FKBP4	None	AT5G48520 ^a
Rhsp51	FKBP5	None	ROTI
Cyp40	PRD	Cyp6/Cy7	SQN
Alp	Alp	None	None
CHP	STB1	None	CHP
Pp5	PP5C	Pp1	Pp5.2
Tp2	DNAK7	None	ATP581
Sg1	SG1T1	Sg1	PK ^a
Ubc45	UNC45B	S84	None
Tvc4	TTC4	Con1	ATT04130 ^a
Tom70	TOMM70A	Tom70p	None
None	None	None	To64
Tab1	RPM1P3/RLJ21908	Tab1	ATT026440 ^a
Non-TTR co-chaperones			
Abi1	AHS1A1	Abi1	AT3G12050 ^a
p23	PTGE53	S841	AT3K013773 ^a
GAG7	GDC37	Cd637	None
Chp1/Melulin	CHOBCT1	None	Rat1
NudC	MEDC	NudC	AT4G27890 ^a

^a Several homologues are underlined in plants. The listed gene names are for *Arabidopsis thaliana*.

Table 2 : résumé des principaux co-chaperons de la Hsp90

L'activité de la Hsp90 dépend de la présence d'autres protéines ayant un rôle de co-chaperons. Le domaine C-terminal comporte un motif MEEVD, qui est reconnu par le domaine TPR (Tetratricopeptide repeat) présent dans la séquence de ces co-chaperons, ce qui permet leur association avec la Hsp90 (Pearl and Prodromou, 2006). Dans les cellules eucaryotes, plus de 20 co-chaperons ont été identifiés ; Ces co-chaperons régulent la

fonction de Hsp90 par différentes voies, telles que l'inhibition et l'activation de son activité ATPasique, et le recrutement de protéines- clientes spécifiques. (Li et al., 2012). Ainsi, le co-chaperon Hop (HSP70/HSP90 organizing protein) permet le recrutement de la Hsp70 et la Hsp40 et l'arrimage de la protéine-cliente.

2.2.4 Inhibition de la Hsp90 et conséquences

Le cancer est une maladie associée à une instabilité génétique qui permet aux cellules cancéreuses d'acquérir six caractéristiques distinctives (marques du cancer ou hallmarks en anglais), y compris : autosuffisance en signaux de croissance, résistance à l'apoptose, insensibilité aux signaux antiprolifératifs, angiogenèse soutenue, invasion des tissus et métastases, et potentiel proliférative sans limites. De nombreuses protéines de signalisation associées à ces caractéristiques du cancer sont des protéines-clientes de la Hsp90.

Hallmark of cancer	Hsp90 client protein
Evasion of apoptosis	Akt,Rip,P53,Survivin, Apaf-1,Bcl-2,IGF-1R
Sustained angiogenesis	VEGFR, HIF1,Akt, Fit-3,FAK,Src
Limitless replicative potential	n-TERT, telomerase
Tissue invasion and metastasis	c-MET,MMP2
Self-sufficiency in growth signals	EGFR/Her-2,Raf, Bcr-Abl,ErbB-2, Src,Akt,MEK
Insensitivity to anti-growth signals	Plk-1,Cdk4,Cdk6, Myt-1,cyclin D

Table 3 : relation entre protéines clientes de Hsp90 et les marques du cancer.

En outre, la croissance rapide des cellules cancéreuses nécessite la participation de grandes quantités cytosoliques de Hsp90. Ainsi, il est utile, pour inhiber la croissance et la métastase du cancer, de supprimer la formation des complexes protéiques de Hsp90 avec ses clientes, qui non seulement réduit la tolérance des cellules cancéreuses aux chimiothérapies, mais aussi accroît, par synergie, l'efficacité de plusieurs médicaments anticancéreux (Li et al., 2012).

Donc, l'inhibition de la Hsp90 permettrait de cibler simultanément les différentes marques du cancer, en bloquant le fonctionnement de nombreuses protéines oncogènes. Un blocage simultané de plusieurs points de régulation permettrait d'induire l'arrêt du cycle cellulaire et conduirait à l'apoptose dans les cellules cancéreuses, via une dégradation des protéines-clientes par le protéasome (Ferrarini et al., 1992).

Aujourd'hui, si les mécanismes de l'inhibition de la Hsp90 sont mieux compris, les conséquences de cette inhibition et le mécanisme de l'activité antitumorale des inhibiteurs sont très complexes et ne restent que partiellement élucidés.

Les études réalisées sur des modèles animaux de xénogreffes de tumeur humaine, ont montré que les inhibiteurs de la Hsp90 s'accumulaient dans les cellules tumorales de manière très sélective (Chiosis and Neckers, 2006).

C'est le cas de la 17-AAG dont la distribution est plus importante dans les cellules cancéreuses que dans les cellules saines. Kamal et al. (Kamal et al., 2003) rapportent

que la Hsp90 issue de cellules tumorales présente une affinité pour des inhibiteurs du site N-terminal (comme 17-AAG) environ 100 fois supérieure à celle de la Hsp90 des cellules normales. Ils montrent que la Hsp90 des cellules cancéreuses se trouve sous forme de complexes multichaperons, c'est-à-dire liée à des protéines-clientes et des protéines partenaires, tandis que la Hsp90 issue de cellules saines se trouve sous sa forme libre homodimérique. De ce fait, le complexe multiprotéique d'origine tumorale présente une activité ATPasique bien plus importante et une affinité également accrue pour les inhibiteurs. Ainsi, l'activité ATPasique des complexes multiprotéiques des cellules cancéreuses est inhibée par la 17-AAG à des concentrations largement inférieures à celles nécessaires à l'inhibition de la Hsp90 des cellules saines.

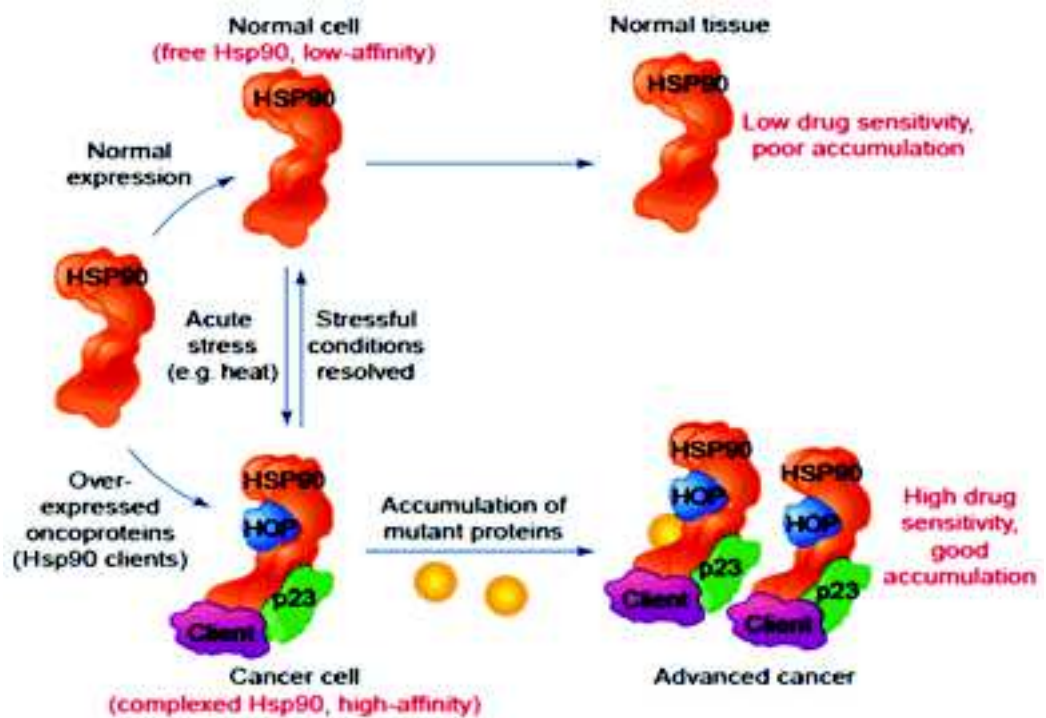


Figure 21 : modèle de la sélectivité des inhibiteurs de Hsp90 dans les tumeurs (Kamal, Boehm and Burrows, 2004)

Dans une cellule tumorale, les complexes multichaperons présenteraient donc la Hsp90 dans une conformation hautement affine, avec une activité ATPasique augmentée, ce qui faciliterait la progression maligne en permettant la maturation des protéines oncogènes plus efficacement. En effet, puisque les protéines-clientes oncogènes de la Hsp90 sont surexprimées dans les cellules cancéreuses, ces dernières sont très dépendantes de cette machinerie superchaperon pour continuer leur croissance. Dans les tissus normaux, la Hsp90 resterait dans un état homodimérique latent, non complexé. Ces observations expliqueraient la sélectivité tumorale observée pour la 17-AAG.

Ces données suggèrent que la Hsp90 des cellules cancéreuses peut être ciblée de manière sélective. Cependant, l'index thérapeutique d'un inhibiteur ne pourra être évalué qu'en stade clinique.

Il existe plusieurs mécanismes d'inhibition de l'activité chaperon de Hsp90, mais le mécanisme principal est l'inhibition du site de liaison de l'ATP du domaine N-terminal. Presque tous les inhibiteurs qui sont entrés en essai clinique et préclinique aujourd'hui,

sont de ce type; cela semble logique, car le premier inhibiteur identifié de Hsp90, la geldanamycine, en fait partie, ce qui y dirigea la majorité des études et des recherches.

2.2.4.1 Les inhibiteurs de du site d'ATP N-terminal de Hsp90

Aujourd'hui, Il existe plusieurs inhibiteurs naturels et synthétiques du site ATP du domaine N-terminal de Hsp90, avec différents squelettes moléculaires et appartenant à différentes familles chimiques.

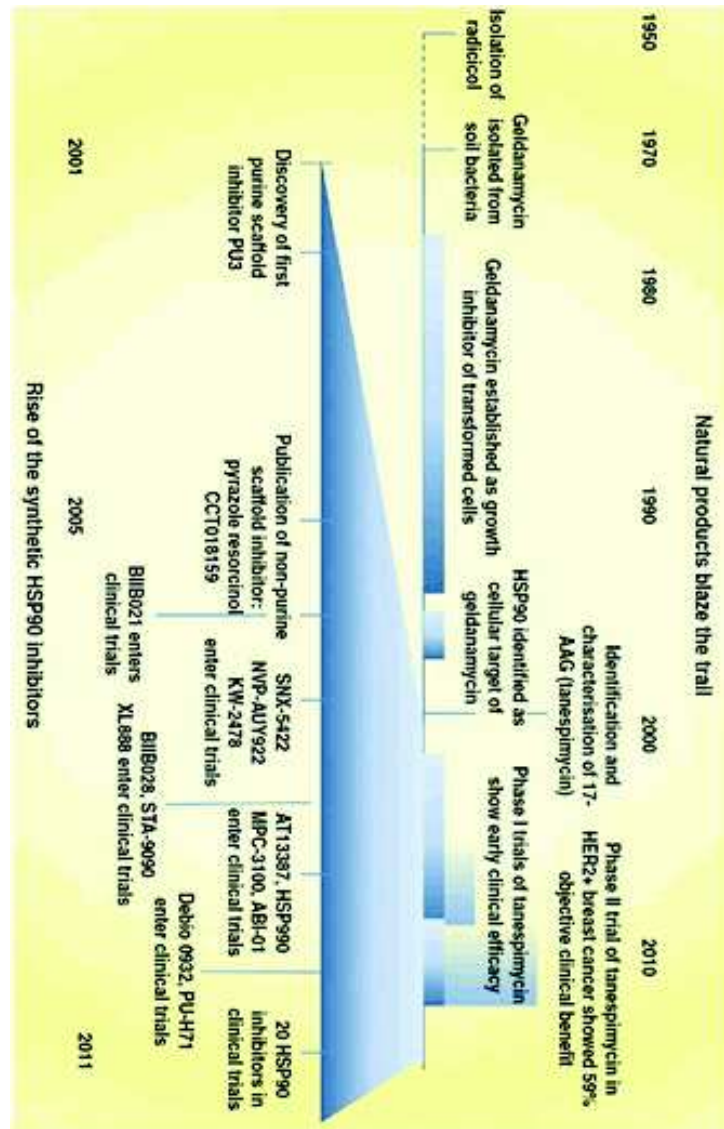


Figure 22 : chronologie de la découverte et du développement des inhibiteurs de Hsp90 (Travers, Sharp and Workman, 2012)

Actuellement, il y a 20 inhibiteurs qui sont en essai clinique sur différents types de cancers, et beaucoup plus d'autres composés qui sont en développement préclinique. Cela représente une augmentation considérable de la production de la recherche en la matière par rapport à ce qu'elle était, il y a plus de dix ans, lorsque le premier dérivé de la geldanamycine est entré en développement préclinique (Travers, Sharp and Workman, 2012).

a. Les ansamycines benzoquinones

Le premier inhibiteur spécifique de la Hsp90 identifié fut la geldanamycine (GA), un antibiotique isolé de *Streptomyces hygroscopicus*, dans les années 1970 (Stebbins et al., 1997).

La GA appartient à la classe des ansamycines benzoquinones qui se lient spécifiquement au site N-terminal de la Hsp90 (Maloney and Workman,

2002). Ce macrolactame possède une puissante activité antitumorale in vitro (Schulte and Neckers, 1998) et dans des modèles animaux in vivo, mais son utilisation en phase clinique a été interrompue en raison d'une forte hépatotoxicité indépendante de l'inhibition de la Hsp90 (Supko et al., 1995). Le 17-allylamino-17 déméthoxygeldanamycine, (17-AAG, KOS-953) ou tanespimycine, un dérivé hémisynthétique de la GA (Schulte and Neckers, 1998), L'alvespimycine, la 17-[(diméthylamino)éthylamino]-17 déméthoxygeldanamycine (17-DMAG, KOS-1022), un analogue aminé substitué en position 17 de la GA (Burger et al., 2004), La retaspimycine (IPI-504), le sel chlorhydrate de la forme réduite de la 17-AAG (Sydor et al., 2006), représentent la deuxième génération des inhibiteurs de la classe des ansamycines.

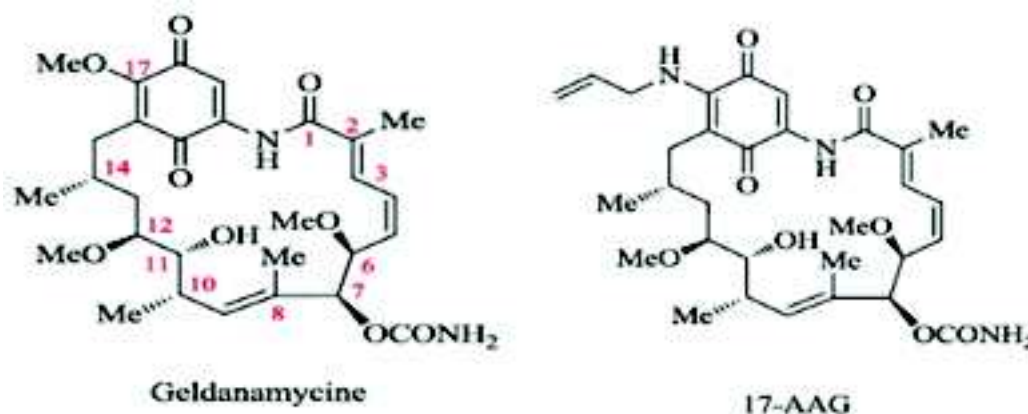


Figure 23 : structures de quelques ansamycines.

À cause de problèmes de toxicité, de solubilité, et de manque d'efficacité, tous les essais cliniques concernant les analogues de la GA ont été suspendus ou abandonnés (<http://clinicaltrials.gov/>).

b. Le radicicol et autres macrolactones résorcyliques

Le radicicol (Rd) est un antibiotique de la famille des macrolactones résorcyliques, isolé de *Monosporium bonorden* en 1953 (Delmotte and Delmotte-Plaqué, 1953).

Cette macrolactone inhibe la Hsp90 en se liant spécifiquement à son site ATPasique N-terminal ; son affinité pour la Hsp90 est cent fois supérieure à celle de la GA ou la 17-AAG. Cependant, bien que très actif in vitro, le radicicol ne montre aucune activité antitumorale in vivo (Schulte et al., 1998).



Figure 24 : structures de quelques macrolactones résorcycliques.

La pochonine D, macrolactone naturelle analogue du Rd isolée de *Pochonia chlamydosporia*, a montré une affinité pour la Hsp90 proche de celle du Rd (Moulin et al., 2005).

Plusieurs analogues du radicol et de la pochonine D ont été développés, mais la pro-pochoxime B semble la plus prometteuse (Barluenga et al., 2009).

c. Les purines

La première molécule décrite de cette classe, PU3, mime les effets cellulaires des inhibiteurs de la famille des ansamycines, mais avec un potentiel plus faible. Bien que moins puissant que la 17-AAG, PU3 est plus hydrosoluble et présente une meilleure biodisponibilité. PU3 représente le premier inhibiteur synthétique de la Hsp90 (Chiosis et al., 2001; Vilenchik et al., 2004).

BIIB021 (ou CNF2024), est une 2-amino-6-chloropurine, avec une puissante activité inhibitrice de la Hsp90 in vitro, et une activité antitumorale in vivo similaire à celle de la 17-AAG. BIIB021, est entrée en évaluation clinique en 2005 (Kasibhatla, Hong and Biamonte, 2007).

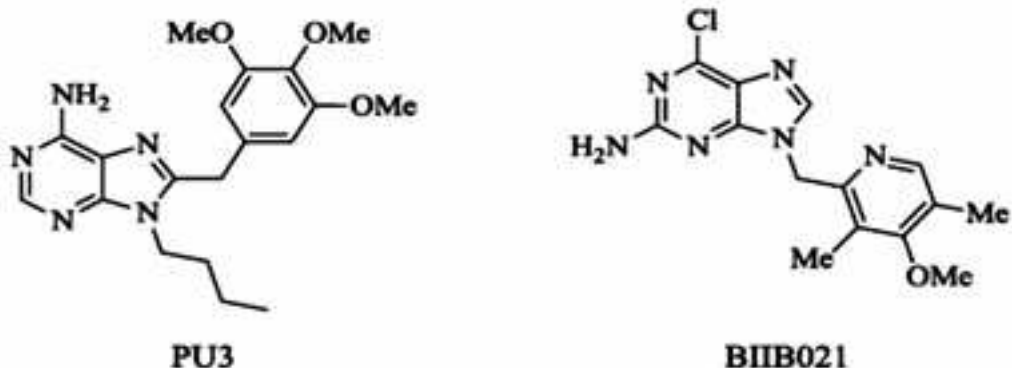


Figure 25 : structures de quelques purines.

d. Les pyrazoles et isoxazoles

Le 3,4-diarylpirazole (CCT018159) est une molécule qui présente un motif résorcinol, également présent dans le radicicol, inhibiteur naturel de la Hsp90 (Cheung et al., 2005).

Une optimisation a conduit au composé VER49009, possédant une bonne activité inhibitrice ATPasique (CI50 de 140 nM) (Dymock et al., 2005). Le remplacement du pyrazole par une isoxazole donnant des liaisons hydrogène plus stables conduit au NVPAUY922 (ou VER52296), qui est entré en essai clinique en 2007 (Brough et al., 2007).



Figure 26 : structures de quelques pyrazoles et isoxazoles

e. Les phényltétrahydroindolones et analogues

Le benzamide A possède une CI50 voisine de 50 μ M (sur une lignée cellulaire non spécifiée). Son originalité relève de son nouveau squelette chimique, par rapport aux précédents inhibiteurs de la Hsp90 (Huang et al., 2009).



Figure 27 : structures de quelques phényltétrahydroindolones et analogues.

Des modifications structurales ont mené à l'analogue tétrahydroindazolone SNX2112 et à sa prodrogue glycinate, le SNX5422. Ces derniers sont actifs à des CI50 de l'ordre du nanomolaire sur plusieurs lignées cellulaires. Les deux molécules sont entrées en évaluation clinique en 2007 (Biamonte et al., 2010).

Chapitre III : Screening Virtuel et Docking Moléculaire

1 Introduction

L'identification et la conception de molécules d'intérêt thérapeutique a bénéficié ces dernières décennies des progrès issus de diverses disciplines scientifiques telles que la biologie, la pharmacochimie et l'informatique.

Ainsi la recherche, qui consistait autrefois à synthétiser et tester les composés sélectionnés sur la base de l'intuition et de l'expérience du chimiste médicamenteux, a radicalement évolué.

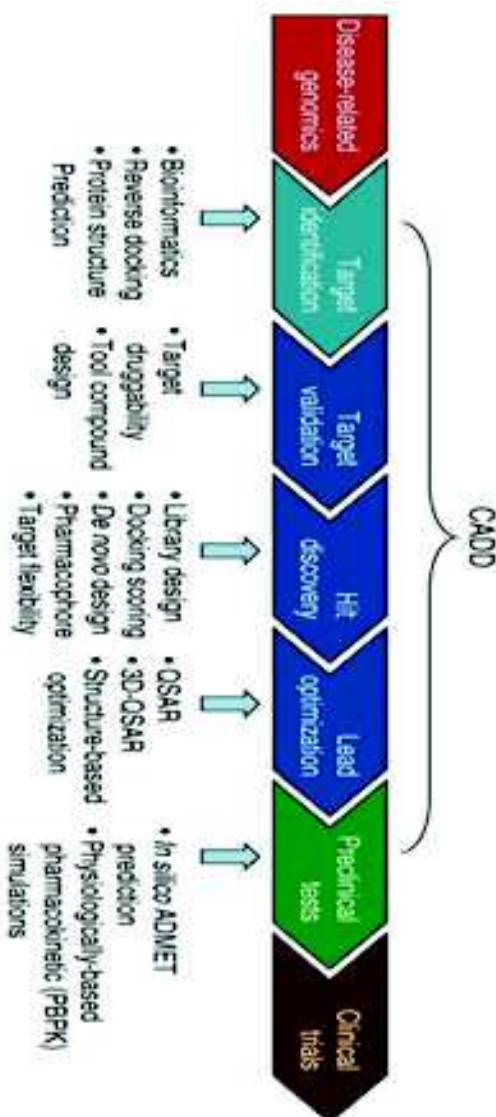


Figure 28 : Pipeline de développement d'un médicament sur lequel sont indiqués les endroits où les méthodes in silico peuvent intervenir (CADD : Computer-Aided Drug Design). (Tang et al., 2006)

L'essor de l'outil informatique a particulièrement changé la donnée, en conduisant à l'émergence d'une nouvelle discipline pouvant participer aux étapes initiales de la recherche pharmaceutique en complément des méthodes expérimentales déjà reconnues. On parle alors de conception de médicaments in silico ou assistée par ordinateur, qui correspond à un ensemble de techniques informatiques spécifiques souvent désigné par l'acronyme CADD ("Computer-Aided Drug Design") (Jorgensen, 2004; Tang et al., 2006; Talele, Khedkar and Rigby, 2010).

Bien que ces outils aient un large champ d'application dans le processus de recherche de nouveaux médicaments, nous nous limiterons à la description des méthodes utilisées pour ce travail, à savoir le criblage virtuel (Walters, Stahl and Murcko, 1998; Shoichet, 2004) et, plus particulièrement, une de ses sous-composantes : l'arrimage moléculaire, ou "docking" (Kitchen et al., 2004).

2 La stratégie du criblage virtuel

Le criblage virtuel (VS, virtual screening) est la stratégie *in silico* la plus utilisée pour l'identification de touches ("hits") dans le cadre de la recherche de nouveaux médicaments (Shoichet, 2004; Oprea and Matter, 2004). Celui-ci fait désormais partie intégrante de la plupart des programmes de recherche de composés bioactifs, que ceux-ci se déroulent en milieu académique ou industriel, car il constitue un complément essentiel au criblage biologique haut-débit.

Le criblage virtuel permet l'exploration de larges chimiothèques (> 10⁶ molécules) à la recherche de composés actifs vis-à-vis d'une cible thérapeutique donnée. Ce processus vise à réduire de façon significative la chimiothèque de départ à une liste limitée de composés jugés les plus prometteurs.

Cette approche conduit souvent à une nette amélioration de la "concentration" de molécules actives pour la cible ("hit rate"), tandis qu'une sélection aléatoire de molécules de la chimiothèque ne saurait fournir un tel enrichissement. Ainsi, le temps aussi bien que les coûts de l'identification de nouvelles touches peuvent être réduits de façon significative (Stahura and Bajorath, 2004; Ripphausen et al., 2010).

1.1 Préparation d'une chimiothèque

Ces bases de données peuvent être seulement des collections de structures (2D, 3D et autres formats) ou fournir des données supplémentaires, telles que la bio-activité mesurée des composés et leurs cibles protéiques, ainsi que les maladies ciblées. Mais très peu de ces bases de données (comme ChEMBL) tentent de lier les données sur les petites molécules à des informations sur leurs caractéristiques biologiques, leurs cibles, ainsi que leurs données d'essai disponibles.

Database	Publisher	License type
Open National Cancer Institute Database	National Cancer Institute	Publicly available
PubChem	National Center for Biotechnology Information	Publicly available
BindingDB	University of Maryland, USA	Publicly available
Rebase	Cambridge Crystallographic Data Centre	Freely accessible for academia commercial version available
CHEMBLdb	European Bioinformatics Institute, Hinxton, UK	Publicly available
ChEMSpider	Royal Society of Chemistry, UK	Publicly available
Human Metabolome Database	University of Alberta, Canada	Publicly available
DrugBank	University of Alberta, Canada	Publicly available
Therapeutic Target Database	National University of Singapore, Singapore	Publicly available
ZINC	University of California, San Francisco, USA	Publicly available
Research Library	ChemNavigator	Commercial
GVKBIO databases	GVK Biosciences Private Limited, India	Commercial
MDDR	Accelrys Inc.	Commercial
Wombat	Sunset Molecular Discovery	Commercial
World Drug Index	Thomson Reuters	Commercial

Table 4 : bases de données de petites molécules les plus connues et les plus pertinentes (Liao et al., 2011).

La pertinence de la chimiothèque employée est la première condition pour le succès d'un criblage virtuel, bien avant celle des algorithmes utilisés pour la recherche de touches au sein de la chimiothèque (Bleicher et al., 2003). En effet, seule une librairie de composés suffisamment diverse peut garantir une exploration satisfaisante de l'espace chimique, maximisant ainsi les chances de découvrir de nouvelles touches (Perez, 2005).

1.2 Préfiltrage de la chimiothèque ADME/T

Pour éviter de perdre du temps avec des molécules possédant des caractéristiques incompatibles avec celles de composés d'intérêt pharmaceutique, le processus de criblage comporte généralement une étape préliminaire de filtrage. Plusieurs critères de base sont pris en compte, tels que le faible cout de production, la solubilité, la stabilité, mais aussi, les

propriétés pharmacologiques d'absorption, de distribution, de métabolisme, d'excrétion et de toxicité (ADME/T) (Gleeson, Hersey and Hannongbua, 2011).

Cette tâche, qui peut être prise en charge par des programmes spécialisés, consiste à exclure les composés toxiques ou supposés tels ou comportant des groupements jugés trop réactifs. Ensuite, ne sont retenus que les composés obéissant à des définitions empiriques simples du profil de molécule active (caractère dit "drug-like"), telle que la populaire "règle des 5" de Lipinski (Lipinski et al., 2001) et ses extensions.

Durant leur phase d'optimisation, la complexité, le poids moléculaire et le caractère lipophile des touches ont généralement tendance à augmenter. Par conséquent, des filtres physico-chimiques additionnels peuvent être employés pour restreindre la chimiothèque aux composés ayant un profil de tête de série, dits "lead-like" (Hann and Oprea, 2004).

On estime alors que les risques d'échec aux tests cliniques sont minimisés, ce qui est souvent un impératif économique. Toutefois, appliquées de façon stricte, ces règles de préfiltrage peuvent assez facilement conduire à l'exclusion de molécules d'intérêt. Par exemple, de nombreux médicaments commercialisés ne respectent pas les règles de Lipinski (Oprea, 2000) ; c'est pourquoi ces règles empiriques sont souvent employées de façon plus ou moins permissive.

Program	Developed by	Free for academia	Prediction Spectrum
ADMET Predictor	Simulations Plus, Inc.	No	ADME/T
SarDrop	Opabrium, Ltd	No	ADME/T
ADME Suite Tox Suite	Advanced Chemistry Development, Inc.	No	ADME Toxicity
ADMEWORKS Predictor	Fujitsu FQS	No	ADME/T
SarClasct	Strand Life Sciences	No	ADME/T
QikProp	Schrodinger, Inc.	No	ADME/T
TOPKAT	Accelrys, Inc.	No	Toxicity
Leadscope	Leadscope, Inc.	No	Toxicity
Meteor Derek Nexus	Lhasa, Ltd	No	Metabolism Toxicity
PASS	Russian Academy of Medical Sciences	No	Toxicity
HazardExpert Pro MetabolExpert ToxAlert MEXAlert RenomEx	CompuDrug, Ltd	No	Toxicity Metabolism Toxicity Metabolism Metabolism
METAPC CASETOX	Molnasec, Inc.	No	Metabolism Toxicity
VolSurf- Metasite	Molecular Discovery, Ltd	No	ADME Metabolism
Bioclipse	Uppsala University, Sweden and European Bioinformatics Institute	Yes	Metabolism
Metadrag	GeneGo, Inc.	No	Metabolism Toxicity
TMES	OASIS Linc	No	Metabolism Toxicity

Table 5 : programmes disponibles de la prédiction ADME/T (Liao et al., 2011).

1.3 Obtention de la structure de la cible

Afin d'obtenir de manière expérimentale la structure d'une protéine, plusieurs étapes sont nécessaires. Il faut isoler le gène, le cloner et l'exprimer dans un organisme adéquat, puis purifier la dose de protéine obtenue et parfois éliminer le fragment rajouté à la protéine pour la chromatographie d'affinité (Edwards et al., 2000). Ces étapes de biologie moléculaire restent encore délicates et limitent le travail. La structure tridimensionnelle est le plus souvent obtenue à partir de l'échantillon par diffraction de rayons X. pour ces méthodes, les molécules doivent auparavant être préparées sous forme de cristal, mais la cristallogenèse reste une étape hautement empirique et difficile. La possibilité d'obtenir des cristaux complexe protéine-ligand est bien plus intéressante, mais aussi plus difficile à réaliser (Abola et al., 2000). Il est également possible d'obtenir des structures par résonance magnétique nucléaire (RMN) dans les cas où la protéine ne comporte pas plus de 200

acides aminés ou par microscopie cryo-électronique (Siega, Van Duynhoven and Baldus, 1999; Kühlbrandt and Williams, 1999).

La majorité des structures protéiques sont disponibles et téléchargeables via la PDB (Protein Data Bank) (Berman et al., 2000), un grand nombre de structures d'une même molécule avec ou sans ligand permet d'avoir une information pertinente. Quand la méthode expérimentale n'est pas possible ou n'a pas été encore réalisée, la solution à envisager en premier est une approche par modélisation comparative (homomogy modeling) (Cavasotto and Phatak, 2009). Il s'agit de prédire un modèle d'après d'autres molécules le plus proche possible de la cible, et dont les structures sont connues expérimentalement.

1.4 Criblage virtuel

Suivant la nature de l'information expérimentale disponible sur la cible et/ou les ligands existants, on distingue deux approches distinctes pour le criblage virtuel. La première se base sur la structure de la cible et est connue sous le nom de "structure-based virtual screening" (SBVS) (Ghosh et al., 2006; Lyne, 2002). La seconde, reposant sur la connaissance d'un nombre suffisant d'informations concernant une ou plusieurs molécules actives (ligands) de référence, est appelée "ligand-based virtual screening" (LBVS) (Lengauer et al., 2004).

Bien que ces deux approches soient surtout utilisées de manière exclusive (souvent parce que la nature des données de départ ne laisse qu'un seul choix possible), leur combinaison lors d'une campagne de criblage permet de maximiser les chances de succès pour identifier de nouvelles touches (Muegge and Oloff, 2006; Peach and Nicklaus, 2009).

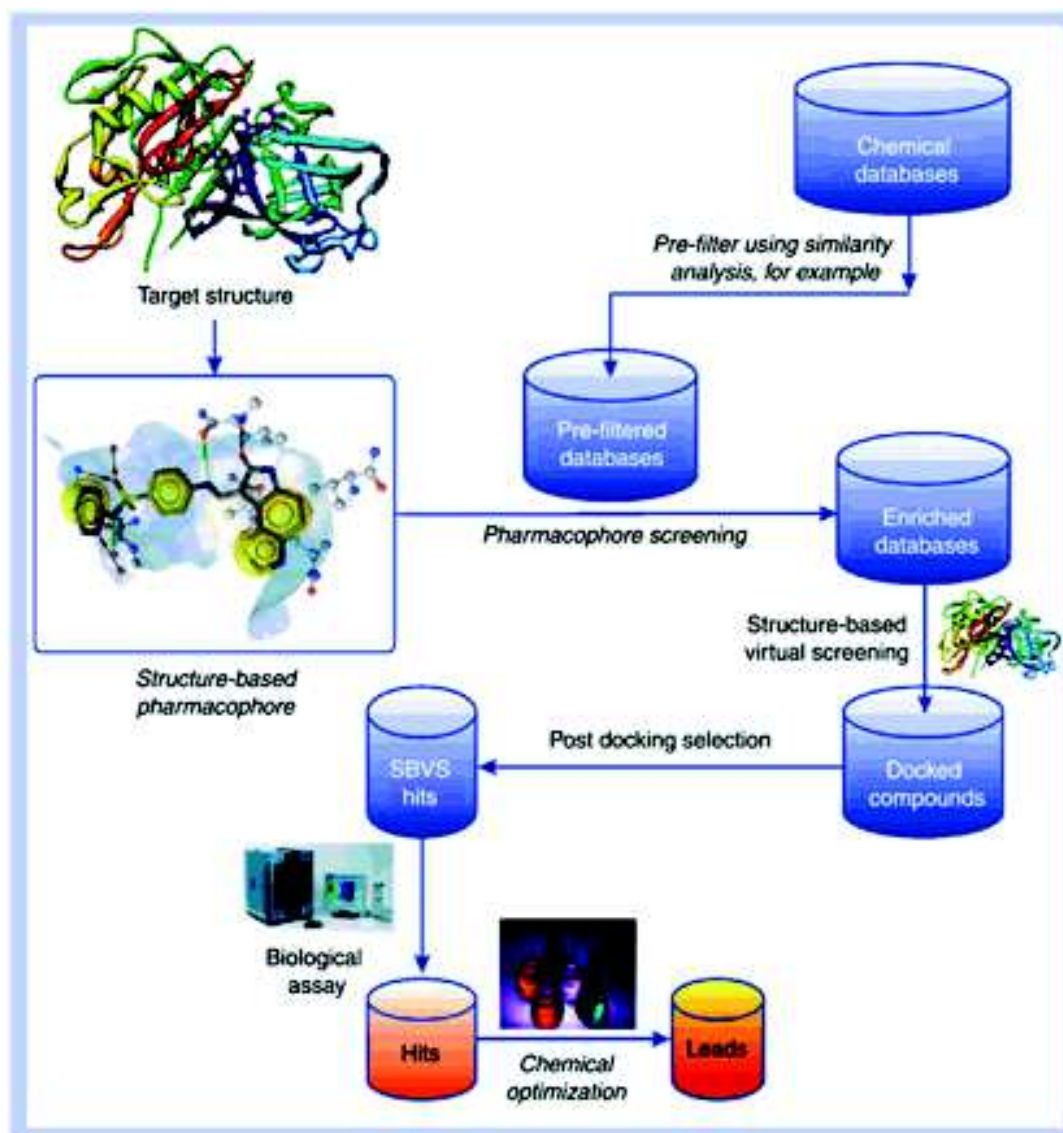


Figure 29 : protocole typique d'un criblage virtuel.

1.4.1 Les méthodes "ligand-based" (LBVS)

Elles consistent en premier lieu à classer la chimiothèque testée à travers divers descripteurs, comme les motifs pharmacophoriques 3D (Mason, Good and Martin, 2001) et autres (Todeschini and Consonni, 2000), décrivant les caractéristiques physico-chimiques et structurales des molécules. Ensuite, il s'agit d'identifier les molécules comportant une similarité avec les ligands connus pour être actifs sur la cible étudiée, ainsi qu'une dissimilarité avec les molécules connues pour être inactives (ou présentant des caractéristiques indésirables).

Cette approche, rapide et relativement simple à mettre en œuvre, présente comme inconvénient majeur son interdépendance envers les informations de référence utilisées pour construire le modèle de prédiction d'affinité. Par conséquent, les résultats ne sont souvent pas satisfaisants en termes de diversité. En particulier, il est a priori peu probable

d'identifier de nouvelles classes de molécules actives dont les propriétés diffèreraient de celles des ligands actifs connus.

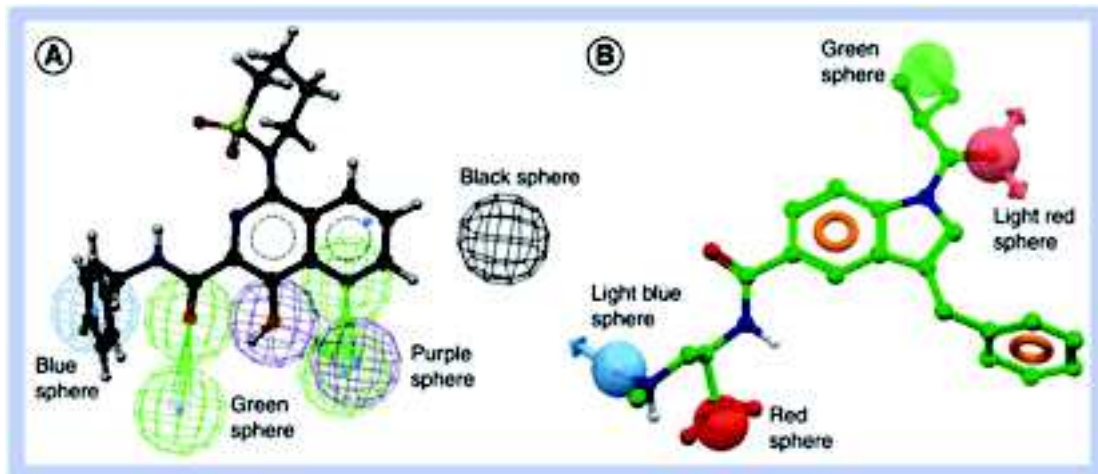


Figure 30 : Exemples de pharmacophores utilisés en (LBVS) (Liao et al., 2011).

(A) Un pharmacophore produit par Catalyst. Sphère bleue: caractère hydrophobique; sphère verte: accepteur de liaison hydrogène ; sphère pourpre : donneur de liaison hydrogène ; sphère noir: volume exclu. (B) Un pharmacophore produit par Phase. Sphère rouge clair: accepteur de liaison hydrogène; sphère bleu clair: donneur de liaison hydrogène; sphère verte: caractère hydrophobique; sphère rouge : caractère négative e; orange torus: noyau aromatique.

1.4.2 Les méthodes "structure-based" (LBVS)

Elles se rapportent souvent aux algorithmes de docking protéine-ligand et quelquefois aux recherches basées sur un motif pharmacophorique du site actif de la protéine cible. Ces méthodes ne consistent pas à rechercher des touches comportant une similarité avec des composés connus, mais à estimer la complémentarité structurale de chaque molécule criblée avec le site actif considéré. Ainsi, contrairement à l'approche basée sur des ligands de référence, cette approche peut potentiellement identifier de nouvelles classes de molécules actives. En revanche, ces méthodes sont généralement plus coûteuses en puissance de calcul et leur emploi requiert souvent une expertise plus importante.

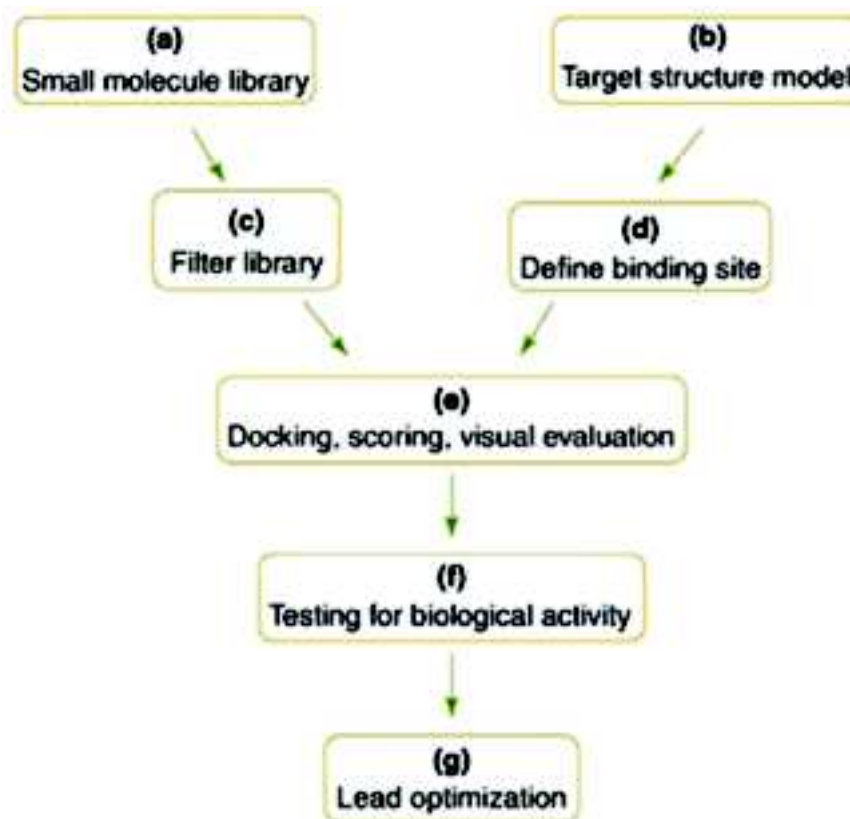


Figure 31 : Déroulement d'une expérience de criblage virtuel par docking. (Ghosh et al., 2006).

Dans un premier temps, la chimiothèque à cribler (a) est préparée (uniformisation des données, ajout des protons à pH physiologique, conversion en 3D, etc.), puis préfiltrée (c) comme décrit précédemment.

Concernant la cible, il faut disposer d'une structure tridimensionnelle fiable de la protéine ciblée (en particulier au niveau de son site actif) (b). Une fois la cible sélectionnée et préparée (vérification de sa structure, sa protonation, prise en compte de certaines molécules d'eau, etc.), le site actif de celle-ci est défini ; cela consiste à déterminer les résidus clés du site de liaison qui forment le lieu de reconnaissance moléculaire avec les ligands (d).

Quand les petites molécules et la cible sont prêtes, un algorithme de docking, associé à sa fonction de score, est employé pour identifier et classer les composés de la chimiothèque ayant les meilleures affinités avec la cible (e).

Une inspection visuelle des solutions de docking des ligands les mieux classés par la fonction de score peut avoir lieu, afin de vérifier leurs modes d'interaction avec les résidus clés du site actif (e). Les molécules candidates obtenant les meilleures prédictions d'affinité constituent enfin des touches (hits) potentielles qu'il faut valider en les soumettant aux tests expérimentaux (f). A l'issue de ces essais, les composés démontrant une activité biologique peuvent passer en phase d'optimisation en vue de déterminer des têtes de série (leads) (g).

2 Le docking moléculaire

Le docking moléculaire consiste en un algorithme informatique automatisé qui prédit et à reproduit les interactions protéine-ligand ; Cette méthode tente d'imiter le processus réunissant une protéine et un ligand pour former un complexe non covalent, et de révéler la complémentarité électrostatique et stérique entre les deux.

Le docking moléculaire et son utilisation dans le criblage virtuel ont fait l'objet de plusieurs revues (Brooijmans and Kuntz, 2003; Abagyan and Totrov, 2001; Kitchen et al., 2004; Sousa, Fernandes and Ramos, 2006), ainsi que l'évaluation et la comparaison de plusieurs programmes de docking (Zhou et al., 2007; Onodera, Satou and Hirota, 2007; Cross et al., 2009; Plewczynski et al., 2011).

Même si du point de vue de la précision, certains programmes actuels se distinguent, il peut être difficile de déterminer dans l'absolu qu'un programme de docking donné "fonctionne mieux" qu'un autre ou de tirer des conclusions générales sur leur pouvoir prédictif ; ceci pour plusieurs raisons (Cole et al., 2005; Liao et al., 2011). En effet, chaque programme de docking est unique par la représentation du système, la méthode de recherche et la fonction de score employée. Cette spécificité se répercute sur sa pertinence en termes de rapport entre précision et efficacité de calcul, rapport qui dépend aussi des paramètres employés pour les calculs et de la nature de la cible étudiée. C'est pourquoi, on assiste par exemple à l'émergence de fonctions de score spécifiques à certaines classes de protéines (Laederach and Reilly, 2003). Enfin, le jugement de la qualité d'un programme peut se faire à plusieurs niveaux, notamment sur la qualité des poses des ligands les mieux classés, sur sa prédiction de l'affinité de liaison ou encore sur son efficacité à cribler des chimiothèques.

Name	Developed by	Incorporated into software package	Free for academia	Drug-design applications
AutoDock	Scripts Research Institute	-	Yes	Allosteric reductase inhibitors Rac1 inhibitors Trypanothione reductase inhibitors
DOCK	University of California, San Francisco	-	Yes	STAT3 dimerization inhibitors Death-associated protein kinase inhibitors Inhibitors of osteoclast formation and bone resorption
FlexX	Bioscience GmbH	LeadIT	No	Inhibitors of penicillin binding protein Inhibitors of ATP-phosphoribosyl transferase Human histamine H4 receptor ligands
FRLED	OpenEye Scientific Software	OpenEye	Yes	Proteasome inhibitors Heat-shock protein 90 inhibitors
Guide	Schrodinger, Inc	Schrodinger	No	Inhibitors of dengue virus methyltransferase EGFR1 kinase inhibitors HIV-1 integrase inhibitors
GOLD	Cambridge Crystallographic Data Centre	-	No	Topoisomerase I inhibitors MINK1 inhibitors Mol tyrosine kinase inhibitors
ICM	Molsoft LLC	Molsoft	No	TNF- α inhibitors Aryl hydrocarbon receptor ligands GTP competitive inhibitors
Surfer-Dock	Tripos Inc	SYBYL	No	Glycogen synthase kinase inhibitors Proteinase inhibitors HIV-1 reverse transcriptase inhibitors

Table 6 : programmes de docking les plus utilisés dans les criblages virtuels (Liao et al., 2011)

2.1 Principe

Le processus du docking est itératif et chaque passe de calcul s'articule en deux étapes. Tout d'abord, une portion limitée de l'espace des conformations du complexe protéine-ligand est explorée afin d'améliorer la pose (orientation et/ou conformation) du ligand dans le site actif de la protéine. L'étape suivante fait intervenir une fonction de score qui évalue la qualité de la pose générée (Gohlke and Klebe, 2002). Cette estimation in silico de l'affinité du ligand pour la cible est basée sur un examen simplifié des interactions entre les deux partenaires. La répétition du cycle de recherche associé à l'estimation du score guidant l'exploration de l'espace doit assurer la convergence de l'algorithme vers un état représentatif du minimum global de l'énergie libre d'association.

2.1.1 Représentation du système

La plupart des méthodes de docking reposent sur une représentation simplifiée qui peut réduire la protéine aux atomes du site actif et du voisinage de celui-ci, et dans laquelle les molécules du solvant ne sont pas représentées explicitement.

Les modèles utilisés sont généralement analogues à ceux de la mécanique moléculaire, dans lesquels chaque atome des structures moléculaires est représenté comme un point unique sur lequel ses propriétés sont projetées. Moins fréquemment, les ligands et/ou les sites actifs peuvent être modélisés par l'intermédiaire de surfaces.

Les propriétés des ligands et/ou des sites actifs réparties dans l'espace peuvent également être représentées à travers leur projection sur une grille virtuelle.

2.1.2 La recherche de la pose optimale

Un système biologique réduit à la zone d'interaction protéine-ligand fait intervenir trois acteurs dans le phénomène de reconnaissance moléculaire : la protéine, le ligand et le solvant. Un algorithme de docking idéal prendrait en compte tous les paramètres associés à ces trois composantes, tels que la complète flexibilité du ligand et de la protéine, les effets de solvant, etc. Toutefois, une telle approche implique l'exploration d'un espace composé d'un nombre très important de degrés de liberté. La résolution du docking par une recherche exhaustive nécessiterait alors des temps de calcul considérables. Pour cette raison, la complexité du système est généralement réduite au niveau de la modélisation structurale. Les approximations les plus courantes à ce niveau consistent à considérer la protéine comme une entité rigide et à ne pas représenter explicitement les molécules du solvant.

2.1.2.1 Docking rigide

Dans le cas des méthodes de docking rigide, la protéine et le ligand sont traités comme étant des entités rigides ; la recherche de la pose optimale se limite au positionnement (translation et/ou rotation) du ligand. Cette opération consiste en la recherche exhaustive dans l'espace discrétisé des degrés de liberté.

Il est possible de considérer indirectement la flexibilité des ligands en utilisant des programmes de docking rigide. Pour cela, un jeu de conformères de basse énergie pour chaque molécule à tester sur le site actif ciblé peut être généré et utilisé. Une telle approche est nécessaire dans le contexte du docking rigide, car il est admis que la conformation complexée d'un ligand peut varier considérablement par rapport à sa conformation "libre" (Nicklaus et al., 1995).

2.1.2.2 Docking semi-flexible

Dans le cas des méthodes de docking semi-flexible, la protéine est traitée comme étant rigide et le ligand comme étant totalement flexible. Lorsque l'espace conformationnel des ligands est exploré, le nombre de degrés de liberté de l'espace de recherche peut être conséquent dans le cas de molécules très flexibles. Dans un tel contexte, l'emploi de méthodes de recherche exhaustives apparaît souvent inapproprié, car nécessitant des simplifications importantes au niveau de l'échantillonnage, des algorithmes de recherche de la flexibilité du ligand par plusieurs approches sont utilisés.

a. L'approche combinatoire (incrémentale)

Dans un premier temps le ligand est découpé en parties rigides et flexibles. Entre les points ou des rotations sont possibles, une ou plusieurs ancres rigides sont définies, ensuite une première partie rigide est mise en interaction avec le récepteur, puis les parties flexibles sont ajoutées de manière successive avec une exploration des angles de torsion.

b. L'approche aléatoire ou stochastique

Les programmes de docking semi-flexible considérés comme les plus efficaces emploient des méthodes de recherche aléatoires ou stochastiques. L'exploration de l'espace de recherche se fait de façon plus ou moins aléatoire et les états générés sont soit acceptés, soit rejetés, suivant des règles spécifiques. Les deux principales classes de méthodes aléatoires sont :

- les méthodes de Monte Carlo, générant aléatoirement des états du système acceptés ou rejetés sur la base probabiliste de la fonction de Boltzmann.
- les méthodes évolutionnaires basées sur des algorithmes génétiques, s'inspirant des théories de l'évolution pour sélectionner les états.

2.1.2.3 Docking flexible

De nombreuses études utilisant le docking par l'approche ligand flexible/protéine rigide ont montré que cette stratégie semi-flexible conduisait à des résultats concluants (Alvarez, 2004; Ghosh et al., 2006). Toutefois, les réussites issues de son utilisation concernent surtout l'étude de protéines relativement rigides. À la suite de leur liaison avec un ligand, de nombreuses protéines peuvent pourtant subir des réarrangements structuraux de plus ou moins grande amplitude. Négliger cet aspect, désigné sous le terme d'"induced fit" (Koshland Jr., 1994), peut avoir des conséquences fâcheuses sur la pertinence des résultats issus d'un docking (Erickson et al., 2004). Même concernant des changements conformationnels mineurs, la flexibilité du site actif de la protéine peut avoir une grande influence dans le phénomène de reconnaissance moléculaire avec un ligand (Teague, 2003).

La mise au point d'algorithmes de docking prenant en compte explicitement la flexibilité de la protéine est toutefois une tâche délicate, car le nombre de degrés de liberté associé à une telle représentation peut se révéler très important (>> 50).

Beaucoup d'efforts ont été fournis afin de surmonter cet obstacle (Kokh and Wenzel, 2008), et les approches qui en résultent peuvent être classés en deux grandes catégories:

a. L'approche ensemble-docking

Qui consiste en la reconstruction du récepteur en combinant plusieurs conformations (expérimentales ou issues de simulations MD) de la macromolécule cible.

Les premières techniques de cette approche impliquent l'utilisation d'une grille composite incorporant multiples structures cristallographiques et/ou RMN des complexes protéines-ligands (Knegtel, Kuntz and Oshiro, 1997), et ont donné des résultats encourageants, même si elles ne tiennent pas compte de la nature combinatoire des conformations utilisées.

Les méthodes des simulations MD (Dynamique Moléculaire) (Pak and Wang, 2000), de Monte Carlo (Morris et al., 1998), et la génération de bibliothèques de rotamères (conformations énergétiquement raisonnables) pour un ensemble de chaînes latérales de la protéine

autorisées à être flexibles (Kallblad and Dean, 2003) ont été aussi mises en œuvre dans le but d'introduire, au moins partiellement, la flexibilité du récepteur de la macromolécule.

Des techniques alternatives (Sudbeck et al., 1998) impliquent l'utilisation des structures cristallines de nombreux complexes protéine-ligand, et la superposition de leurs régions les plus stables, dans le but de créer une carte composite du site de liaison.

Il est aussi possible d'effectuer un docking sur chacune des multiples conformations (Barril and Morley, 2005; Wong et al., 2005). Cette approche systématique peut se révéler coûteuse en temps et en ressources computationnelles, mais a l'avantage de ne pas dépendre des paramètres et de la pertinence d'un programme spécifique.

b. L'approche induced fit

Qui échantillonne la flexibilité du récepteur lors d'une simulation de docking. Cette approche semble plus prometteuse, car elle permet l'exploration complète de l'espace conformationnel de la protéine, en couvrant un plus large ensemble d'orientations du récepteur (Scherman et al., 2006). Néanmoins, cette méthode prend beaucoup plus de temps, parce que la flexibilité de la protéine nécessite l'échantillonnage d'un nombre important de degrés de liberté dans la simulation. Ces méthodes ne peuvent pas généralement être utilisées dans les SBVS, principalement en raison de leurs exigences de calcul trop élevées (basse vitesse) pour le criblage de grandes chimiothèques.

Le docking rigide permissif (soft docking) considère la flexibilité de la protéine de manière indirecte, en atténuant certains termes de répulsion. Cela peut permettre au ligand de pénétrer légèrement la surface de la protéine en prévision des réarrangements qui auraient eu lieu lors de l'association de partenaires flexibles (Ferrari et al., 2004). Une telle approche indirecte est cependant limitée pour couvrir le spectre des effets d'induced fit.

Enfin, de nouveaux programmes de docking qui traitent la flexibilité du site actif en employant la même stratégie que celle des ligands (chaque rotamère constitue un degré de liberté), sont actuellement en développement (Cozzini et al., 2008).

2.1.3 L'évaluation du score des poses

L'évaluation et le classement des différentes conformations du ligand positionné dans le site actif de la protéine constituent le second aspect crucial du docking, après la nature de la recherche conformationnelle. Une fonction de score doit théoriquement pouvoir distinguer, parmi les différentes poses générées pour un ligand, celles qui correspondent aux modes de liaison les plus représentatifs de la reconnaissance moléculaire. Ainsi, elle doit permettre, sur la base d'une pose optimale proposée, de différencier les molécules bioactives des autres composés inactifs vis-à-vis de la protéine.

Des techniques basées sur le calcul d'énergie libre ont été développées et constituent un moyen quantitatif rigoureux pour estimer l'affinité de liaison d'un ligand pour une protéine (Kollman, 1993; Simonson, Archontis and Karplus, 2002). Bien qu'elles soient très précises, leur coût (en temps, en contraintes et en expertise) ne les rend pas appropriées pour une application de docking dans le cadre d'un criblage virtuel. Pour représenter le meilleur compromis entre vitesse et précision, les fonctions de score implémentées dans les programmes de docking sont donc basées sur la simplification des phénomènes impliqués dans la reconnaissance moléculaire, en particulier de ceux qui sont délicats à évaluer en dehors des calculs d'énergie libre (comme l'entropie).

Les différentes fonctions de score implémentées dans les programmes de docking ont fait l'objet de nombreuses publications (Ferrara et al., 2004; Cheng et al., 2009; Wang et al., 2004). Elles sont généralement classées suivant trois catégories : les fonctions de score basées sur un champ de force, les fonctions de score empiriques et celles basées sur des connaissances statistiques. Enfin, plusieurs fonctions de score de nature différente peuvent être combinées pour former des fonctions de score dites de consensus.

2.1.3.1 Les fonctions de score basées sur un champ de force

Les champs de force dans leur forme standard évaluent la somme de deux énergies : l'énergie entre atomes liés au sein d'une molécule donnée (énergie interne) et l'énergie entre atomes non-liés. Ce deuxième terme correspond au terme principal de l'énergie d'interaction protéine-ligand dans le cas d'une application à un problème de docking. La plupart du temps, les fonctions de score basées sur un champ de force ne considèrent qu'une conformation donnée de la protéine. Ainsi, si l'on compare l'activité de deux ligands par rapport à cette même conformation, cela permet de faire abstraction du terme d'énergie interne de la protéine qui s'annule dans l'expression de la différence d'énergie libre d'interaction entre les deux ligands.

L'affinité d'un ligand donné pour le site actif, mesuré en tant qu'énergie d'interaction, correspond à la somme des énergies d'interaction de van der Waals (souvent représentée par un potentiel de Lennard-Jones) et électrostatiques (potentiel de Coulomb). On peut ajouter à la fonction de score un terme d'énergie interne du ligand. Tous ces termes peuvent s'exprimer à partir des paramètres du champ de force considéré.

De telles fonctions de score présentent certaines limitations qui s'ajoutent à celles induites par la représentation du système (p. ex. la non représentation explicite du solvant). En particulier, les effets d'entropie, qui peuvent varier d'un ligand à l'autre pour un site actif donné, aussi bien que pour un ligand donné d'un site actif à l'autre, ne sont pas pris en compte. Seule la contribution enthalpique de l'énergie libre d'interaction est ainsi prise en compte.

2.1.3.2 Les fonctions de score empiriques

Ce type de fonction de score approxime l'énergie libre de liaison en sommant de façon pondérée différents termes d'interaction dérivés de paramètres structuraux. Les différents poids de la fonction de score sont ajustés pour reproduire en priorité des données expérimentales, telles que les constantes de liaison tirées d'un jeu d'entraînement de complexes protéines-ligands.

La plupart des programmes de docking implémentent ce type de fonction de score témoignant de leur efficacité (en termes de rapport précision/rapidité). Cependant, le principal inconvénient de ces fonctions empiriques est leur forte dépendance aux données utilisées pour les calibrer qui, en cas de mauvaise paramétrisation, peut limiter leur transférabilité sur des systèmes différents.

2.1.3.3 Les fonctions de score basées sur des connaissances statistiques

Ces fonctions de score sont construites à partir de règles fondées sur une analyse statistique des complexes protéines-ligands résolus expérimentalement. Elles partent du principe que les distances interatomiques les plus représentées statistiquement dans les

complexes constituent des contacts énergétiques favorables et, qu'à l'inverse, les plus rares représentent des interactions moins stables.

Ainsi, leur paramétrisation dépend de la quantité d'informations expérimentales disponibles, et on doit leur apparition à la profusion de données structurales accessibles dans des bases de données telles que la Protein Data Bank.

2.1.3.4 Les fonctions consensus

Ces fonctions hybrides combinent les résultats issus de diverses fonctions de score. On estime qu'il est possible de compenser partiellement les faiblesses intrinsèques de chacune des fonctions de score employées, évitant leurs erreurs individuelles et ainsi d'augmenter la probabilité d'identifier des composés actifs. Cependant, si les termes des différentes fonctions de score sont fortement corrélés, l'intérêt du consensus devient limité, car il peut entraîner une amplification des erreurs, au lieu de les atténuer (Feher, 2006).

2.1.3.5 Autres types de fonctions de score

Les méthodes de docking reposant sur des surfaces utilisent des fonctions de score qui sont adaptées et principalement centrées sur la complémentarité géométrique surface-surface, et auxquelles il peut être adjoint une estimation d'interactions sur le modèle des fonctions de score plus conventionnelles.

Chapitre IV : Matériels et Méthodes

1 Matériels

Pour tout le travail in silico réalisé dans la présente étude, le package Schrödinger Suite 2011 a été utilisé, installé sur trois machines : un PC portable (processeur Intel Core 2 Duo 2GHz, 4 Go de RAM, plateforme : Windows 7), un PC portable (processeur Intel Core I 5, 2.5 GHz, 4 Go de RAM, plateforme : Windows 7) et un PC de bureau (processeur Intel Core I3 3GHz, 4 Go de RAM, plateforme : Windows 8).

Schrödinger Suite 2011 est un package de Modélisation Moléculaire très complet et très puissant qui comprend les programmes et applications, et les solutions les plus sophistiqués et les plus performants dans le monde de la bio-informatique, la chimio-informatique et la conception des médicaments assistée par ordinateur

1.1 Les programmes et applications Schrödinger

1.1.1 Maestro 9.2

Maestro est une interface graphique intuitive pour toutes les applications, et les solutions Schrödinger. Maestro comprend aussi une gamme d'outils de visualisation (plusieurs types de représentation 3D et 2D, de colorations, de labélisation ...etc.), et de manipulations des structures (table de structures, alignement de séquences et de structures, calcul RMSD, 2D skecter...etc.) très pratiques et sophistiqués.

1.1.2 Glide 5.7

Glide est une solution complète pour le docking ligand-récepteur ; il offre un spectre complet de vitesses et de précisions, allant du criblage virtuel à haut débit de millions de composés aux prédictions de modes de liaison extrêmement précises, en fournissant des enrichissements constamment élevés à tous les niveaux.

1.1.3 Phase 3.3

Phase, est un programme polyvalent et à haute performance pour la conception des pharmacophores, l'alignement des structures, la prédiction de l'activité, et la génération et la recherche de bases de données de structures 3D.

Phase fournit un support rapide, précis et hautement configurable, pour la découverte de touches (hits), le développement SAR (structure activity relationship), l'optimisation de têtes de série. Phase peut être également utilisé comme une source d'alignements moléculaires par tiers programmes de 3D QSAR (quantitative structure activity relationship).

1.1.4 Prime 3.0

Prime, est un progiciel puissant et innovateur pour la prédiction précise des structures de protéines. **Prime** est un outil très précis de la modélisation comparative et la reconnaissance des repliements de protéines. Le chemin de modélisation comparative comprend le processus complet de la prédiction de structures de protéines et la reconnaissance de motifs, à l'alignement, jusqu'à la construction du modèle, au raffinement des structures (prédiction de chaînes latérales, prédiction de boucles, et minimisation).

1.1.5 PrimeX 1.8

PrimeX, est Un package complet pour le raffinement précis des structures cristallines de protéines, utilisant le champ de force OPLS-AA avec des technologies très sophistiquées.

1.1.6 Desmond 3.0

Desmond, programme extrêmement performant pour les simulations de MD des systèmes biomoléculaires.

1.1.7 CombiGlide 2.7

CombiGlide, programme de technologie combinatoire et resquillage d'ancres (core hopping) pour la découverte et l'optimisation de têtes de série (leads).

1.1.8 QSite 5.7

QSite, un programme de QM/MM (mécanique quantique/mécanique moléculaire) à haute performance. **QSite** applique la mécanique quantique au centre réactif du site actif de la protéine, et la mécanique moléculaire pour le reste du système.

1.1.9 Impact 5.7

Impact, est un programme de mécanique et dynamique moléculaires, qui constitue la composante des calculs de mécanique moléculaire de Glide, de Liaison, et de QSite. Certaines fonctionnalités de base d'Impact, comme la minimisation d'énergie par mécanique moléculaire (MM), les simulations de dynamique moléculaire (MD), les simulations hybrides de Monte Carlo (HMC), et la solvatation explicite des structures sont disponibles à travers l'interface Maestro.

1.1.10 MacroModel 9.9

MacroModel, un programme polyvalent, riche en fonctionnalités pour la modélisation moléculaire. **MacroModel** comprend les principaux champs de force, des modèles de solvatation précis, et des méthodes de recherche de conformations de pointe pour fournir un package des plus complets et performants de modélisation moléculaire.

1.1.11 MCPRO+ 2.0

MCPRO+, un package Monte Carlo à usage général utilisant la théorie des perturbations d'énergie libre pour la prédiction précise de l'énergie libre relative, et les affinités de liaison relatives des complexes protéines- ligands peuvent être prédites par des approches multiples.

1.1.12 Liaison 5.7

Liaison, est un programme qui prédit les affinités de liaison des complexes ligands-récepteurs utilisant un modèle linéaire d'approximation (LIA) d'interaction qui est adapté à un ensemble d'énergies de liaison libres connues.

1.1.13 Jaguar 7.8

Jaguar, est un package de mécanique quantique ab initio à haute performance pour la simulation la phase gaz et solution, avec une puissance particulière dans le traitement des systèmes contenant des métaux.

1.1.14 QikProp 3.4

QikProp, Prédiction ADME/T rapides des médicaments candidats. **QikProp** évalue efficacement les propriétés pharmaceutiques pertinentes de plus d'un demi-million de composés pendant 1 heure, en générant un grand nombre de propriétés et de descripteurs moléculaires.

1.1.15 Strike 2.0

Strike, un logiciel puissant pour la modélisation statistique et QSAR.

1.1.16 LigPrep 2.5

LigPrep, un programme pour la génération de modèles moléculaires 3D précises. LigPrep va bien au-delà des simples conversions de structures

2D en 3D, en incluant des variations tautomériques, stéréochimiques, et d'ionisation, ainsi que la minimisation d'énergie et des filtres souples pour générer des banques de ligands entièrement personnalisés, qui sont optimisés pour de plus amples analyses informatiques.

1.1.17 Epik 2.2

Epik, Prévisions de pKa rapides et robustes, en combinant la fiabilité éprouvée des méthodes Hammett et Taft avec des outils de tautomérisation puissants, pour la prédiction précise des états de protonation des ligands dans des conditions biologiques.

1.1.18 SiteMap 2.5

SiteMap, un outil rapide, précis et pratique pour l'identification de sites de liaison, et la détermination de leurs caractéristiques et propriétés, et même le scorer.

1.2 Les Solutions Schrödinger

1.2.1 Protein Preparation Assistant

Protein Preparation Assistant fournit une interface pratique pour corriger les défauts dans les structures des protéines et les préparer pour différents types d'utilisation.

1.2.2 Induced Fit Docking:

Le protocole IFD de Schrödinger est nouvelle méthode de prédiction rapide, précise, et innovatrice pour modéliser les changements conformationnels du site actif induits par la liaison du ligand ; en fusionnant le pouvoir prédictif de Prime avec les capacités de score de Glide, IFD est un nouvel outil qui prédit les géométries induites du site actif avec un succès remarquable.

1.2.3 Ligand and Structure Based Descriptors

Ligand and Structure Based Descriptors est une solution pratique et efficace pour la hiérarchisation des composés et la prévision des affinités de liaison, intégrant de façon cohérente cinq différentes méthodes pour prédire les affinités de liaison.

1.2.4 Quantum-Polarized Ligand Docking

Quantum-Polarized Ligand Docking est solution de recherche innovatrice qui combine la puissance de Glide avec la précision de QSite Le premier algorithme de ce type, QM-Polarized Ligand Docking utilise la méthode ab initio pour calculer les charges du ligand dans l'environnement de la protéine, offrant une précision nettement améliorée par rapport aux algorithmes de docking MM pures.

1.2.5 Virtual Screening Workflow

Virtual Screening Workflow fournit une interface pratique pour la mise en place de préparations flexibles des ligands, de filtrages, et de dockings par Glide, pour cribler virtuellement un grand nombre de ligands contre une protéine cible.

2 Méthodes

Un protocole de criblage virtuel a été élaboré suite à plus d'une centaine d'essais de docking sur différentes structures Hsp90 et mettant en jeu différents types de ligands ; ces essais avaient pour but de déterminer les méthodes et les paramètres les plus appropriés pour la préparation des ligands et des récepteurs, ainsi que pour les simulations de docking.

Ce protocole comprend les étapes suivantes :

- Sélection de huit structures cristallographiques pour servir de cibles pour le criblage.
- Test de la fiabilité du programme de docking utilisé.
- Criblage virtuel poussé d'une petite chimiothèque de flavonoïdes, les plus abondants dans l'alimentation.
- Criblage virtuel d'une chimiothèque de flavonoïdes et de coumarines relativement grande.
- Evaluation visuelle et optimisation des meilleures poses.

Ce protocole comprend plusieurs enrichissements pertinents qui sont le fruit d'une analyse approfondie de la Hsp90 et ses ligands, à savoir :

- L'utilisation de plusieurs structures cristallographiques (ensemble- docking) pour introduire partiellement le phénomène d'induced fit de la Hsp90 et diminuer le taux de faux négatifs.
- L'implication d'une contrainte spécifique à la Hsp90 : une liaison hydrogène caractéristique avec le résidu Asp93 (Asp79 pour la levure) dans toutes les simulations de docking, et la prise en compte de quatre molécules d'eau conservées dans les dockings et dans l'évaluation des poses ; cette approche permet d'éliminer un grand nombre de poses incorrectes et d'obtenir des modes de liaison très précis, et par conséquent, l'élimination d'un grand nombre de faux positifs.
- l'utilisation des technologies de pointe de la suite Schrödinger 2011 d'une manière bien pensée et pertinente comme : la méthodologie Glide XP, la recherche conformationnelle sur deux champs de force (OPLS et MMFFs), et Prime MM/GBSA, ce qui donne une grande précision et une bonne qualité des simulations.
- L'utilisation de plusieurs critères dans l'évaluation des poses (XP Score, Prime MM-GBSA ΔG libre de liaison, le mode de liaison) ce qui filtre finement les faux positifs et augmente considérablement les chances de réussite.

2.1 Choix des cibles

L'un des facteurs limitants majeurs des techniques de docking est la flexibilité de la protéine cible, notamment les réarrangements conformationnels au niveau du site actif induits par la liaison du ligand (induced fit), qui conduit souvent, au cours du docking, à la génération de faux négatifs (molécules actives biologiquement, prédites à tort avec une faible ou non-activité in silico). La génération de faux négatifs correspond à une perte d'information pertinente au cours du criblage. La génération de faux positifs, inévitable au cours d'un criblage, est moins handicapante, car elle ne correspond qu'à une perte de temps dans le cadre d'un protocole multi-étapes.

La Hsp90 en est un exemple, avec des réarrangements importants au niveau de son site ATPasique N-terminal induits par la liaison des ligands, ce qui fait, par exemple, que l'inhibiteur PU3 (9-butyl-8-(3,4,5- trimethoxy-benzyl)-9h-purin-6-ylamine) cocrystallisé avec la Hsp90 code PDB : 1UY6 ne peut être amarré correctement dans le site actif de la Hsp90 : 1AMW cocrystallisée avec le substrat naturel ADP.

Plusieurs approches ont été proposées pour pallier ce problème (voir : III, 3.1.2.3 Docking flexible) ; pour le présent protocole, l'approche ensemble- docking a été adoptée, et pour considérer, ne serait-ce que partiellement, la flexibilité du site de liaison de la Hsp90, huit structures Hsp90 cristallographiques ont été choisies.

code PDB	Isomère Hsp90	Ligand cocrystallisé	Résolution (angströms)	Référence
1AMW	Hsp90 de la levure	ADP	1.85	(Prodromou et al., 1997)
1BGQ	=	Radiol col	2.50	(Prodromou et al., 1997)
1UY6	Hsp90 α de l'homme	PU3	1.9	(Wright et al., 2004)
2BRC	Hsp90 de la levure	3,4-diarylpyrazole	1.6	(Cheung et al., 2005)
2BSM	Hsp90 α de l'homme	5 VER-49009	2.05	(Dymock et al., 2005)
2BZ5	=	1-(2-phenol)-2-naphthol	1.9	(Barril et al., 2005)
2H55	=	PU-DZ8	2.00	(Immormino et al., 2006)
2IWX	Hsp90 de la levure	Analogue du radiol col	1.5	(Proisy et al., 2006)

Tableau 1 : structures des Hsp90 utilisées dans le criblage virtuel.

Le choix des structures Hsp90 utilisées assure, outre les critères conventionnels (bonne résolution, absence de mutations, présence d'un ligand cocrystallisé...), une bonne diversité conformationnelle du site de liaison qui dérive de celle des ligands cocrystallisés, qui sont dans ce cas le substrat naturel ADP et d'autres puissants inhibiteurs naturels et synthétiques.

2.2 Préparation des cibles

Un fichier de structure PDB typique n'est pas adapté pour une utilisation immédiate en modélisation moléculaire. Un fichier PDB typique comporte uniquement des atomes lourds, et peut inclure un ligand cocrystallisé, des molécules d'eau, des ions métalliques, et des cofacteurs. Certaines structures sont multimériques, ce qui peut nécessiter leur réduction à une seule unité, et elles peuvent éventuellement manquer d'informations sur les connectivités, qui doivent être attribuées, ainsi que l'ordre des liaisons et les charges formelles.

Pour la préparation des structures protéiques cibles, **Protein Preparation Assistant** de Schrödinger a été utilisé, selon les étapes suivantes :

- Les structures cocrystallisées protéines-ligands sont importées et visualisées dans Maestro.
- Les complexes multimères sont simplifiés ; une seule chaîne est retenue.
- La structure de la protéine est ajustée en ajoutant les hydrogènes, en assignant l'ordre des liaisons et les charges formelles, en connectant les ponts disulfures, et en complétant éventuellement les chaînes latérales des résidus manquants (elles sont prédites, ajoutées, et minimisées par Prime 3.0).
- Les molécules d'eau sont supprimées, et ne sont retenues que quatre molécules interstitielles très conservées, dont trois, entourant le groupement carboxyle du résidu Asp93 (Asp79 pour la Hsp90 de la levure), sont maintenues dans tous les dockings réalisés ; la quatrième, pouvant être déplacée ou remplacée par certains groupements, n'est maintenue que dans les dockings de test (ligand cocrystallisé avec son récepteur) lorsqu'elle existe dans le récepteur Figure 32, Figure 33.
- La structure du ligand cocrystallisé est ajustée en assignant l'ordre des liaisons et les charges formelles, et en assignant l'état d'ionisation et de tautomérisation (prédits et assignés par Epik 2.2).
- Enfin, les complexes sont raffinés et minimisés en deux étapes: en premier lieu, l'orientation des hydrogènes et des molécules d'eau est échantillonnée exhaustivement et les hydrogènes sont minimisés afin d'optimiser les liaisons d'hydrogène, puis, en deuxième lieu, tout le complexe est minimisé (relaxation) en ne permettant qu'un très léger mouvement du squelette de la protéine (pour ne pas altérer sa structure) avec une convergence vers un RMSD de 0.30 (la minimisation est réalisée par Impact 5.7 en utilisant le champ de force OPLS 2005).

2.3 Test de la fiabilité du programme de docking utilisé, Glide 5.7

Les simulations de docking ont été réalisées par Glide 5.7 incorporé dans le package Schrödinger Suite 2011.

2.3.1 Glide 5.7

L'algorithme Glide (Grid-Based Ligand Docking With Energetics)

(Friesner et al., 2004; Halgren et al., 2004) entreprend une recherche systématique des positions, des orientations, et des conformations du ligand dans le site de liaison du récepteur utilisant une série de filtres hiérarchiques. La forme et les propriétés du récepteur sont représentées sur une grille par plusieurs différents ensembles de champs fournissant des scores progressivement précis de la pose du ligand. Les champs sont calculés avant le docking. Le site de liaison est défini par une boîte rectangulaire confinant les translations du centre de masse du ligand. Un ensemble de conformations initiales du ligand sont générées par une recherche exhaustive des minima de torsion, et les conformères sont clustérisés d'une façon combinatoire. Chaque cluster, caractérisé par une conformation commune de l'ancre (core) et un ensemble exhaustif de conformations des groupes rotamères, est amarré comme objet simple dans la première phase.

La recherche commence par une phase de positionnement et de score approximative qui rétrécit de manière significative l'espace de recherche et réduit le nombre de poses à considérer plus tard à une centaine. Dans l'étape suivante, les poses sélectionnées sont minimisées sur des grilles d'interactions van der Waals et électrostatiques précalculées pour le récepteur sur le champ de force OPLS. Dans l'étape finale, les 5 à 10 poses aux plus basses énergies obtenues de cette façon sont soumises à un procédé de Monte Carlo,

à partir duquel les minima de torsion voisins sont examinés, et l'orientation des groupes périphériques du ligand est raffinée.

Les poses minimisées sont rescorées alors en utilisant la fonction de GlideScore, qui est une version développée de la fonction de score empirique ChemScore (Eldridge et al., 1997), avec des composantes basées sur le champ de force et des termes additionnels relevant de la désolvatation et des interactions répulsives.

La sélection de la meilleure pose est effectuée en utilisant un modèle énergétique de score Emodel, qui combine le score des énergies des grilles, l'affinité de liaison prédite par GlideScore, et l'énergie interne de la souche du ligand (le conformer représentant la pose à scorer).

Glide utilise deux modes de score : le mode de score PS (précision standard) et le mode de score XP (extra-précision) ; Le point de départ pour le mode XP est une version modifiée de ChemScore, comme dans le cas du mode PS ; cependant, des termes originaux sont employés pour manipuler des effets physiques absents dans ChemScore. Des pénalités relatives au coût de désolvation sont appliquées en amarrant des molécules d'eau explicites dans les complexes protéines-ligands ayant les scores les plus élevés issus du docking, et par évaluation de la désolvatation des groupes polaires et chargés du ligand et de la protéine en comptant le nombre de molécules d'eau voisines, et en comparant ces valeurs à des statistiques extraites d'une base de données de ligands actifs correctement amarrés.

Des motifs de reconnaissance moléculaire basés sur le concept du confinement hydrophobe du ligand par la protéine sont définis, et des augmentations incrémentales de l'affinité de liaison sont ajoutées aux scores des ligands, quand les motifs appropriés sont reconnus. Afin d'évaluer correctement ces nouveaux termes, l'échantillonnage de l'algorithme, réalisé à une résolution plus élevée augmente considérablement (surcoût de temps et de ressources computationnelles).

Des termes additionnels, tels que le traitement spécial des ponts en sel et des interactions de π -cation et de π - π stacking, et de divers autres motifs spécialisés sont aussi impliqués dans le mode XP (Friesner et al., 2006).

2.3.2 Test de Glide 5.7 par cross-docking

Plusieurs techniques sont utilisées pour évaluer les performances des différents algorithmes de docking. La prédiction du mode d'interaction consiste à déterminer le positionnement correct du ligand par rapport à son récepteur. La capacité d'un programme à réussir ce travail est habituellement jugée au moyen de la déviation quadratique moyenne ou RMSD (root-mean-square deviation) du modèle conçu par le logiciel vis-à-vis de la structure du ligand de référence (ligand cocrystallisé). Le limite admis est une différence maximum de 2 angströms, au-delà de laquelle la prédiction est considérée comme non adéquate (Vieth et al., 1998).

Pour tester la capacité de Glide 5.7 à prédire correctement les modes de liaison pour la Hsp90, un cross-docking a été réalisé en amarrant chaque ligand cocrystallisé des huit structures Hsp90 choisies dans chacune de celles-ci, et en calculant le RMSD de la meilleure pose du ligand avec son mode de liaison cristallographique. Ce cross-docking a été réalisé par la solution **Virtual Screening Workflow** de Schrödinger qui comprend cinq boîtes de dialogue : une pour l'introduction des ligands, la deuxième pour le préfiltrage des ligands, la troisième pour la préparation des ligands, la quatrième pour l'introduction ou la génération directe des grilles des récepteurs, et la cinquième pour le docking.

2.3.2.1 Préparation des ligands

Les huit ligands cocrystallisés avec les structures Hsp90 choisies ont été séparés de leurs récepteurs par Maestro ; leur état (ionisation et tautomérisation) a été prédit par Epik 2.2 à pH physiologique, puis ils ont été stockés dans un fichier mae (format Maestro) pour être introduits directement dans Virtual Screening Workflow (pas de préfiltrage).

2.3.2.2 Génération des grilles des récepteurs (recepteur grids)

Le docking par Glide se fait en deux étapes, la première consiste à générer une grille du récepteur (voir IV, 2.3.1 Glide 5.7) et la deuxième consiste à amarrer le ou les ligand(s) dans la grille déjà générée.

Les grilles des huit récepteurs Hsp90 ont été générées par le panneau Receptor Grid Generation de Glide 5.7 qui comprend quatre boîtes de dialogue ; la première pour la définition du récepteur et le facteur de graduation des interactions van der Waals, la deuxième pour la définition du site de liaison, la troisième pour la définition des contraintes (contraintes de position, contraintes de liaisons hydrogène et/ou métalliques, et contraintes hydrophobiques), et la quatrième pour le traitement (flexible ou rigide) des groupes rotatoires (les hydroxyles) du site de liaison.

Le site de liaison a été défini par une boîte rectangulaire de 13 Å de diamètre confinant les translations du centre de masse du ligand cocrystallisé avec le récepteur. Le facteur de graduation (scaling factor) des rayons des interactions van der Waals a été défini de 1 pour tout atome du récepteur ayant une charge partielle (valeur absolue) de [0,25] ou inférieure. Une seule contrainte de liaison hydrogène a été définie, avec un oxygène carboxylique du résidu Asp93 (Asp79 pour la levure).

Les grilles ainsi générées sont stockées dans des fichiers Glide-grid.zip, et sont utilisables dans autant de simulations de docking que l'on veut, tant qu'elles sont gardées.

2.3.2.3 Docking

Virtual Screening Workflow donne la possibilité de réaliser un docking multi-stages (4 stages), chaque stage des trois premiers correspond à un niveau de précision de la méthode de docking avec plusieurs options :

- Glide HTVS (High-Throughput virtual screening) : très rapide, mais avec une précision médiocre, utile pour chercher les grandes chimiothèques.
- Glide SP (standard precision) : un compromis entre la rapidité et la précision.
- Glide XP (extra-precision) : grande précision, mais relativement lente, utile pour chercher les petites chimiothèques.
- Postprocess with Prime MM-GBSA (Molecular Mechanics- Generalized-Born/Surface-Area) : pour obtenir les énergies libres de liaison des ligands en utilisant un modèle de désolvation en continuum (solvant implicite) GBSA.

Selon le choix, on peut impliquer les trois premiers stages dans le docking en déterminant le nombre ou le pourcentage des poses à retenir dans chaque stage pour qu'elles soient (les poses précédentes) investiguées dans le stage suivant, comme on peut en impliquer deux ou un seulement dans le docking ; le quatrième stage, lorsqu'il est impliqué, n'est appliqué qu'aux poses finales (issues du stage final, quel que soit ce dernier).

Pour le test de Glide, vu le nombre relativement petit des ligands mis en jeu et pour des raisons de précision, la méthodologie Glide XP a été utilisée directement, avec un stage final de Postprocess with Prime MM-GBSA ; les paramètres utilisés sont :

- Force field : Le champ de force OPLS 2005, qui est un champ de force très puissant adapté aux protéines et aux petites molécules organiques, mise à jour de OPLS 2001, développé par Schrödinger sur la base du champ de force OPLS original de W.L. Jorgensen (Jorgensen and Tirado-Rives, 1988), avec la modification et l'amélioration de plusieurs paramètres, et l'introduction d'autres nouveaux.
- Use Epik state penalties for docking : pour appliquer les pénalités énergétiques des états des ligands prédits par Epik (le coût énergétique de l'état d'ionisation).
- Scaling of ligand van der Waals radii for nonpolar atoms : Le facteur de graduation (scaling factor) des rayons des interactions van der Waals a été défini de 0.80 pour tout atome du ligand ayant une charge partielle (valeur absolue) de [0,15] ou inférieure.

Pour la méthodologie Glide XP, les options suivantes ont été choisies :

- Generate multiple input conformations : pour exécuter une recherche conformationnelle par MacroModel 9.9 sur les structures des ligands d'entrée correspondant aux meilleures poses de l'étape précédente pour localiser le conformer ayant la plus basse énergie (le minimum global), utilisant deux champs de force différents, OPLS
- 2005 et MMFFs (Halgren, 1996), puissant champ de force adapté aux protéines et aux petites molécules organiques. Cette option produit de deux structures supplémentaires d'entrée pour le docking XP en plus de la structure originale, une qui correspond au minimum global OPLS 2005 et l'autre au minimum global MMFFs, et par conséquent, augmente le coût de temps considérablement (plus du triple, y compris le temps de la recherche conformationnelle). Les variations des structures d'entrée produisent souvent de meilleurs résultats finals.
- Docking method : la méthode du docking flexible (flexibilité totale des ligands, récepteur rigide) a été choisie, avec pénalisation des conformations non-planaires des liaisons amides.
- Perform post-docking minimization : pour exécuter une minimisation des poses issues du docking final. La minimisation optimise des longueurs des liaisons et les angles pour rescorer les poses.
- Use selected constraints for each grid : la contrainte de liaison hydrogène, avec un oxygène carboxylique du résidu Asp93 (Asp79 pour la levure) définie pour toutes les grilles (IV, 2.3.2.2 La génération des grilles des récepteurs) a été appliquée, et toute pose ne vérifiant pas cette contrainte (c.-à-d., ne formant pas une liaison hydrogène avec un oxygène carboxylique du résidu Asp93, Asp79 pour la levure) est automatiquement rejetée.
- Generate up to N poses per compound state : 1 pose a été générée pour chaque ligand.
- After docking, keep : 100 % des poses générées ont été retenues pour être ensuite traitées par Postprocess with Prime MM-GBSA.

2.4 Criblage virtuel d'une petite chimiothèque de flavonoïdes

Vu qu'un nombre d'études épidémiologiques a associé des apports élevés en flavonoïdes avec une réduction du risque de différents cancers, et vu le grand nombre de voies de la carcinogenèse impliquant la Hsp90, une supposition a été faite que cet effet anticancéreux pourrait être lié à un certain mécanisme qui relie les flavonoïdes et la Hsp90.

Pour vérifier cette supposition, une chimiothèque composée de 25 flavonoïdes des plus abondants dans l'alimentation humaine a été passée à un protocole de criblage virtuel poussé, dans une tentative d'y découvrir de nouvelles molécules naturelles actives contre la Hsp90, et ce qui correspondrait à un nouveau mécanisme anticancéreux de ces substances végétales.

Tout comme pour le test de la fiabilité de Glide, la solution Virtual Screening Workflow de Schrödinger a été mise en œuvre pour réaliser le criblage virtuel de la petite chimiothèque.

2.4.1 Préparation de la petite chimiothèque

Cette petite chimiothèque comporte les flavonoïdes suivants :

La quercétine, le kaempférol, la myricétine, l'isorhamnétine, l'apigénine, la lutéoline, la chrysin, la (+)-catéchine, l'(-)-épicatéchine, la (+)- gallocatéchine, l'(-)-épicatéchine gallate, l'(-)-épigallocatéchine, l'(-)- épigallocatéchine gallate, la naringénine, l'hespérétine, l'ériodictyol, la cyanidine, la delphinidine, la pétunidine, la péonidine, la pèlargonidine, la malvidine, la daidzéine, la génistéine, et la glycitéine.

Les structures des molécules ont été téléchargées de la base de données PubChem (<http://pubchem.ncbi.nlm.nih.gov/>)

du National Center for Biotechnology Information, sous format SD 2D. Les structures ont été converties au format 3D par LigPrep 2.5, en générant jusqu'à quatre stéréo-isomères pour tout stéréo-centre non spécifié ; un seul conformer de basse énergie a été retenu pour chaque structure en échantillonnant seulement les petits noyaux (de 5/6 membres). Les états des structures (ionisation et tautomérisation) ont été prédits par Epik 2.2) à pH physiologique.

2.4.2 Génération des grilles des récepteurs

Six récepteurs, à savoir : 1AMW, 1UY6, 2BRC, 2BSM, 2BZ5, 2H55 ont été utilisées pour le criblage de la petite chimiothèque. Pour les récepteurs ne comportant que trois molécules d'eau conservées, (la quatrième étant déplacée par les ligands cocrystallisés, voir IV, 2.2

Préparation des cibles) : 2H55, 2BRC, et 2BSM, leurs grilles générées pour le test de Glide (voir, IV, 2.3.2.2 génération des grilles des récepteurs) ont été réutilisées ; et pour les récepteurs comportant quatre molécules d'eau conservées : 1AMW, 1UY6, et 2BZ5, les grilles ont été générées de nouveau avec les mêmes paramètres utilisés pour la génération des grilles du test de Glide, sauf que la quatrième molécule d'eau conservée, pouvant être déplacée ou remplacée par certains groupements chimiques des ligands, a été éliminée ici, mais elle sera prise en considération lors de l'évaluation des poses du docking.

2.4.3 Docking

La méthodologie Glide XP avec les mêmes paramètres et les mêmes options utilisés dans le test de Glide (voir, IV, 2.3.2.3 docking) a été choisie.

2.5 Criblage virtuel d'une chimiothèque relativement grande de flavonoïdes et de coumarines

Une chimiothèque de flavonoïdes et de coumarines plus grande et plus diversifiée a été criblée virtuellement par le présent protocole avec une stratégie un peu différente, pour des raisons de temps et de ressources computationnelles, vu le nombre relativement grand des structures à traiter.

2.5.1 Préparation de la grande chimiothèque

Plus de 5700 structures de flavonoïdes redondantes (un flavonoïde peut être représenté par plusieurs structures) et 400 structures de coumarines uniques ont été téléchargées de la base de données PubChem sous format

2D SD ; une base de données Phase Database a été générée à partir des structures téléchargées par Phase 3.3, selon les étapes suivantes :

- **Skip duplicate**structures : afin de n'introduire dans la Phase Database que des structures uniques (un flavonoïde est représenté par une seule structure) en raison de la redondance des 5700 structures téléchargées ; cette tâche est assurée par Ligfilter.
- **Preparation** : Les structures ont été converties au format 3D par LigPrep 2.5, en générant jusqu'à quatre stéréo-isomères pour tout stéréo-centre non spécifié ; un seul conformer de basse énergie a été retenu pour chaque structure en échantillonnant seulement les petits noyaux (de 5/6 membres). Les états des structures (ionisation et tautomérisation) ont été prédits par Epik 2.2 à pH physiologique.
- **Filtering** : un préfiltrage a été effectué sur les structures à introduire dans la Phase Database basé sur des propriétés ADME/T et des descripteurs moléculaires parmi un grand nombre d'autres générés par QikProp 3.4, à savoir (les valeurs entre parenthèses représentent les limites choisies pour le présent protocole ; toute molécule ayant une valeur hors ces limites est automatiquement exclue de la Phase Database) :
 - **#stars (0 – 5)** : nombre des propriétés et/ou des descripteurs dont les valeurs sortent de l'intervalle des valeurs similaires pour 95 % des médicaments connus (violations). Un grand nombre d'étoiles suggère qu'une molécule est moins drug-like que les molécules avec quelques étoiles. Les propriétés et les descripteurs suivants (pour leur signification voir Annexe 01) sont inclus dans la détermination des
 - **#stars**: MW, dipole, IP, EA, SASA, FOSA, FISA, PISA, WPSA, PSA, volume, #rotor, donorHB, accptHB, glob, QPpolrz, QPlogPC16, QPlogPoct, QPlogPw, QPlogPo/w, QlogS, QPLogKhsa, QPlogBB,#metabol.
 - **mol_MW (≤ 600)**: le poids moléculaire du ligand.
 - **accptHB (>0)**: Estimation du nombre de liaisons hydrogène que le ligand peut donner.
 - **donorHB (>0)**: Estimation du nombre de liaisons hydrogène que le ligand peut accepter.
 - **#rtvFG (0-1)** : Nombre des groupes fonctionnels réactifs ; la présence de ces groupes peut conduire à des faux positifs dans des bio-essais,

et à une décomposition, une réactivité ou à des problèmes de toxicité in vivo.

- **RuleOfFive (0-2)** : Nombre des violations de la règle des cinq de
- Lipinski (Lipinski et al., 2001). Les règles sont les suivantes: mol_MW
- <500, QPlogPo/w <5, donorHB ≤ 5, accptHB ≤ 10. Les composés qui satisfont à ces règles sont considérés drug-likes.

La Phase Database de flavonoïdes et de coumarines ainsi générée contient 1445 structures, dont 1050 sont des structures uniques (le reste sont des stéréo-isomères) prêtes à être passées au criblage virtuel.

2.5.2 Docking

Les 1445 structures de la Phase Database ont été passées au criblage utilisant Virtual Screening Workflow en deux stages avec les paramètres (champ de force, pénalité Epik, facteur de graduation) utilisés pour le test de Glide (voir IV, 2.3.2.3 docking) :

- Le premier stage en mode Glide SP (standard precision) contre les huit grilles des récepteurs générées auparavant, à savoir : 1AMW, 1BGQ, 1UY6, 2BRC, 2BZ5, 2H55, 2IWX, avec les options suivantes : méthode du docking flexible, minimisation post-docking, et contrainte d'une liaison hydrogène avec un oxygène carboxylique du résidu Asp93 (Asp79 pour la levure).
- Le deuxième stage en mode Glide XP (extra-precision) a passé les structures issues du premier stage au criblage contre les grilles des deux récepteurs qui ont enregistré les meilleurs scores en mode Glide SP, à savoir : 1UY6 et 2H55, avec les mêmes options utilisées pour le test de Glide (voir IV, 2.3.2.3 docking) et le post-traitement par Prime MM-GBSA.

2.6 Evaluation des résultats

Les résultats de Virtual Screening Workflow pour chaque récepteur sont stockés dans des fichiers OUT_le nom du récepteur_pv.maegz ; Virtual Screening Workflow génère aussi un fichier FINAL_DOCK_OUT.mae qui comporte la meilleure pose pour chaque ligand et le récepteur avec lequel elle a été obtenue. Ces résultats sont visualisés et étudiés à travers l'interface Maestro.

Dans l'évaluation des poses, les critères suivants ont été pris en compte :

- **Un bon score de docking** : par rapport aux scores des ligands cocrystallisés avec la Hsp90.
- **Une basse Prime MM-GBSA ΔG bind** : l'énergie libre de liaison prédite par Prime MM-GBSA.
- **Le mode de liaison de la pose** : évaluation du mode de liaison de la pose par comparaison aux modes de liaison expérimentaux.

Le critère clé de cette évaluation est basé sur le pouvoir du ligand à déplacer la quatrième molécule d'eau conservée.

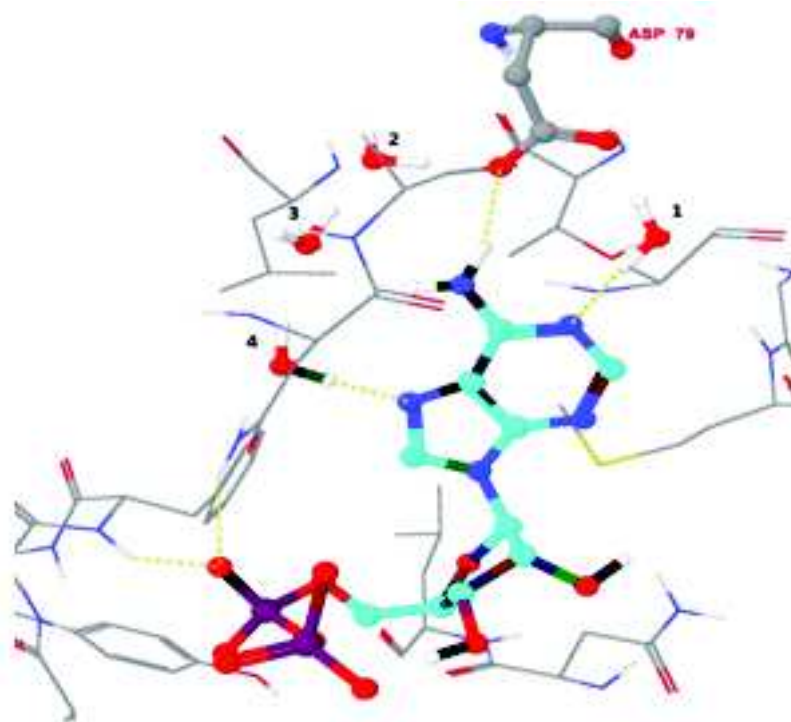


Figure 32 : structure PDB : 1AMW avec le substrat naturel ADP et les quatre HOH conservées autour de l'Asp79 (Asp93 pour la Hsp90 de l'homme). à noter la liaison hydrogène OH---N entre HOH 4 et la purine.

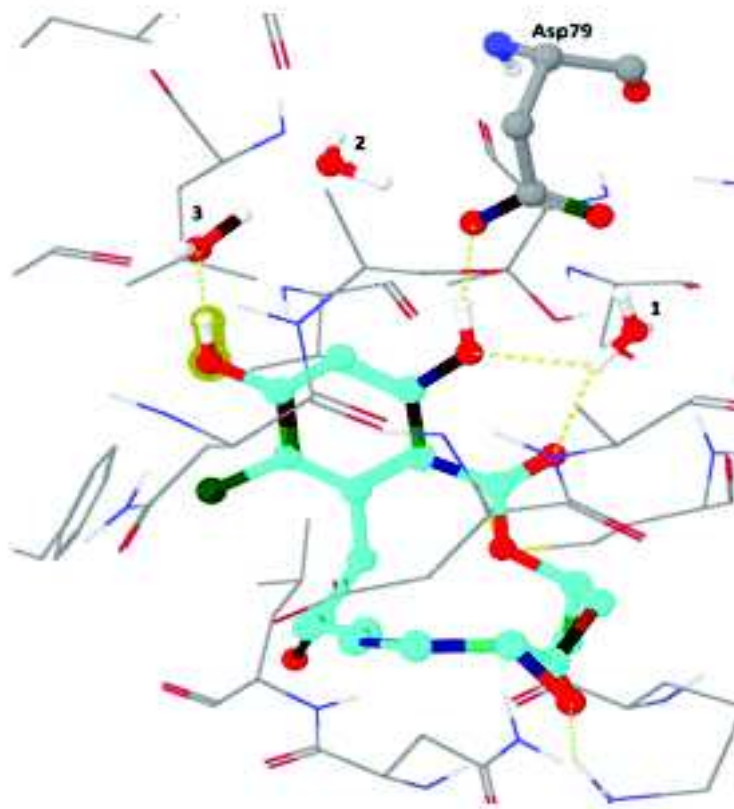


Figure 33 : structure PDB 1BGQ avec l'inhibiteur naturel Radicicol et seulement trois HOH conservées, la quatrième ayant été déplacée et remplacée par l'hydroxyle du radicicol (en surbrillance).

Comme Glide ne simule pas ce phénomène du déplacement des molécules d'eau, la molécule HOH4 a été éliminée des grilles des récepteurs du criblage virtuel (voir IV, 2.4.2 Génération des grilles des récepteurs), mais elle est prise en considération dans l'évaluation des poses. Une pose n'est sélectionnée que si sa molécule comporte soit un atome accepteur à une distance convenable de l'oxygène de HOH4 (comme le cas de l'ADP Figure 32) pour former une liaison hydrogène avec lui, soit un groupement capable de déplacer HOH4 (comme le cas de l'hydroxyle du radicicol Figure 33).

Quantum-Polarized Ligand Docking :

Les meilleures poses issues du criblage de la grande chimiothèque sont réanalysées par la solution QM-Polarized Ligand Docking de Schrödinger LLC.

Le protocole QM-Polarized Ligand Docking vise à améliorer les charges partielles sur les atomes du ligand dans une simulation de docking par Glide en les remplaçant par des charges dérivées des calculs par la mécanique quantique sur le ligand dans l'environnement du récepteur. De cette façon la polarisation des charges sur le ligand par le récepteur est expliquée, et le redocking des ligands avec ces nouvelles charges peut améliorer la précision de la simulation. Le protocole fonctionne en prenant un petit ensemble de poses de meilleurs scores pour chaque ligand, en calculant les charges par QSite 5.7, puis en réamarrant chacune de ces poses dans le récepteur, pour, enfin, sélectionner la meilleure pose de l'ensemble.

Les paramètres suivants ont été choisis pour le protocole :

- Pour le récepteur, les grilles générées précédemment ont été utilisées.
- Pour le docking initial, les charges ont été générées par la méthode semi-empirique NDDO, Glide XP a été choisi, le facteur de graduation (scaling factor) des rayons des interactions van der Waals a été défini de 0.8 pour les atomes du ligand, toute pose ayant un RMSD inférieur à 0.5 Å et un déplacement atomique inférieur à 1.3 Å est considérée comme dupliquée et est éliminée, et enfin, un maximum de 5 poses ont été gardées pour chaque ligand.
- Pour le calcul des charges dérivées du quantum, **Jaguar 7.8** a été utilisé avec le niveau de précision accurate (précis) basée sur la théorie de la densité fonctionnelle pour le ligand dans le contexte du récepteur.
- Pour le redocking, Glide XP a été choisi, le facteur de graduation (scaling factor) des rayons des interactions van der Waals a été défini de 0.8 pour les atomes du ligand, toute pose ayant un RMSD inférieur à 0.5 Å et un déplacement atomique inférieur à 1.3 Å est considérée comme dupliquée et est éliminée, et enfin, la meilleure pose a été gardée pour chaque ligand.
- La sélection finale est basée sur GlideScore.

Chapitre V : Résultats et Discussion

1 Résultats du test de la fiabilité du programme Glide 5.7

Name	GScore (XP)	Prime MMGBSA ΔG bind	RMSD	ensemble receptor	conformer field	#stars	#rtvFG	Rule Of Five
2H55-ligand	-13.784	-97.953	2.431	H55	OPLS	0	2	1
2BSM-ligand	-12.886	-104.310	0.507	BSM	MMFFS	0	0	0
2BRC-ligand	-12.253	-84.065	1.763	BRC	Original	0	0	0
1BGQ-ligand	-12.212	-85.945	0.448	BRC	Original	0	3	0
1AMW-ligand	-12.160	-43.049	0.859	AMW	OPLS	4	2	2
1UY6-ligand	-12.045	-87.564	0.976	H55	OPLS	0	0	0
2IWX-ligand	-11.819	-80.252	0.309	BRC	Original	0	2	0
2BZ5-ligand	-11.433	-71.155	1.726	BRC	OPLS	0	0	0

Tableau 2 : résultats du test de Glide

GScore (Glide Score) : est le score de la pose basé sur un modèle énergétique (Emodel) ; ce score est une approximation de l'énergie de liaison et donc, plus il est négatif, meilleure est l'affinité prédite. **GScore = (0.065 × vdW) + (0.130 × Coul) + Lipo + Hbond + Metal + BuryP + RotB + Site.**

- **vdW** : l'énergie de Van der Waals.
- **Coul** : l'énergie de Coulomb.
- **Lipo** : l'énergie des contacts lipophiliques.
- **HBond** : l'énergie des liaisons hydrogène.
- **Metal** : l'énergie des liaisons métalliques.
- **BuryP** : pénalité des groupes polaires enfouis.
- **RotB** : pénalité de l'immobilisation des liaisons rotatives du ligand.
- **Site** : l'énergie des contacts polaires dans le site actif.

Pour le mode Glide XP, des termes additionnels sont pris en compte (voir IV, 2.3.1 Glide 5.7)

- Prime MM-GBSA ΔG bind : l'énergie libre de liaison prédite par Prime sur le modèle de solvatation électrostatique GBSA.

- ensemble receptor : le récepteur avec lequel la meilleure pose du ligand a été obtenue.
- conformer field : le champ de force sur lequel a été obtenu le conformer de la pose suite à la recherche conformationnelle.
- #stars, #rtvFG, et Rule Of Five : pour leur signification voir (IV, 2.5.1 Préparation de la grande chimiothèque)

Discussion :

On voit que les RMSD des poses sont tous ≤ 2 Å (valeurs positives), sauf celui du ligand de 2H55 (2H55-ligand) qui a un RMSD de 2.431, mais même pour ce ligand Glide a été capable de reproduire le mode de liaison expérimental exactement Figure 34, la valeur de RMSD légèrement supérieure à 2 étant due seulement à une mauvaise conformation d'une hétéro-chaine exposée au solvant peu importante. Pour le reste, on peut dire que le pouvoir prédictif de Glide pour les modes de liaison de la Hsp90 est plus ou moins fiable, du moins pour les structures PDB choisies. Ces bons résultats sont dus en partie aux enrichissements du protocole utilisé, notamment la contrainte de la liaison hydrogène, qui ont aidé Glide à prédire avec fiabilité les modes expérimentaux et d'éviter les poses incorrectes.

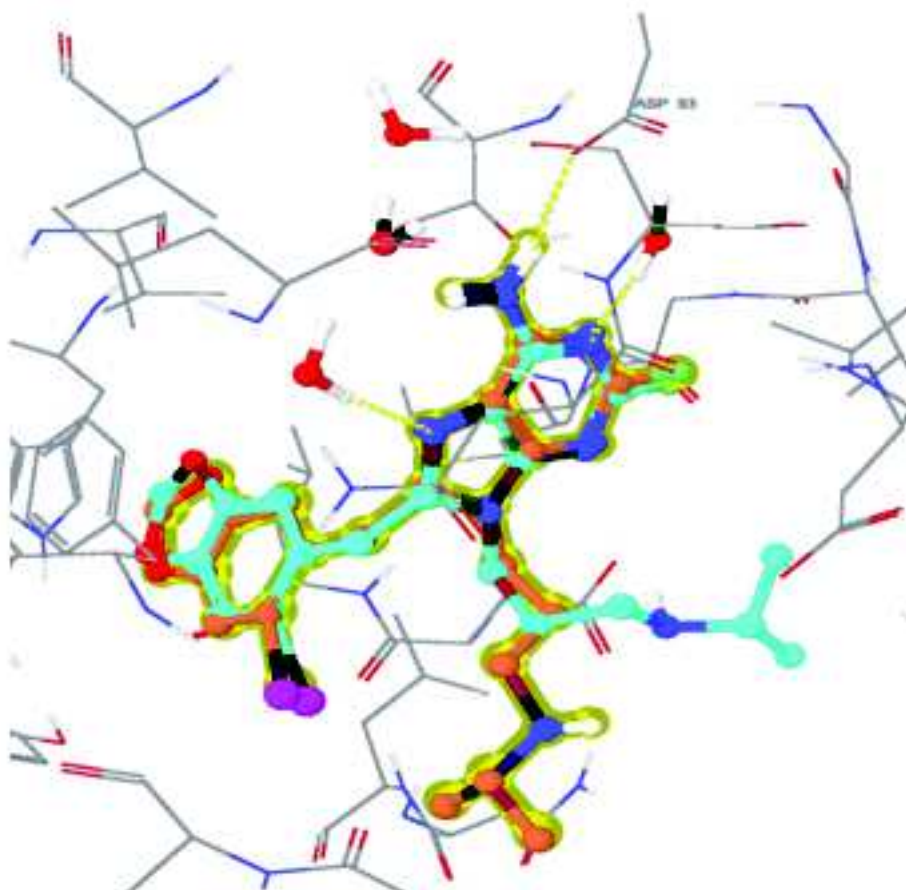


Figure 34 : superposition de la meilleure pose de H55-ligand (en surbrillance) et de son mode de liaison cristallographique.

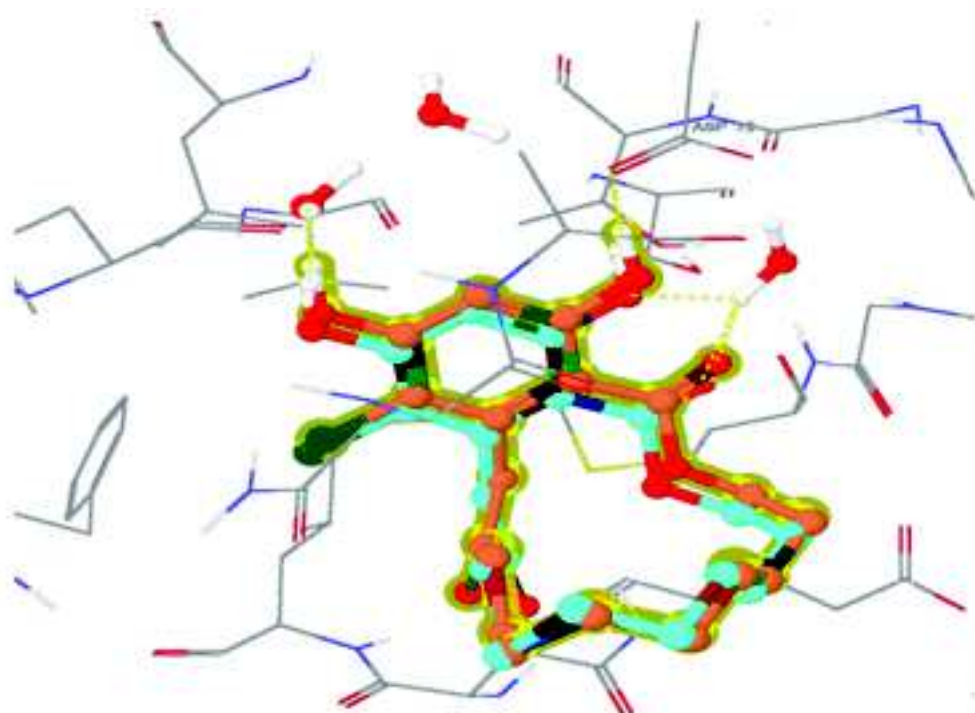


Figure 35 : superposition de la meilleure pose de 2IWX-ligand (meilleur RMSD : 0.309, en surbrillance) et de son mode de liaison cristallographique.

2 Résultats du criblage virtuel de la petite chimiothèque de flavonoïdes

name	GScore (XP)	Prime MMGBSA ΔG bind	ensemble receptor	conformer field	#stars	#rtvFG	Rule Of Five
(-)-Epigallocatechin gallate	-18.198	-99.996	UY6	MMFFS	5	1	2
(-)-Epicatechin gallate	-17.667	-97.083	UY6	MMFFS	4	1	1
(-)-Epigallocatechin	-13.243	-45.286	BRC	MMFFS	0	0	1
(+)-Gallocatechin	-12.921	-62.044	UY6	OPLS	0	0	1
Myricetin	-12.859	-49.519	BRC	Original	0	0	1
(+)-Catechin	-12.436	-68.205	UY6	Original	0	0	0
Eriodictyol	-12.315	-65.511	UY6	Original	0	0	0
(-)-Epicatechin	-12.192	-54.230	H55	Original	0	0	0
Delphinidin	-11.749	-48.105	BRC	MMFFS	0	0	1
Luteolin	-11.699	-56.619	H55	OPLS	0	0	0
Quercetin	-11.520	-53.973	BRC	Original	0	0	0
Peonidin	-11.399	-66.185	H55	Original	0	0	0
Naringenin	-11.361	-66.241	UY6	MMFFS	0	0	0
Glycitein	-11.276	-62.179	BRC	MMFFS	0	0	0
Genistein	-11.001	-56.579	BRC	MMFFS	0	0	0
Hesperetin	-10.988	-49.186	BRC	MMFFS	0	0	0

Tableau 3 : résultats du criblage virtuel des 25 flavonoïdes les plus abondants dans l'alimentation

Discussion

l'évaluation des 25 poses basée sur leur GScore, leur énergie libre de liaison, et leur mode de liaison (pour les critères de l'évaluation, voir, IV, 2.6 Evaluation des résultats), a révélé que la majorité des flavonoïdes criblés seraient peu ou pas actifs contre la Hsp90, du moins pour ce criblage virtuel, excepté trois molécules qui pourraient avoir une activité potentielle plus ou moins puissante contre la Hsp90.

En effet, l'(-)-épigallocatechine comme exemple, bien qu'elle ait un GScore de -13.243 (\uparrow), n'a pas une assez bonne MM-GBSA ΔG libre de liaison : -45.286 (\downarrow). En plus, le mode de liaison de sa meilleure pose indique qu'elle n'aurait pas une grande affinité à la Hsp90, vu qu'il ne présente aucun groupement susceptible de remplacer HOH4 et aucun atome accepteur pour former une liaison hydrogène avec elle. Son C8-H du noyau phényle A serait en clash stérique et électrostatique avec HOH4 Figure 37.

Davies et al (Davies et al., 2012) ont remarqué que le remplacement d'un N par un CH lors de la modification de la purine en pyrrolopyrimidine a causé la perte de l'affinité de liaison de la molécule Figure 36.

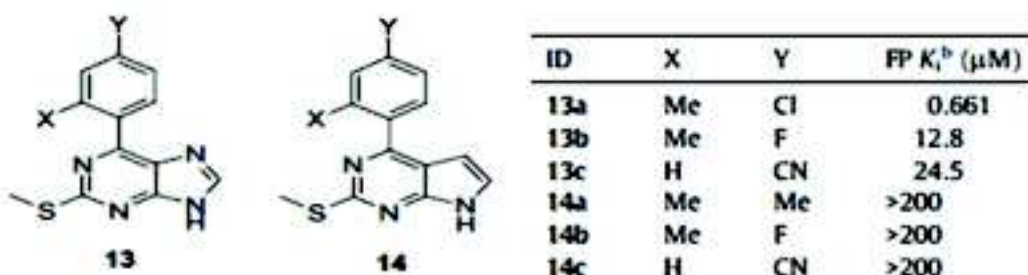


Figure 36 : la modification de la purine (13) en pyrrolopyrimidine et la perte de son affinité (>200). Le N de la purine (13) est remplacé par un CH dans la pyrrolopyrimidine (14).

Ceci a été attribué à un changement dans les interactions type liaison hydrogène entre le core du ligand (purine ou pyrrolopyrimidine) et une molécule d'eau structurellement conservée (HOH4).

La modification de la purine en pyrrolopyrimidine remplace une liaison hydrogène O—H---N entre HOH4 et la purine avec un clash stérique du O—H----H—C5 entre HOH4 et la pyrrolopyrimidine, très probable explication de la perte d'affinité pour les composés 14a-c Figure 36.

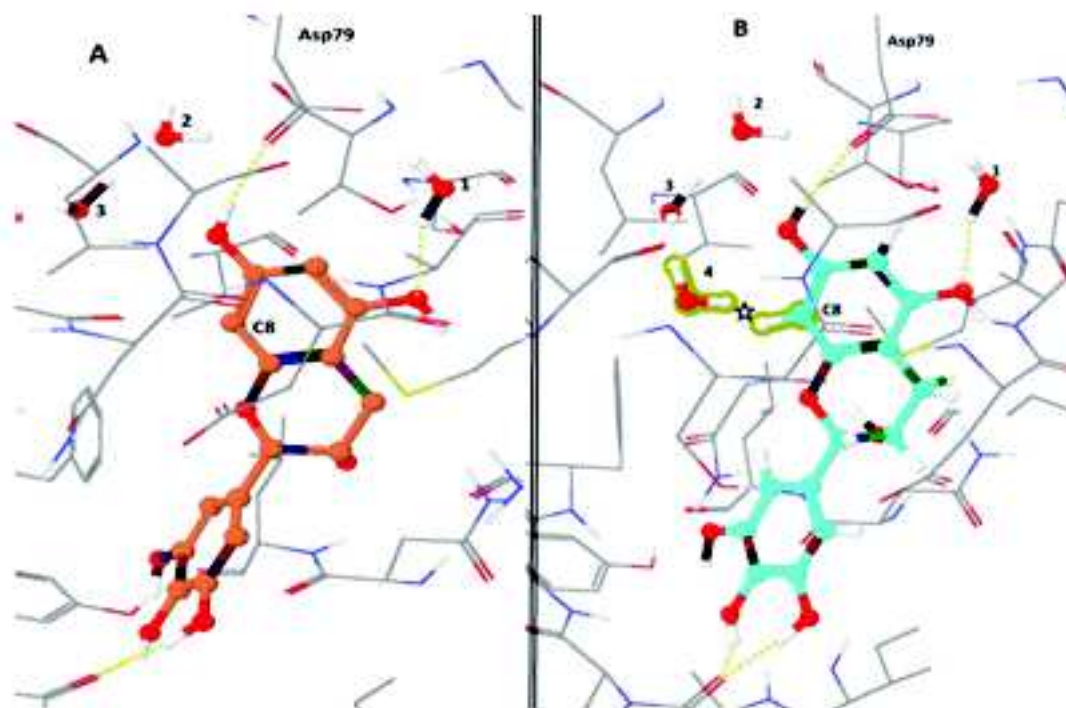


Figure 37 : A : mode de liaison de l'(-)-épigallocatéchine ; B : rétablissement de HOH4 et simulation de son clash stérique et électrostatique avec C8—H.

Par contre, l'(-)-épigallocatéchine gallate, l'(-)-épicatéchine gallate, et la (+)-gallocatéchine, avec leur groupement hydroxyle en C5 du noyau phényle A, seraient capables de déplacer HOH4 Figure 38, et par conséquent éviteraient le clash stérique et électrostatique décrit ci-dessus.

Ce raisonnement est basé sur (a) le fait que les deux molécules d'eau HOH3 et HOH4 soient dans un environnement semi-hydrophobe (Leu48, Ile91 et Val186), qui n'est pas, à priori, le plus favorable environnement pour une molécule d'eau ; (b) plusieurs structures cristallographiques montrent clairement que HOH4 peut être remplacée par certains groupements chimiques, entre autres, les hydroxyles Figure 33.

Le déplacement des molécules d'eau « piégées » peut parfois avoir un gain entropique, en permettant des interactions favorables directes additionnelles entre le ligand et la protéine (Davies et al., 2012).

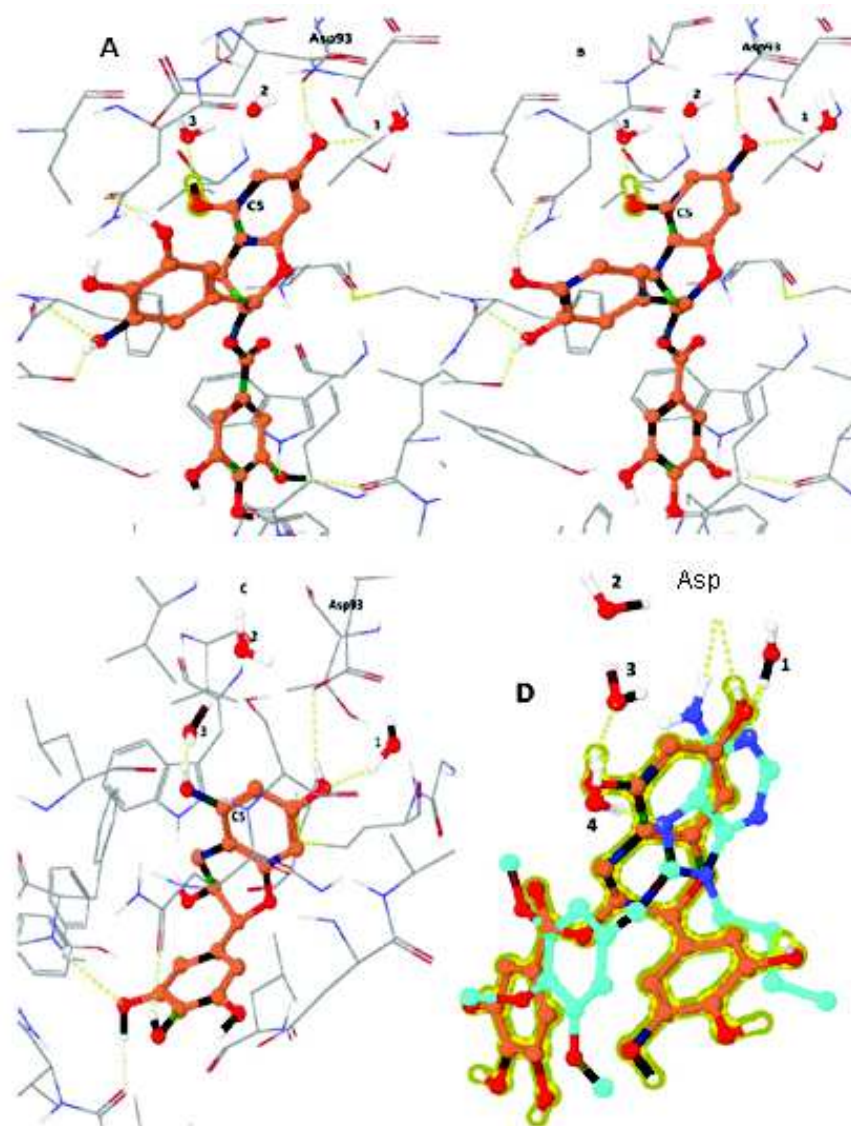


Figure 38 : A: mode de liaison de 1'(-)-épigallocatechine gallate; B: mode de liaison de 1'(-)-épicatéchine gallate; C: mode de liaison de la gallocatéchine; D: superposition de 1'(-)-épigallocatechine gallate (en surbrillance) et PU3 (1UY6-ligand).

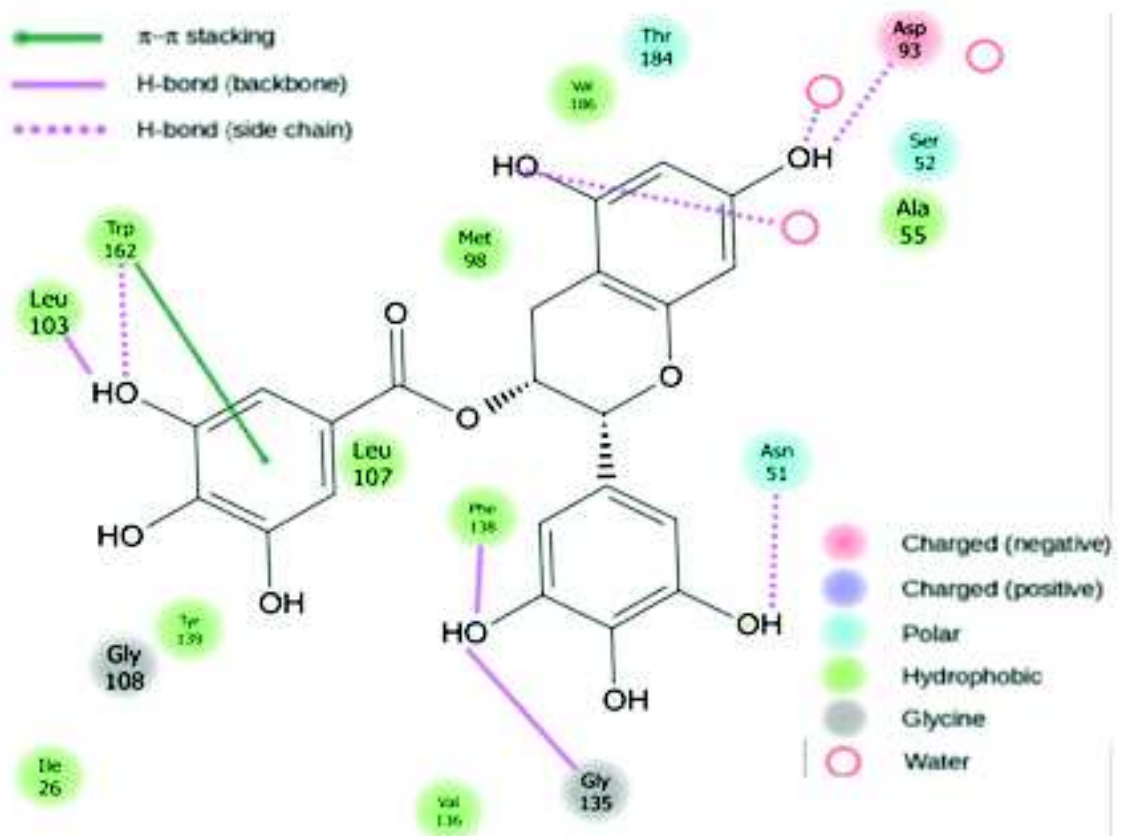


Figure 39 : diagramme d'interaction de l'(-)-épigallocatechine gallate avec la Hsp90 (1YU6).

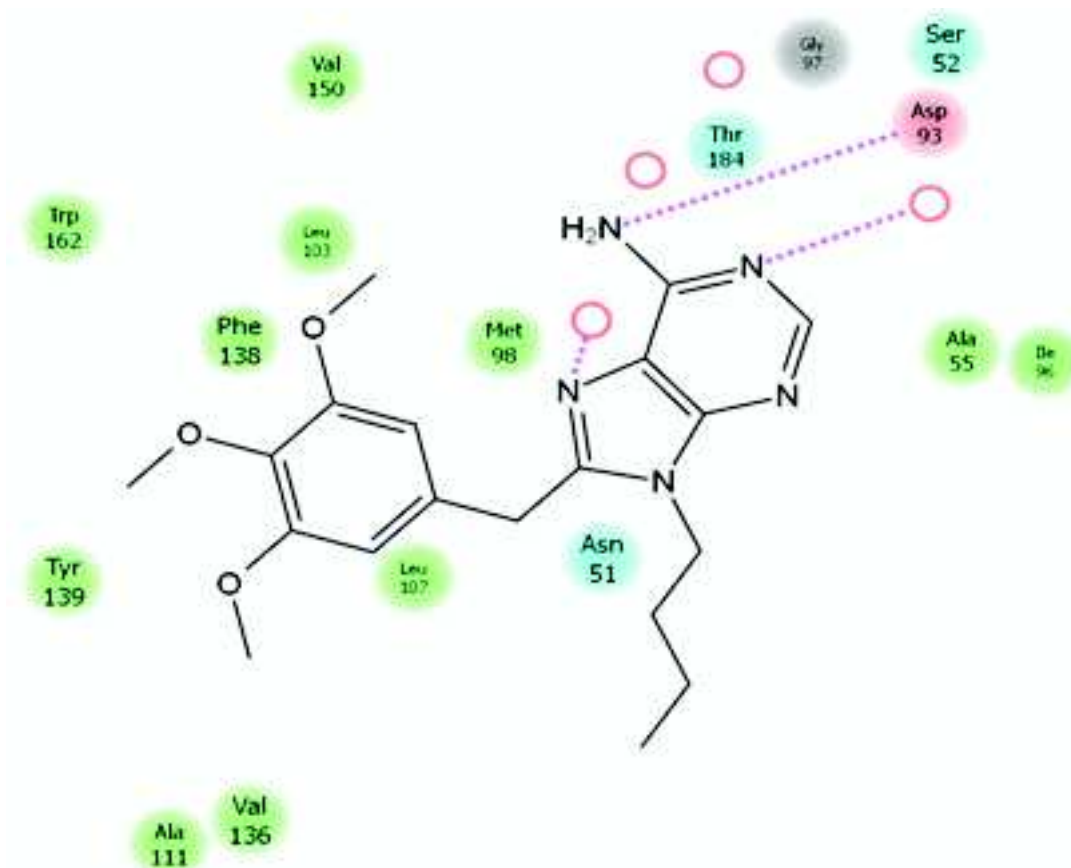


Figure 40 : diagramme d'interaction de PU3 (1UY6-ligand) avec la Hsp90 (1YU6).

L'(-)-épigallocatechine gallate et l'(-)-épicatéchine gallate ayant des GScores excellents : -18.198 et -17.667 (↑↑) respectivement (le meilleur score des ligands cocrystallisés étant celui de H55-ligand : -13.784) et de très bonnes MM-GBSA ΔG libre : -99.996 et -97.083 (↑) respectivement (la meilleure ΔG libre des ligands cocrystallisés étant celle de 2BSM- ligand : -104.310), pourraient avoir une activité potentielle très puissante contre la Hsp90. Elles ont un mode de liaison proche de 1UY6-ligand (PU3) Figure 38, cocrystallisé avec le récepteur avec lequel leurs meilleures poses ont été obtenues.

Elles forment par leur noyau phényle A une liaison hydrogène clé avec Asp93 et une autre avec HOH1; leur groupement hydroxyle en C5 remplace HOH4 et forme une liaison hydrogène avec HOH3 ; elles forment par leur noyau phényle B une liaison hydrogène avec la chaîne latérale de Asn51, et deux autres avec le squelette peptidique au niveau de Gly135 et Phe138 ; elles forment également par leur ester gallique une liaison hydrogène avec la chaîne latérale de Trp162 et une interaction de π - π stacking avec son noyau, et une liaison hydrogène avec le squelette peptidique au niveau de Leu103, tandis que PU3 (2YU6-ligand) forme seulement une liaison hydrogène clé avec Asp93, et deux autres avec HOH1 et HOH4.

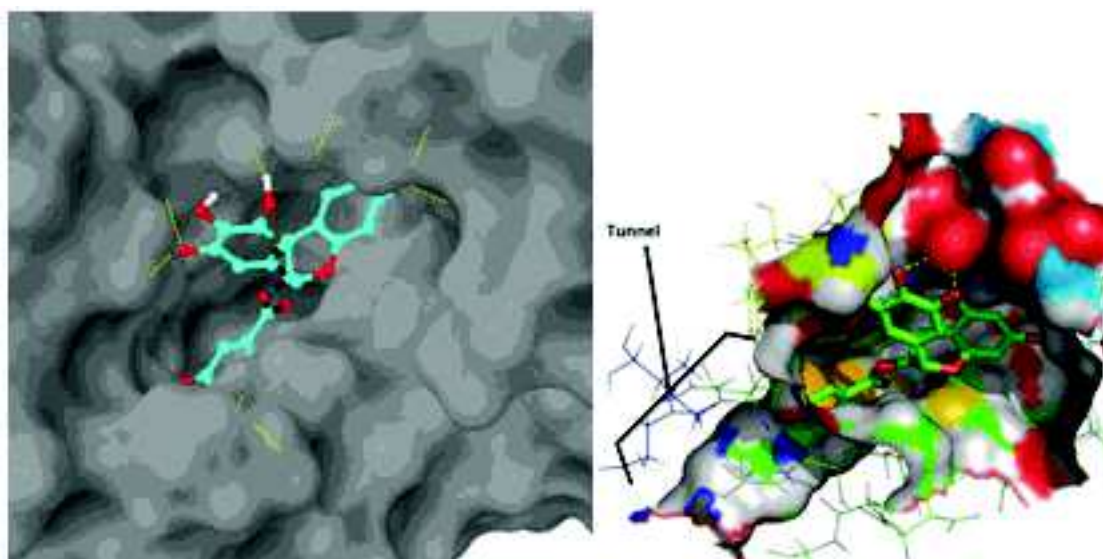


Figure 41 : emplacement de l'(-)-épigallocatechine dans le site actif de la Hsp90 (PDB : 1UY6) ; à noter la présence d'un tunnel (à gauche) dans lequel s'enfonce l'ester gallique de l'(-)-épigallocatechine.

L'ester gallique s'enfonce dans un tunnel lipophile Figure 41 qui n'est pas présent dans toutes les structures PDB de la Hsp90 (induced fit pour PU3), 1AMW, entre autres, ne présente pas ce tunnel, son entrée étant fermée par les chaînes latérales des résidus avoisinants Figure 42.

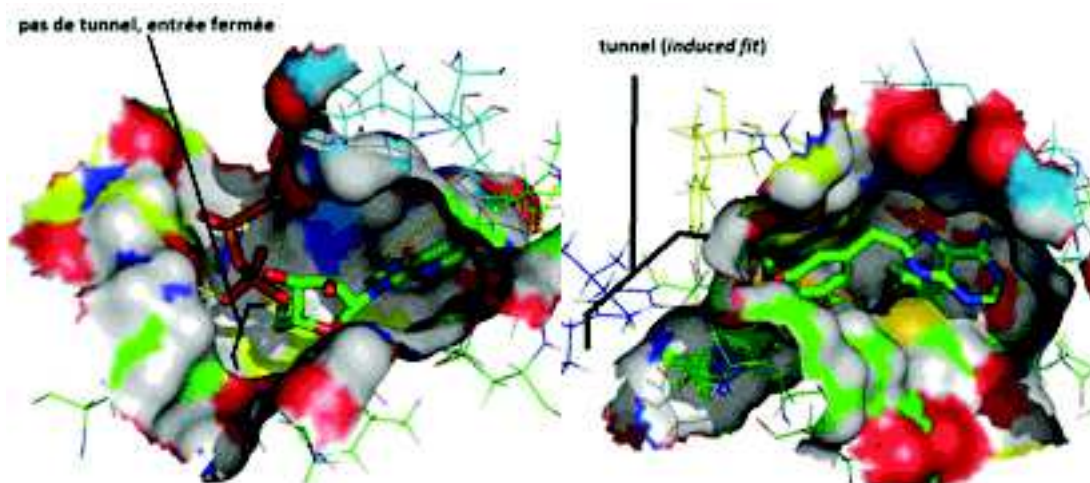


Figure 42 : à gauche 1AMW avec l'ADP, pas de tunnel ; à droite 1UY6 ave PU3, tunnel présent.

L'(-)-épigallocatechine gallate et l'(-)-épicatéchine gallate établissent des interactions lipophiles importantes à travers leur noyau A, leur hétérocycle C, et l'enfouissement de leur ester gallique dans le tunnel décrit ci-dessus avec les résidus hydrophobes avoisinants Figure 43.

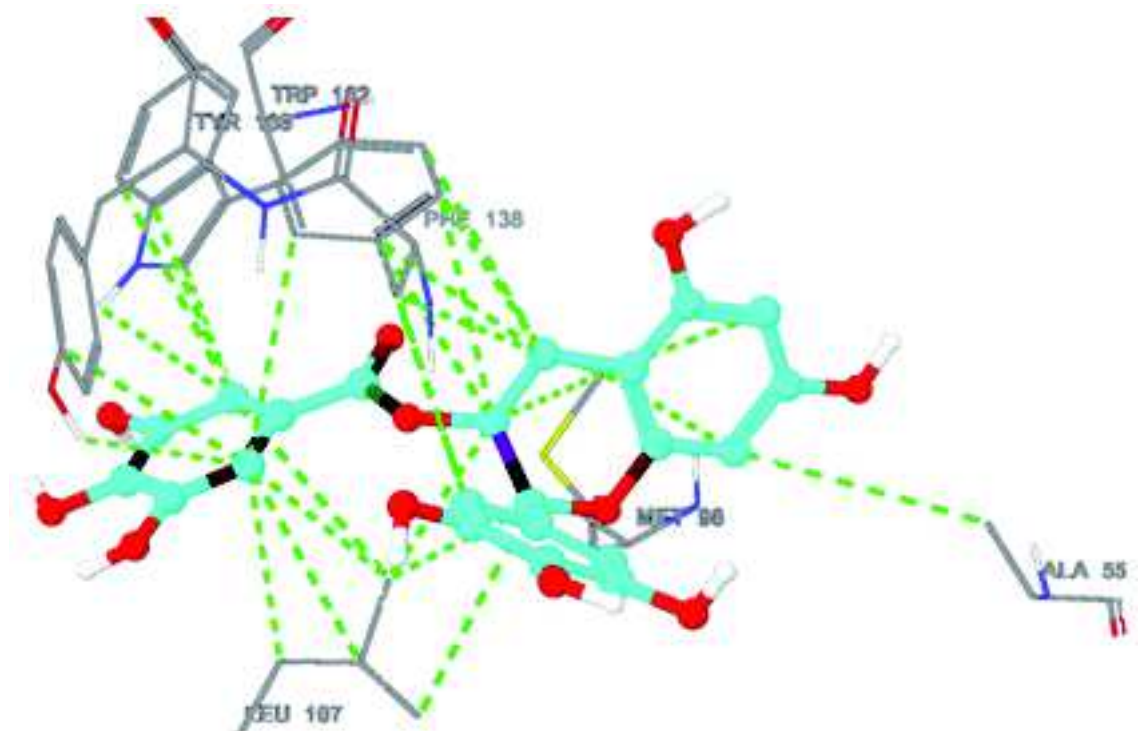


Figure 43 : contacts hydrophobes de l'(-)-épigallocatechine gallate dans le site actif de 1UY6.

Presque tous les hydroxyles de l'(-)-épigallocatechine gallate et l'(-)- épicatechine gallate sont placés dans des régions hydrophiles, et l'emplacement de certains d'entre eux est presque le même emplacement de certaines molécules d'eau cocrystallisées avec 1UY6.

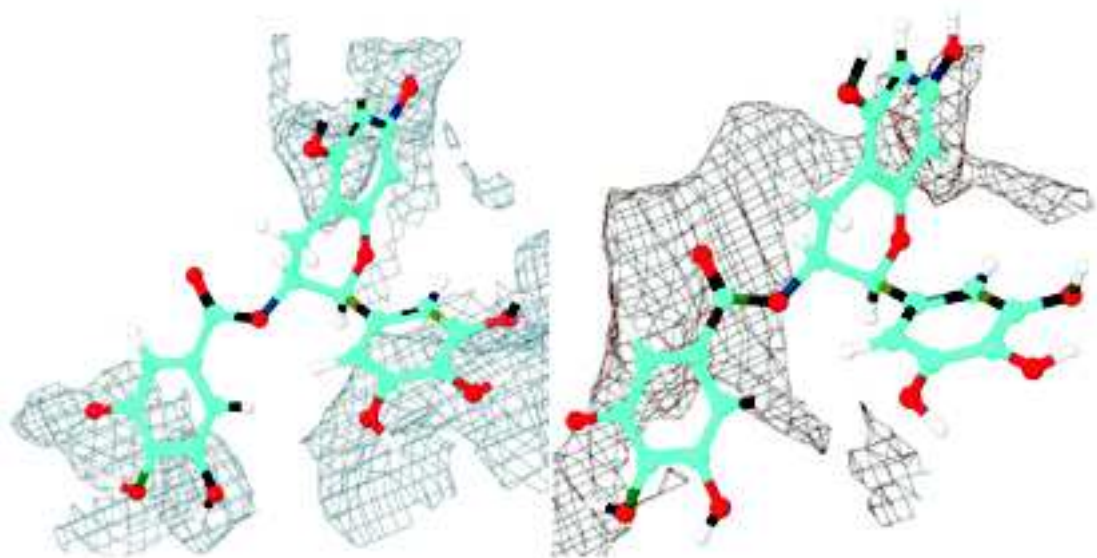


Figure 44 : à gauche, régions hydrophiles (mailles en turquoise) du site de liaison de 1UY6 ; à droite, les régions hydrophobes (mailles en rouge).

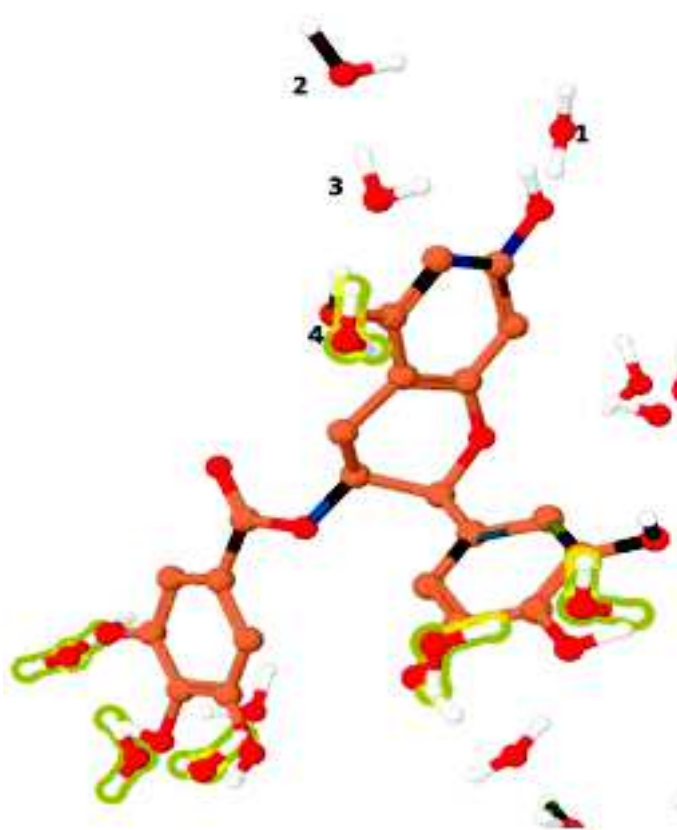


Figure 45 : emplacement de certains hydroxyyles de de l'(-)-
 épigallocatechine gallate superposés avec certaines molécules d'eau
 (en surbrillance) cocrystallisées dans le site de liaison de 1UY6.

Des évidences précliniques grandissantes suggèrent que l'(-)- épigallocatechine gallate (flavan-3-ol), la catéchine principale trouvée dans le thé vert, a un effet potentiel sur diverses maladies humaines. Elle fonctionne comme antioxydant puissant, empêchant les dommages oxydatifs dans les cellules saines, mais également comme agent antiangiogénique et antitumoral, et comme modulateur de la réponse des cellules tumorales à la chimiothérapie. Une grande partie des propriétés anticancéreuses du thé vert sont médiées par l'(-)-épigallocatechine gallate qui induit l'apoptose et favorise l'arrestation de la croissance cellulaire, et bloque la carcinogenèse en affectant une grande sélection de voies de transduction de signal. Diverses études cliniques ont indiqué que le traitement par l'(-)-épigallocatechine gallate empêche l'incidence et la prolifération des tumeurs dans différents sites, tels que le foie, l'estomac, la peau, le poumon, la glande mammaire et le colon. L'(-)-épigallocatechine gallate a un grand potentiel dans la prévention du cancer en raison de sa sûreté, son coût bas et sa biodisponibilité (Singh, Shankar and Srivastava, 2011).

Dans un premier temps, Palermo et al. (Palermo, Westlake and Gasiewicz, 2005) ont suggéré que l'(-)-épigallocatechine gallate inhibait la Hsp90 en se liant au domaine C-terminal dans un site proche du site ATPasique C-terminal. Après, Li et al. (Li et al., 2009) ont confirmé les suggestions précédentes et ont avancé que l'(-)-épigallocatechine gallate, en se liant au domaine C-terminal, n'interfère pas avec la liaison de l'ATP. Au contraire, Yin et al. (Yin, Henry and Gasiewicz, 2009) ont trouvé que l'(-)- épigallocatechine gallate se liait à au domaine C-terminal et à la Hsp90 d'une manière compétitive avec l'ATP. Tran et al. (Tran et al., 2010) ont trouvé que l'(-)-épigallocatechine gallate se liait à la Hsp90 et la Hsp70 d'une manière compétitive avec l'ATP et inhibait même leur expression.

Nos résultats suggèrent que l'(-)-épigallocatechine gallate se lie au domaine N-terminal d'une manière compétitive de la même façon dont il se lie au chaperon moléculaire Grp78 (glucose-regulated protein) (Ermakova et al., 2006).

3 Résultats du criblage de la grande chimiothèque de flavonoïdes et coumarines

name	GScore (XP)	QM-Polarized GScore (XP)	Prime MMGBSA ΔG bind	ensemble receptor	conformer field	#stars	#rtvFG	Rule Of Five
(-)-Epigallocatechin gallate	-18.195	-18.300	-100.082	UY6	OPLS	5	1	2
Gossypin	-18.021	-18.150	-100.857	UY6	MMFPS	5	1	2
Epicatechin 3-O-(4-O-methylgallate)	-17.609	-17.733	-103.252	UY6	OPLS	2	1	1
3,4,2',4'-Tetrahydroxy-2-geranyldihydrochalcone	-16.533	-16.771	-101.235	UY6	Original	1	0	0
Artonin E	-16.053	-16.177	-100.719	2H55	Original	0	0	0
Brousochalcone A	-15.885	-15.943	-95.0791	UY6	MMFPS	0	1	0
Bolusanthol C	-15.730	-15.846	-100.249	2H55	Original	2	0	1
Sophoraflavanone G	-14.931	-15.077	-89.946	UY6	Original	1	0	0
Bavachalcone	-14.706	-14.806	-85.683	2H55	Original	0	1	0
Sophoraisoflavanone D	-14.637	-14.688	-101.704	UY6	MMFPS	2	0	1
Heptyl-3-catechin	-14.521	-14.623	-96.460	UY6	MMFPS	0	0	0
Desmethyloxanthohumol	-14.438	-14.499	-87.367	UY6	OPLS	0	1	0
5,4'-Dimethoxy-3'-prenylbiochanin A	-14.190	-14.296	-92.003	UY6	Original	0	0	0
Morusin	-13.817	-13.930	-98.971	UY6	OPLS	1	0	0

Tableau 4 : résultats finals du criblage de la grande chimiothèque de flavonoïdes et de coumarines.

QM-Polarized GScore : le score de la réévaluation par le protocole **QM-Polarized Ligand Docking**.

Discussion

Le premier stage de docking en mode Glide SP a généré plus de 1000 poses (la meilleure pose pour chaque ligand ayant respecté la contrainte imposée et le récepteur avec lequel elle a été obtenue). Toute pose ayant un score > -8 a été exclue, et il n'a été retenu qu'un seul stéréo-isomère pour chaque ligand (ayant le meilleur score), ce qui a réduit le nombre des poses à passer au deuxième stage à 571.

Le deuxième stage en mode Glide XP a généré 571 poses qui ont été évaluées minutieusement (pour les critères de l'évaluation voir : V, 2

Résultats du criblage virtuel de la petite chimiothèque de flavonoïdes), et 14 molécules, dont les structures sont présentées Figure 46, ont été sélectionnées et réévaluées par le protocole : QM-Polarized Ligand Docking (tableau ci-dessus)

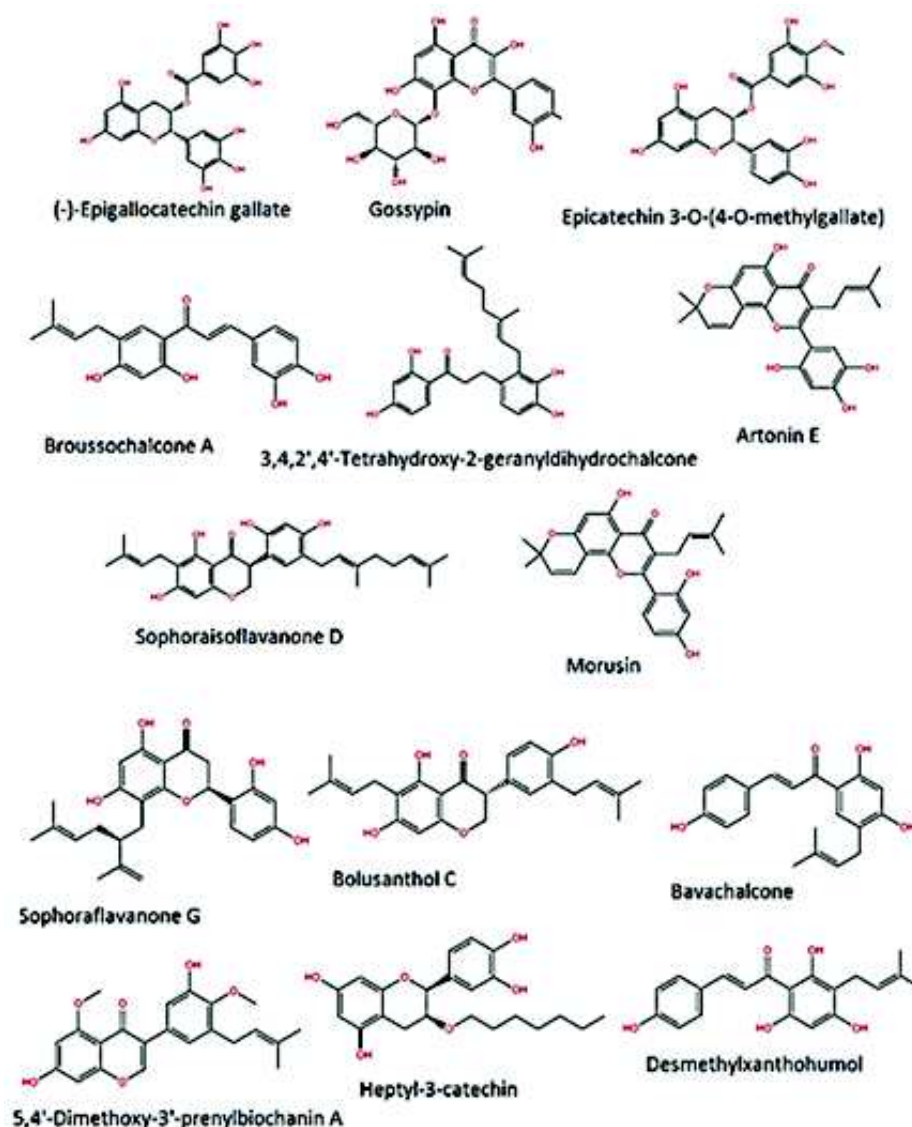


Figure 46: structures des molécules des résultats finals du criblage de la grande chimiothèque.

Le meilleur score est celui de 1'(-)-épigallocatéchine gallate (QM Polarized GScore (XP) : -18.300 (↑↑), MMGBSA DG bind: -100.082 (↑↑)), tout comme dans le criblage de la petite chimiothèque et avec exactement le même mode de liaison, ce qui prouve la reproductibilité des résultats de Glide et la pertinence du protocole de criblage.

Le dérivé méthylé épicatechine 3-O-(4-O-methylgallate) avec un score très légèrement inférieur (QM-Polarized GScore (XP) : -17.733 (↑↑), MMGBSA DG bind : -103.252 (↑)) et exactement le même mode de liaison de l'(-)-épigallocatéchine gallate, a l'avantage d'avoir le nombre de #stars : 2 et seulement 1 violation de la règle de Lipinski par rapport à 5 et 2, respectivement, pour l'(-)-épigallocatéchine gallate. L'heptyl-3- catéchine avec un bon QM-Polarized GScore (XP) : -14.623 et une bonne MMGBSA DG bind : -96.460, a un profil pharmaceutique encore meilleur, avec un nombre de #stars : 0 et 0 violations de la règle des cinq. Elle a presque le même mode liaison que les deux précédents, sauf qu'à la place de l'ester gallique, elle a une chaîne aliphatique qui s'enfonce dans le tunnel hydrophobe où elle établit un réseau de contacts lipophiles Figure 47.

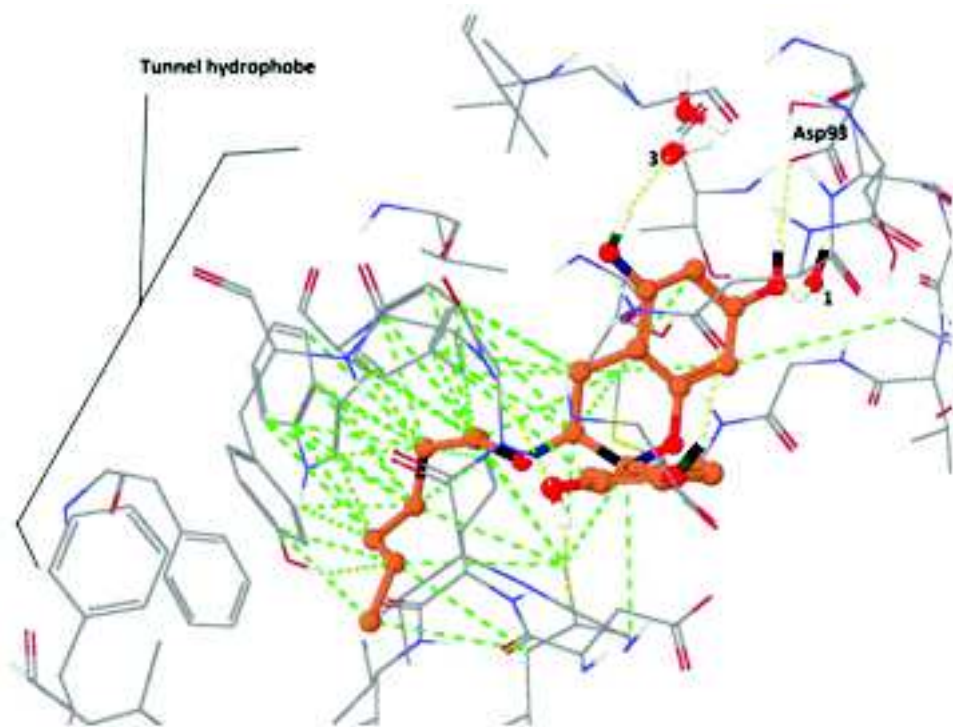


Figure 47 : mode de liaison et contacts hydrophobes de l'heptyl-3-catéchine (PDB : 1UY6).

La gossypine (gossypetine-8-glucoside), un glucoside du flavonol gossypetine (8-hydroxyquercetine), ayant un excellent QM-Polarized GScore (XP) : -18.150 (↑↑) et une très bonne MM-GBSA ΔG : -100.857 (↑), pourrait avoir une puissante activité sur la Hsp90.

Son désavantage est qu'elle est glucoside, les flavonoïdes glycosidiques étant déglycosilés lors de leur absorption.

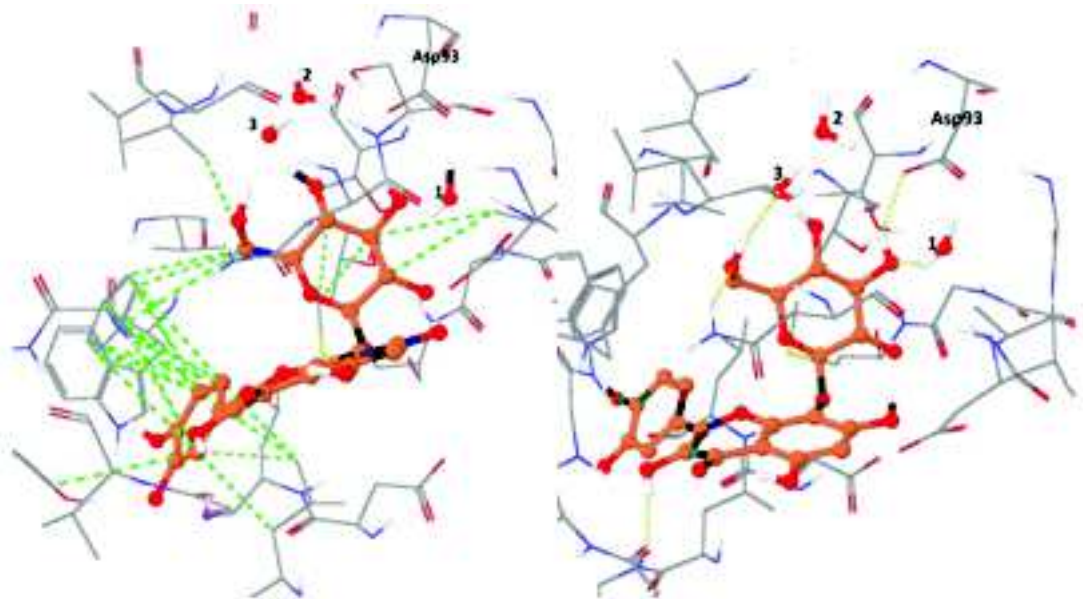


Figure 48 : à droite, mode de liaison de la gossypine; à gauche, ses contacts hydrophobes (PDB: UY6).

La 3,4,2',4'-tetrahydroxy-2-géranyldihydrochalcone, qui est une dihydrochalcone, a un très bon QM-Polarized GScore (XP) : -16.771 (↑), une très bonne MMGBSA ΔG bind : -101.235 (↑), et un meilleur profil, avec un nombre de #stars : 1 et 0 violations de la règle des cinq. Elle est stabilisée par un grand réseau de contacts lipophiles Figure 49.

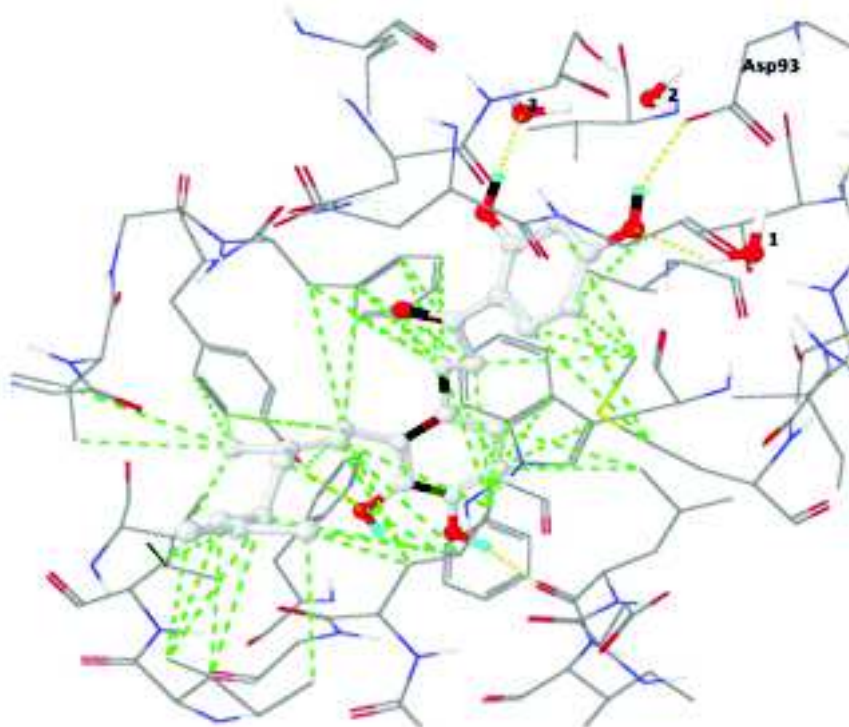


Figure 49 : mode de liaison et contacts hydrophobes de la 3,4,2',4'-tetrahydroxy-2-géranyldihydrochalcone.

L'artonnine E (5'-hydroxymorusine), un flavone, a un très bon QM- Polarized GScore (XP) : -16.177 (↑), une très bonne MMGBSA ΔG bind : -100.719 (↑), et un meilleur profil pharmaceutique, avec un nombre de #stars : 0 et 0 violations de la règle des cinq. Elle a un mode de liaison similaire à celui de 2H55-ligand (PU-DZ8) cocrystallisé avec le récepteur avec lequel sa meilleure pose a été obtenue Figure 50.

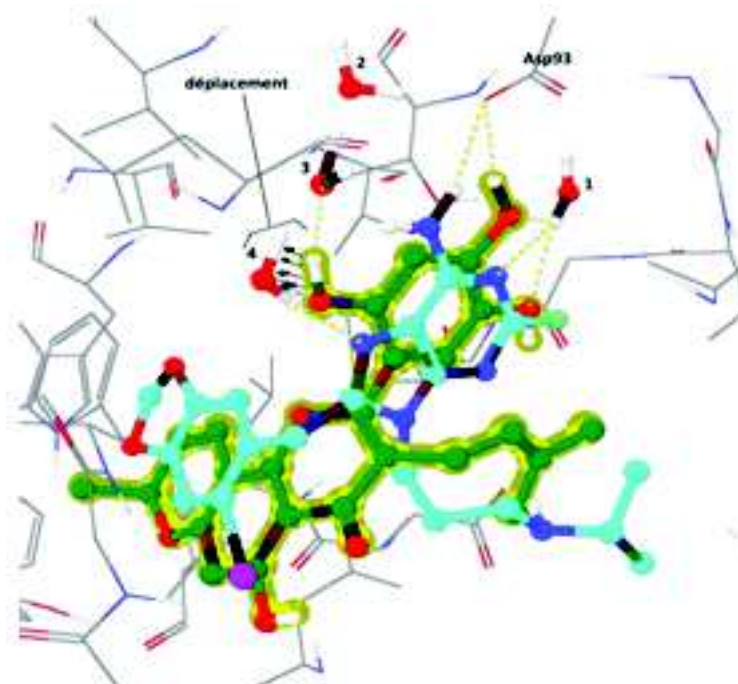


Figure 50 : superposition de l'artinine E (en surbrillance) et de PU-DZ8 (2H55-ligand) dans 2H55.

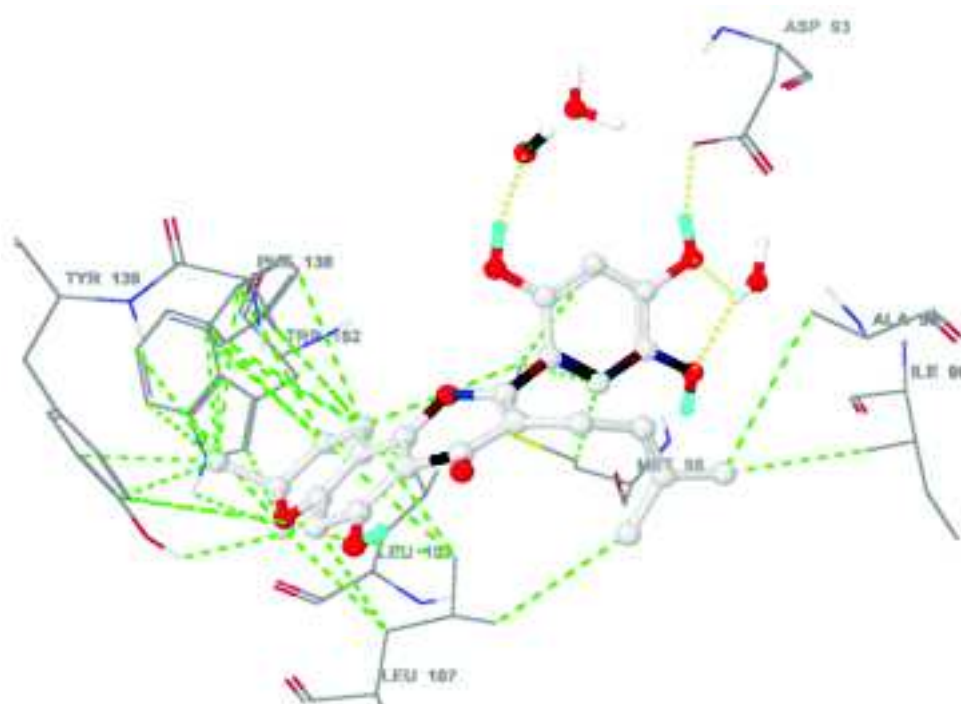


Figure 51 : contacts hydrophobes de l'artinine E dans le site actif de 2H55.

Elle forme par son noyau phényle B une liaison hydrogène clé avec Asp93 et deux autres avec HOH1; son groupement hydroxyle en C2' remplace HOH4 et forme une liaison hydrogène avec HOH3. Son noyau phényle A et son prényle se placent à l'entrée du tunnel hydrophobe où ils forment un réseau de contacts lipophiles Figure 51.

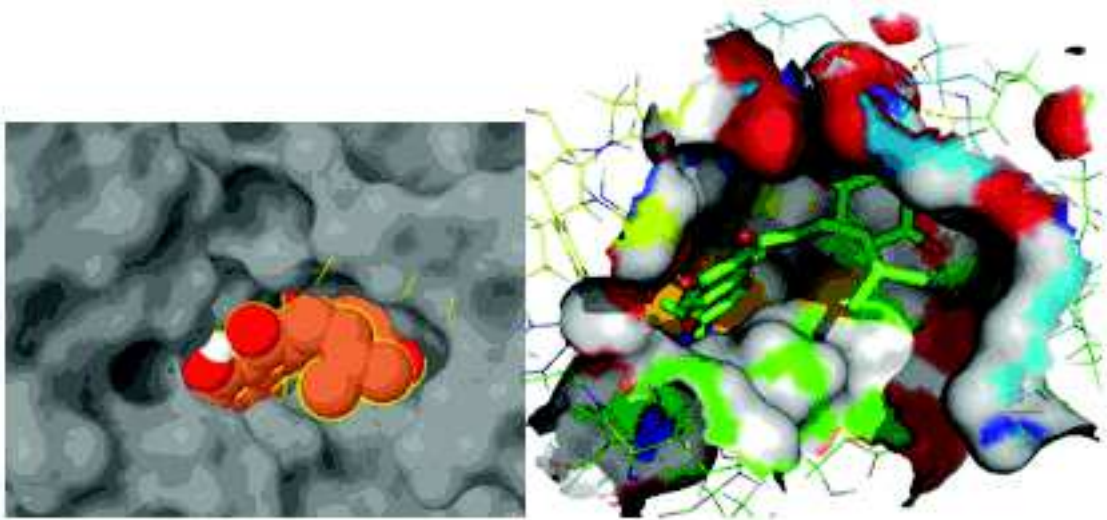


Figure 52 : l'artonine E dans le site actif de 2H55.

L'artonine E présente une bonne complémentarité structurelle avec le site actif de 2H55 Figure 52.

Toutes les autres molécules ont des modes de liaison plus ou moins similaires aux molécules précédentes, avec un noyau phényle assurant la liaison hydrogène clé avec Asp93, et un groupement hydroxyle déplaçant HOH4, et des cycles ou des chaînes s'enfonçant plus ou moins profondément dans le tunnel hydrophobe, excepté deux molécules qui ont, à la place du groupement hydroxyle remplaçant HOH4, un oxygène accepteur, à savoir : le bolusanthol C Figure 53 et la 5,4'-diméthoxy-3'-prénylbiochanine A Figure 54. Ces deux molécules ne déplacent pas HOH4, mais plutôt en acceptent une liaison hydrogène comme l'ADP, PU3, PU- DZ8 et autres ligands. Ces deux molécules ont été réamarrées dans leurs sites actif respectifs en présence de HOH4 (le premier docking ayant été réalisé en absence de HOH4), leurs scores présentés Tableau 4 en sont issus.

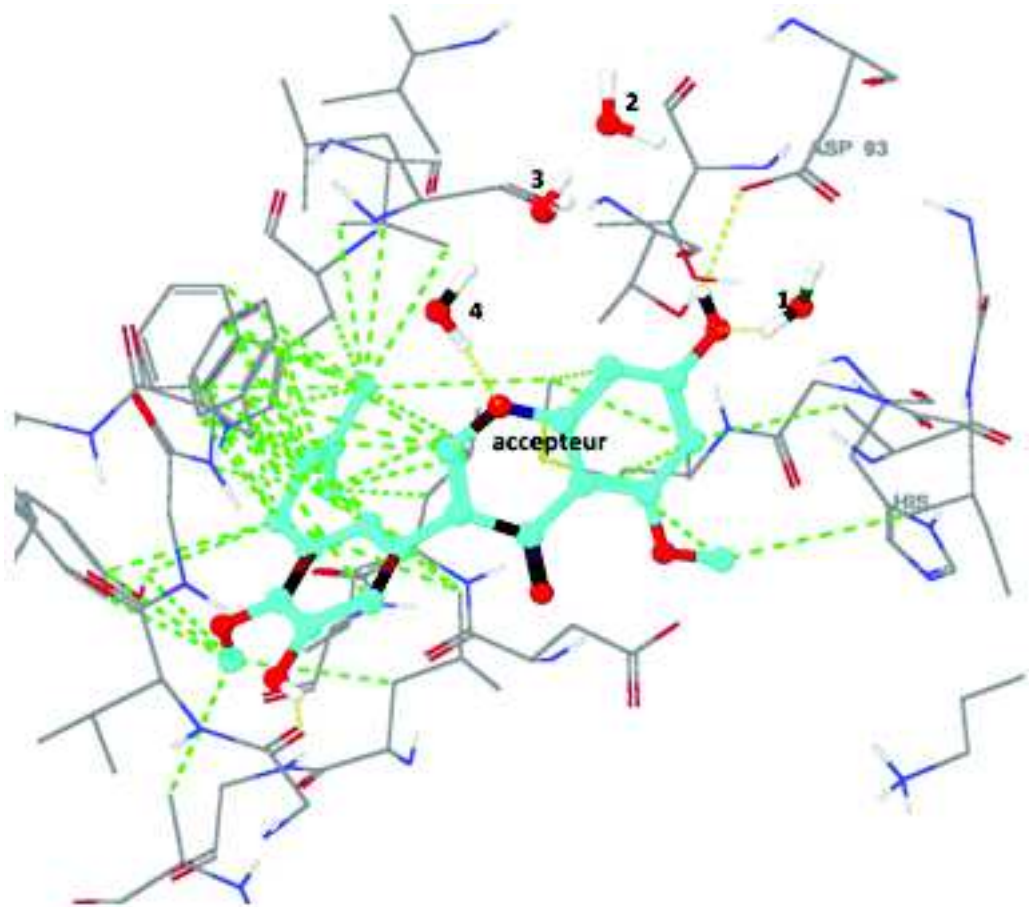


Figure 53 : mode de liaison de la 5,4'-diméthoxy-3'-prénylbiochanine A avec ses contacts hydrophobes.

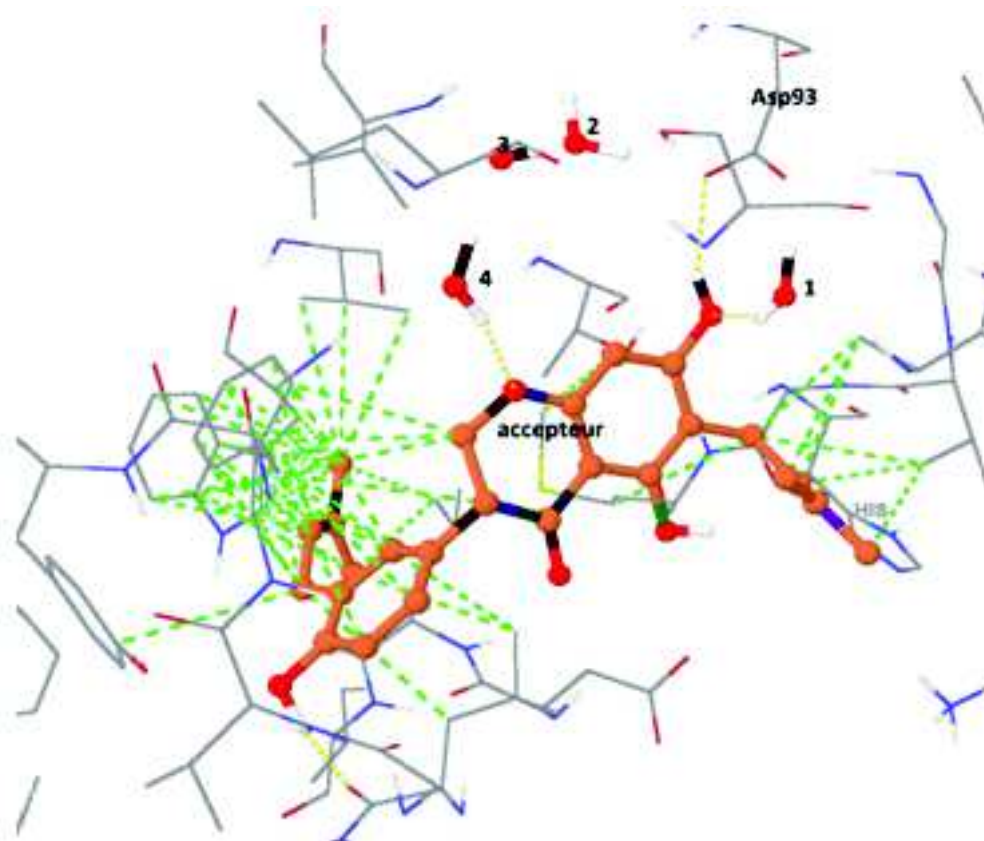


Figure 54 : mode de liaison du le bolusanthol C avec ses contacts hydrophobes.

Conclusions

L'objectif principal de cette étude était d'évaluer *in silico* l'effet inhibiteur potentiel des flavonoïdes sur la Hsp90 comme mécanisme moléculaire anticancéreux. Cette évaluation a été menée au moyen de l'élaboration d'un protocole de criblage virtuel basé sur la structure de la cible (structure-based virtual screening) ; ce protocole comprend plusieurs enrichissements pertinents mis en place suite à une étude approfondie de la cible : Hsp90, qui se sont révélés efficaces, du moins pour les structures Hsp90 étudiées, à reproduire les modes de liaison expérimentaux avec fiabilité lors du test du protocole.

Ce protocole de criblage virtuel, après avoir été testé, a été appliqué d'une manière poussée sur une petite chimiothèque de 25 flavonoïdes des plus abondants dans l'alimentation, et suite à une évaluation multicritères fine des résultats, trois molécules de la sous classe des flavan-3-ols se sont révélées actives *in silico* contre la Hsp90, dont deux sont très prometteuses, à savoir : l'(-)-épigallocatechine gallate et l'(-)- épicatechine gallate ayant des GScores excellents : -18.198 et -17.667 (↑↑) respectivement (le meilleur score des ligands cocrystallisés étant celui de H55-ligand : -13.784) et de très bonnes Prime MM-GBSA ΔG libre : -99.996 et -97.083 (↑) respectivement (la meilleure ΔG libre des ligands cocrystallisés étant celle de 2BSM-ligand : -104.310).

Le protocole de criblage virtuel a été appliqué en deuxième lieu avec une stratégie un peu différente (implication de deux stages de docking) sur une chimiothèque de flavonoïdes et de coumarines relativement plus grande, comprenant 1050 composés préfiltrés selon des critères pharmaceutiques à partir de plus de 5700 structures, et suite à une évaluation multicritères fine des résultats, 14 molécules de différentes sous-classes de flavonoïdes se

sont révélées plus ou moins actives contre la Hsp90 in silico, dont 4 sont très prometteuses, à savoir : l'(-)- épigallocatechine gallate (QM-Polarized GScore (XP) : -18.300 (↑↑), MMGBSA DG bind : -100.082 (↑)), l'épicatéchine 3-O-(4-O- methylgallate) (QM-Polarized GScore (XP) : -17.733 (↑↑), MMGBSA DG bind : -103.252 (↑)), la 3,4,2',4'-Tetrahydroxy-2-geranyldihydrochalcone (QM-Polarized GScore (XP) : -16.771 (↑), MMGBSA ΔG bind : -101.235 (↑)) , et l'artanine E (QM-Polarized

GScore (XP) : -16.177 (↑), MMGBSA ΔG bind : -100.719 (↑)).

Autant que nous sachions, hormis l'(-)-épigallocatechine gallate dont l'effet inhibiteur sur le site N-terminal ATPasique de la Hsp90 est controversé, les molécules jugées actives in silico dans cette étude n'ont jamais été évoquées dans la littérature comme inhibiteurs du site ATPasique N-terminal de la Hsp90, et nous serions les premiers à les découvrir, et si elles s'avèrent tels in vitro, ce serait là un nouveau mécanisme moléculaire découvert de l'effet anticancéreux des flavonoïdes.

Bibliographie

- Abagyan, R. and Totrov, M. (2001) 'Highthroughput docking and lead generation.', *Curr. Opin. Chem. Biol.*, vol. 5, pp. 375-382.
- Abola, E., Kuhn, P., Earnst, T. and Stevens, R.C. (2000) 'Automation of X-ray crystallography.', *Nat. Struct. Biol.*, vol. 7, pp. 973-977.
- Albanese, V., Yam, A.Y., Baughman, J., Parnot, C. and Frydman, J. (2006) 'Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells.', *Cell*, vol. 124, pp. 75-88.
- Aligue, R., Akhavan-Niak, H. and Russell, P. (1994) 'A role for Hsp90 in cell cycle control: Wee1 tyrosine kinase activity requires interaction with Hsp90.', *Embo J*, vol. 13, pp. 6099- 6106.
- Ali, M.M.U., Roe, S.M., Vaughan, C.K., Meyer, P., Panaretou, B., Piper, P.W., Prodromou, C. and Pearl, L.H. (2006) 'Crystal structure of an Hsp90- nucleotide-p23/Sba1 closed chaperone complex', *Nature*, vol. 440, pp. 1013-1017.
- Allan, R.K., Mok, D., Ward, B.K. and Ratajczak, T. (2006) 'Modulation of chaperone function and cochaperone interaction by novobiocin in the C-terminal domain of Hsp90: evidence that coumarin antibiotics disrupt Hsp90 dimerization,', *J. Biol. Chem.*, vol. 281, pp. 7161- 7171.
- Alvarez, J.C. (2004) 'High-throughput docking as a source of novel drug leads.', *Curr. Opin. Chem. Biol.*, vol. 8, no. 4, pp. 365-370.
- Anderson, J.J., Anthony, M.S., Cline, J.M., Washburn, S.A. and Garner, S.C. (1999) 'Health potential of soy isoflavones for menopausal women', *Public Health Nutr*, vol. 2, pp. 489-504.
- Aron, P.M. and Kennedy, J.A. (2007) 'Compositional investigation of phenolic polymers isolated from *Vitis vinifera* L. cv. Pinot Noir during fermentation.', *J Agric Food Chem*, vol. 55, pp. 5670-5680.
- Arrigo, A.P., Viot, S., Chaufour, S., Firdaus, W., Kretz-Remy, C. and Diaz-Latoud, C. (2005) 'Hsp27 consolidates intracellular redox homeostasis by upholding glutathione in its reduced form and by decreasing iron intracellular levels.', *Antioxid Redox Signal*, vol. 7, no. 3-4, pp. 414-422.
- Bardwell, J.C. and Craig, E.A. (1987) 'Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*.', *Proc Natl Acad Sci USA*, vol. 84, pp. 5177-5181.
- Barluenga, S., Fontaine, J.G., Wang, C., Aouadi, K., Chen, R., Beebe, K., Neckers, L.M. and Winssinger, N. (2009) 'Inhibition of HSP90 with pochoximes: SAR and structure-based insights.', *Chembiochem*, vol. 10, pp. 2753-2759.
- Barril, X., Brough, P., Drysdale, M., Hubbard, R.E., Massey, A., Surgenor, A. and Wright, L. (2005) 'Structure-based discovery of a new class of Hsp90 inhibitors', *Bioorganic & Medicinal Chemistry Letters*, vol. 15, pp. 5187-5191.

- Barril, X. and Morley, S.D. (2005) 'Unveiling the full potential of flexible receptor docking using multiple crystallographic structures.', *J. Med. Chem.*, vol. 48, no. 13, pp. 4432-4443.
- Batra, P. and Sharma, A.K. (2013) 'Anti-cancer potential of flavonoids: recent trends and future perspectives', *Biotech*, vol. 3, pp. 1-21.
- Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A. and Forterre, P. (1997) 'An atypical topoisomerase II from Archaea with implications for meiotic recombination.', *Nature*, vol. 386, pp. 414-417.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) 'The Protein Data Bank.', *Nucleic Acids Res.*, vol. 28, pp. 235-242.
- Biamonte, M.A., Van de Water, R., Arndt, J.W., Scannevin, R.H., Perret, D. and Lee, W.C.J. (2010) 'Heat shock protein 90: inhibitors in clinical trials.', *Med. Chem.*, vol. 53, pp. 3-17.
- Binder, R.J., Kelly, J.B., Vatner, R.E. and Srivastava, P.K. (2007) 'Specific immunogenicity of heat shock protein Gp96 derives from chaperoned antigenic peptides and not from contaminating proteins.', *J Immunol*, vol. 179, no. 11, pp. 7254-7261.
- Birt, D.F., Hendrich, S. and Wang, W. (2001) 'Dietary agents in cancer prevention: flavonoids and isoflavonoids', *Pharmacology & Therapeutics*, vol. 90, pp. 157-177.
- Bleicher, K.H., Bohm, H.J., Muller, K. and Alanine, A.I. (2003) 'Hit and lead generation: beyond high-throughput screening.', *Nat. Rev. Drug Discovery*, vol. 5, no. 2, pp. 369-378.
- Bosetti, C., Rossi, M., McLaughlin, J.K., Negri, E., Talamini, R., Lagioui, P. and al. (2007) 'Flavonoids and the risk of renal cell carcinoma.', *Cancer Epidemiol Biomarkers Prev*, vol. 16, no. 1, pp. 98-101.
- Brooijmans, N. and Kuntz, I.D. (2003) 'Molecular recognition and docking algorithms.', *Annu. Rev. Biophys. Biomol. Struct.*, vol. 32, pp. 335-373.
- Brooijmans, N. and Kuntz, I.D. (2003) 'Molecular recognition and docking algorithms.', *Annu. Rev. Biophys. Biomol. Struct.*, vol. 32, pp. 335-373.
- Brough, P.A., Aherne, W., Barril, X., Borgognoni, J., Boxall, K., Cansfield, J.E., Cheung, K.M., Collins, I., Davies, N.G., Drysdale, M.J. and al. (2007) '4,5-Diarylisoazole Hsp90 chaperone inhibitors: potential therapeutic agents for the treatment of cancer.', *J. Med. Chem.*, vol. 51, pp. 196-218.
- Burel, C., Mezger, V., Pinto, M., Rallu, M., Trigon, S. and Morange, M. (1992) 'Mammalian heat shock protein families. Expression and functions.', *Experientia*, vol. 48, pp. 629-634.
- Burger, A.M., Fiebig, H., Stinson, S. and Sausville, E. (2004) '17-DMAG activity in human melanoma models.', *Anticancer Drugs*, vol. 15, pp. 377-387.
- Bush, K., Goldberg, A. and Nigam, S. (1997) 'Proteasome inhibition leads to a heat-shock response, induction of endoplasmic reticulum chaperones, and thermotolerance.', *J Biol Chem*, vol. 272, no. 24, pp. 9086-9092.

- Cao, J., Zhang, Y., Chen, W. and Zhao, X. (2010) 'The relationship between fasting plasma concentrations of selected flavonoids and their ordinary dietary intake', *Br J Nutr*, vol. 103, pp. 249-255.
- Cavasotto, C.N. and Phatak, S.S. (2009) 'Homology modeling in drug discovery: current trends and applications.', *Drug Discov Today*, vol. 14, no. 13-14, pp. 676-683.
- Chene, P. (2002) 'ATPases as drug targets: learning from their structure.', *Nat. Rev. Drug Discov.*, vol. 9, no. 1, pp. 665-673.
- Cheng, T., Li, X., Li, Y., Liu, Z. and Wang, R. (2009) 'Comparative assessment of scoring functions on a diverse test set.', *J Chem Inf Model.*, vol. 49, no. 4, pp. 1079-1093.
- Chen, B., Piel, W., Gui, L., Bruford, E. and Monteiro, A. (2005) 'The HSP90 family of genes in the human genome: insights into their divergence and evolution.', *Genomics*, vol. 86, no. 6, pp. 627-637.
- Cheung, K.M., Matthews, T.P., James, K., Rowlands, M.G., Boxall, K.J., Sharp, S.Y., Maloney, A., Roe, S.M., Prodromou, C., Pearl, L.H. and al. (2005) 'The identification, synthesis, protein crystal structure and in vitro biochemical evaluation of a new 3,4-diarylpyrazole class of Hsp90 inhibitors.', *Bioorg. Med. Chem. Lett.*, vol. 15, pp. 3338-3343.
- Chiosis, G. and Neckers, L. (2006) 'Tumor selectivity of Hsp90 inhibitors: the explanation remains elusive.', *ACS Chem. Biol.*, vol. 1, pp. 279-284.
- Chiosis, G., Timaul, M.N., Lucas, B., Munster, P.N., Zheng, F.F., Sepp-Lorenzino, L. and Rosen, N. (2001) 'A small molecule designed to bind to the adenine nucleotide pocket of Hsp90 causes Her2 degradation and the growth arrest and differentiation of breast cancer cells.', *Chem. Biol.*, vol. 8, pp. 289-299.
- Claussen, H., Buning, C., Rarey, M. and Lengauer, T. (2001) 'FlexE: efficient molecular docking considering protein structure variations.', *J. Mol. Biol.*, vol. 308, no. 2, pp. 377-395.
- Cole, J.C., Murray, C.W., Nissink, J.W., Taylor, R.D. and Taylor, R. (2005) 'Comparing protein-ligand docking programs is difficult.', *Proteins: Struct., Funct., Bioinf.*, vol. 60, no. 3, pp. 325-332.
- Cozzini, P., Kellogg, G.E., Spyraakis, F. and al. (2008) 'Target flexibility: an emerging consideration in drug discovery and design.', *J Med Chem.*, vol. 51, no. 20, pp. 6237-6255.
- Cross, J.B., Thompson, D.C., Rai, B.K. and al. (2009) 'Comparison of several molecular docking programs: pose prediction and virtual screening accuracy.', *J Chem Inf Model.*, vol. 49, no. 6, pp. 1455-1474.
- Croteau, R., Kutchan, T. and Lewis, N. (2000) 'Natural products (secondary metabolites).', in Buchanan, B., Grisse, W. and Jones, R. (ed.) *Biochemistry and Molecular Biology of Plants.*, American Society of Plant Physiologists.
- Crozier, A., Ashihara, H. and Clifford, M. (ed.) (2006) *Plant Secondary Metabolites and the Human Diet.*, Oxford: Blackwell Publishing.
- Crozier, A., Jaganath, I. and Clifford, M. (2006) 'Phenols, polyphenols and tannins: An overview.', in Crozier, A., Ashihara, H. and Clifford, M. (ed.) *Plant Secondary Metabolites and the Human Diet.*, Oxford: Blackwell Publishing.

- Czar, M.J., Welsh, M.J. and Pratt, W.B. (1996) 'Immunofluorescence localization of the 90- kDa heat shock protein to cytoskeleton.', *European journal of cell biology*, vol. 70, pp. 322-330.
- Davies, N.G.M., Browne, H., Davis, B., Drysdale, M.J., Foloppe, N., Geoffrey, S., Gibbons, B., Hart, T., Hubbard, R., Jensen, M., Mansell, H., Massey, A., Matassova, N. and al. (2012) 'Targeting conserved water molecules: Design of 4-aryl-5-cyanopyrrolo[2,3-d]pyrimidine Hsp90 inhibitors using fragment-based screening and structure-based optimization', *Bioorg. Med. Chem.*, vol. 20, pp. 6770-6789.
- Delmotte, P. and Delmotte-Plaquet, J. (1953) 'A new antifungal substance of fungal origin.', *Nature*, vol. 171, p. 344.
- Dewick, P. (2002) *Medicinal Natural Products: A Biosynthetic Approach*, 2nd edition, Chichester: Wiley.
- Dixon, R.A. and Steele, C.L. (1999) 'Flavonoids and isoflavonoids—a goldmine for metabolic engineering.', *Trends Plant Sci*, vol. 4, pp. 394-400.
- Dymock, B.W., Barril, X., Brough, P.A., Cansfield, J.E., Massey, A., McDonald, E., Hubbard, R.E., Surgenor, A., Roughley, S.D., Webb, P. and al. (2005) 'Novel, potent small-molecule inhibitors of the molecular chaperone Hsp90 discovered through structure-based design.', *J. Med. Chem.*, vol. 48, pp. 4212-4215.
- Edwards, A.M., Arrowsmith, C.H., Christendat, D., Dharamsi, A., Friensen, J.D., Greenblat, J.F. and Vedadi, M. (2000) 'Protein production: feeding the crystallographers and NMR spectroscopists.', *Nat. Struct. Biol.*, vol. 7, pp. 970-972.
- Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P. (1997) 'Journal of Computer-Aided Molecular Design', Empirical scoring functions.1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes., vol. 11, no. 5, pp. 425-445.
- Erickson, J.A., Jalaie, M., Robertson, D.H., Lewis, R.A. and Vieth, M. (2004) 'Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy.', *J. Med. Chem.*, vol. 47, no. 1, pp. 45-55.
- Ermakova, S.P., Kang, B.S., Choi, B.Y., Choi, H.S., Schuster, T.F., Ma, W.Y., Bode, A.M. and Dong, Z. (2006) '(-)-Epigallocatechin gallate overcomes resistance to etoposide-induced cell death by targeting the molecular chaperone glucose-regulated protein 78.', *Cancer Res*, vol.66, pp. 9260-9269.
- Eustace, B., Sakurai, T., Stewart, J., Yimlamai, D., Unger, C., Zehetmeier, C., Lain, B., Torella, C., Henning, S., Beste, G. and al. (2004) 'Functional proteomic screens reveal an essential extracellular role for hsp90D in cancer cell invasiveness.', *NCB*, vol. 6, pp. 507-514.
- Feher, M. (2006) 'Consensus scoring for protein-ligand interactions.', *Drug Discov. Today*, vol. 11, no. 9-10, pp. 421- 428.
- Felts, S., Owen, B., Nguyen, P., Trepel, J., Donner, D. and Toft, D. (2000) 'The hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties.', *J Biol Chem*, vol. 275, no. 5, pp. 3305-3312.
- Ferrara, P., Gohlke, H., Price, D.J., Klebe, G. and Brooks III, C.L. (2004) 'Assessing scoring functions for protein-ligand interactions.', *J. Med. Chem.*, vol. 47, pp. 3032-3047.

- Ferrarini, M., Heltai, S., Zocchi, M. and Rugarli, C. (1992) 'Unusual expression and localization of heat-shock proteins in human tumor cells.', *Int J Cancer*, vol. 51, pp. 613-619.
- Ferrari, A.M., Wei, B.Q., Costantino, L. and Shoichet, B.K. (2004) 'Soft docking and multiple receptor conformations in virtual screening.', *J. Med. Chem.*, vol. 47, no. 21, pp. 5076-5084.
- Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P. and Shenkin, P.S. (2004) 'Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.', *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1739-1749.
- Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. and Mainz, D.T. (2006) 'Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.', *Journal of Medicinal Chemistry*, vol. 49, no. 21, pp. 6177-6196.
- Ghosh, S., Nie, A., An, J. and Huang, Z. (2006) 'Structure-based virtual screening of chemical libraries for drug discovery.', *Curr. Opin. Chem. Biol.*, vol. 10, no. 3, pp. 194-202.
- Gleeson, M.P., Hersey, A. and Hannongbua, S. (2011) 'In-silico ADME models: a general assessment of their utility in drug discovery applications.', *Curr Top Med Chem.*, vol. 11, no.4, pp. 358-381.
- Gohlke, H. and Klebe, G. (2002) 'Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors.', *Angew. Chem., Int. Ed. Engl.*, vol. 41, no. 15, pp. 2644-2676.
- Graham, T. (1991) 'Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates.', *Plant Physiol*, vol. 95, pp. 594-603.
- Grammatikakis, N., Vultur, A., Ramana, C., Sigano, A., Schweinfest, C., Watson, D. and Raptis, L. (2002) 'The role of Hsp90N, a new member of the Hsp90 family, in signal transduction and neoplastic transformation.', *J Biol Chem*, vol. 277, pp. 8312-8320.
- Gu, L., Kelm, M.A., Hammerstone, J.F., Beecher, G., Holden, J., Haytowitz, D., Gebhardt, S. and Prior, R.I. (2004) 'Concentration of proanthocyanidins in common foods and estimates of normal consumption.', *J Nutr*, vol. 134, pp. 613-617.
- Gupta, R. (1995) 'Phylogenetic analysis of the 90 kD heat shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species.', *Mol Biol Evol*, vol. 12, no. 6, pp. 1063-1073.
- Halgren, T.A. (1996) *J. Comput. Chem.*, vol. 17, no. 5.
- Halgren, T.A., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T. and Banks, J. (2004) 'Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.', *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1750-1759.
- Hann, M.M. and Oprea, T.I. (2004) 'Pursuing the leadlikeness concept in pharmaceutical research.', *Current Opinion in Chemical Biology*, vol. 8, no. 3, pp. 255-263.

- Harris, S., Shiau, A. and Agard, D. (2004) 'The crystal structure of the carboxy-terminal dimerization domain of HtpG, the Escherichia coli Hsp90, reveals a potential substrate binding site.', *Structure*, vol. 12, pp. 1087-1097.
- Havsteen, B.H. (2002) 'The biochemistry and medical significance of the flavonoids', *Pharmacology & Therapeutics*, vol. 96, pp. 202-267.
- Hessling, M., Richter, K. and Buchner, J. (2009) 'Dissection of the ATP-induced conformational cycle of the molecular chaperone Hsp90', *Nat. Struct. Mol. Biol.*, vol. 16, pp. 287-293.
- Hickey, E., Brandon, S., Potter, R., Stein, G., Stein, J. and Weber, L. (1986) 'Sequence and organization of genes encoding the human 27 kDa heat shock protein.', *Nucleic Acids Res*, vol. 14, no. 10, pp. 4127-4145.
- Hillisch, A., Pineda, L.F. and Hilgenfeld, R. (2004) 'Utility of homology models in the drug discovery process.', *Drug Discovery Today*, vol. 15, no. 9, pp. 659-69.
- Huang, K.H., Veal, J.M., Fadden, R.P., Rice, J.W., Eaves, J., Strachan, J.P., Barabasz, A.F., Foley, B.E., Barta, T.E., Ma, W., Silinski, M.A. and al. (2009) 'Discovery of novel 2-aminobenzamide inhibitors of heat shock protein 90 as potent, selective and orally active antitumor agents.', *J. Med. Chem.*, vol. 52, pp. 4288-4305.
- Immormino, R.M., Kang, Y., Chiosis, G. and Gewirth, D.T. (2006) 'Structural and quantum chemical Studies of 8-ARYL-SULFANYL ADENINE class HSP90 Inhibitors.', *J.MED.CHEM.*, vol.49, p. 4953.
- Jorgensen, W.L. (2004) 'The many roles of computation in drug discovery.', *Science*, vol. 5665, no. 303, pp. 1808-1813.
- Jorgensen, W.L. and Tirado-Rives, J. (1988) *J. Am. Chem. Soc.*, vol. 110, p. 1657.
- Kachlicki, M. and Stobiecki, P. (2006) 'Isolation and Identification of Flavonoids', in Grotewold, E. *The Science of Flavonoids*, ©Springer Science + Business Media, Inc.
- Kallblad, P. and Dean, P.M. (2003) 'Efficient conformational sampling of local sidechain flexibility.', *J. Mol. Biol.*, vol. 326, no. 5, pp. 1651-1665.
- Kamal, A., Boehm, M.F. and Burrows, F.J. (2004) 'Therapeutic and diagnostic implications of Hsp90 activation', *Trends Mol. Med.*, vol. 10, no. 6, pp. 283-290.
- Kamal, A., Thao, L., Sensintaffar, J., Zhang, L., Boehm, M.F., Fritz, L.C. and Burrows, F.J. (2003) 'A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors.', *Nature*, vol. 425, pp. 407-410.
- Kampinga, H.H., Hageman, J., Vos, M.J., Kubota, H., Tanguay, R.M., Bruford, E.A., Cheetham, M.E., Chen, B. and Hightower, L.E. (2009) 'Guidelines for the nomenclature of the human heat shock proteins.', *Cell stress & chaperones*, vol. 14, pp. 105-111.
- Kanadaswami, C., Lee, L.T., Lee, P.P., Hwang, J.J. and al (2005) 'The antitumor activities of flavonoids', *in vivo*, vol. 19, pp. 895-910.
- Kang, B., Plescia, J., Dohi, T., Rosa, J., Doxsey, S. and Altieri, D. (2007) 'Regulation of tumor cell mitochondrial homeostasis by an organelle-specific hsp90 chaperone network.', *Cell*, vol. 131, pp. 257-270.

-
- Kasibhatla, S.R., Hong, K. and Biamonte, M.A.. (2007) 'Rationally designed high-affinity 2- amino-6-halopurine heat shock protein 90 inhibitors that exhibit potent antitumor activity.', *J. Med. Chem.*, vol. 50, p. 2767–2778.
- Katschinski, D., Le, L., Heinrich, D., Wagner, K., Hofer, T., Schindler, S. and Wenger, R. (2002)'Heat induction of the unphosphorylated form of hypoxia-inducible factor-1alpha is dependent on heat shock protein-90 activity.', *J Biol Chem*, vol. 277, pp. 9262-9267.
- Kitchen, D.B., Decornez, H., Furr, J.R. and Bajorath, J. (2004) 'Docking and scoring in virtual screening for drug discovery: methods and applications.', *Nature Reviews Drug Discovery*, vol. 11, no. 3, pp. 935-949.
- Knegtel, R.M.A., Kuntz, I.D. and Oshiro, C.M. (1997) 'Molecular docking to ensembles of protein structures.', *J. Mol. Biol.*, vol. 266, pp. 424-440.
- Kokh, D.B. and Wenzel, W. (2008) 'Flexible side chain models improve enrichment rates in silico screening', *J. Mol. Med.*, vol. 51, p. 5919.
- Kollman, P. (1993) 'Free energy calculations: Applications to chemical and biochemical phenomena.', *Chem. Rev.*, vol. 93, no. 7, pp. 2395-2417.
- Kong, J., Chia, L., Goh, N., Chia, T. and Brouillard, R. (2003) 'Analysis and biological activities of anthocyanins.', *Phytochemistry*, vol. 64, p. 923–933.
- Koshland Jr., D. (1994) 'The key–lock theory and the induced fit theory.', *Angew. Chem., Int. Ed. Engl.*, vol. 33, no. 23-24, pp. 2375-2378.
- Krishna, P. and Gloor, G. (2001) 'The Hsp90 family of proteins in *Arabidopsis thaliana*.', *Cell Stress Chaperones*, vol. 6, pp. 238-246.
- Kühlbrandt, W. and Williams, K.A. (1999) 'Analysis of macromolecular structure and dynamics by electron cryo-microscopy.', *Curr. Opin. Chem. Biol.*, vol. 3, pp. 537-543.
- Laederach, A. and Reilly, P.J. (2003) 'Specific empirical free energy function for automated docking of carbohydrates to proteins.', *J. Comput. Chem.*, vol. 24, no. 14, pp. 1748-1757.
- Lai, B.T., Chin, N.W., Stanek, A.E., Keh, W. and Lanks, K.W. (1984) 'Quantitation and intracellular localization of the 85K heat shock protein by using monoclonal and polyclonal antibodies.', *Mol Cell Biol*, vol. 4, pp. 2802-2810.
- Le Marchand, L. (2002) 'Cancer preventive effects of flavonoids—a review', *Biomed Pharmacother*, vol. 56, pp. 296-301.
- Lengauer, T., Lemmen, C., Rarey, M. and Zimmermann, M. (2004) 'Novel technologies for virtual screening.', *Drug Discov. Today*, vol. 9, no. 1, pp. 27-34.
- Leroux, V. and Maigret, B. (2007) 'Should structure-based virtual screening techniques be used more extensively in modern drug discovery?', *Comp. App. Chem.*, vol. 24, no. 1, pp. 1-10.
- Liao, C., Sitzmann, M., Pugliese, A. and Nicklaus, M.C. (2011) 'Software and resources for computational medicinal chemistry', *Future Med Chem.*, vol. 8, no. 3, pp. 1057-1085.
- Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J. (2001) 'Experimental and computational approaches to estimate solubility and permeability in drug discovery
-

- and development settings.', *Advanced Drug Delivery Reviews*, vol. 46, no. 1-3, pp. 3-26.
- Li, J., Soroka, J., Buchner and J. (2012) 'The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones', *Biochimica et Biophysica Acta*, vol. 1823, pp.624-635.
- Li, Y., Zhang, T., Jiang, Y., Lee, H.F., Schwartz, S.J. and Sun, D. (2009) '(-)- Epigallocatechin-3- gallateinhibitsHsp90functionbyimpairing Hsp90 associationwithcochaperonesinpancreaticcancer cellline Mia Paca-2.', *Mol Pharm*, vol. 6, pp. 1152-1159.
- Li, Y., Zhang, D., Xu, J., Shi, J., Jiang, L., Yao, N. and Ye, W. (2012) 'Discovery and development of natural heat shock protein 90 inhibitors in cancer treatment', *Acta Pharmaceutica Sinica B*, vol. 2, no. 2, pp. 238-245.
- Louvion, J.F., Warth, R. and Picard, D. (1996) 'Two eukaryote-specific regions of Hsp82 are dispensable for its viability and signal transduction functions in yeast.', *Proc Natl Acad Sci USA*, vol. 93, pp. 13937-13942.
- Lyne, P.D. (2002) 'Structure-based virtual screening: an overview.', *Drug Discovery Today*, vol. 20, no. 7, pp. 1047-1055.
- Macheix, J.J., Fleuriet, A., Billot, J. and Jay-Allemand, C. (1990) *Fruit phenolics*, Boca Raton: CRC Press.
- Maki, R.G., Old, L.J. and Srivastava, P.K. (1990) 'Human homologue of murine tumor rejection antigen gp96: 5'-regulatory and coding regions and relationship to stress-induced proteins.', *Proc Natl Acad Sci USA*, vol. 87, no. 17, pp. 5658-5662.
- Maloney, A. and Workman, P. (2002) 'Hsp90 as a new therapeutic target for cancer therapy: the story unfolds.', *Expert Opin Biol Ther*, vol. 2, pp. 3-24.
- Marais, J.P.J., Deavours, B., Dixon, R.A. and Ferreira, D. (2006) 'The Stereochemistry of Flavonoids', in Grotewold, E. (ed.) *The Science of Flavonoids*, ©Springer Science + Business Media, Inc.
- Marcu, M.G., Chadli, A., Bouhouche, I., Catelli, M. and Neckers, L.M. (2000) 'The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP-binding domain in the carboxyl terminus of the chaperone.', *J. Biol. Chem.*, vol. 275, pp. 37181-37186.
- Mason, J.S., Good, A.C. and Martin, E.J. (2001) '3-D pharmacophores in drug discovery.', *Curr. Pharm. Des.*, vol. 7, no. 7, pp. 567-597.
- Meyer, P., Prodromou, C., Hu, B., Vaughan, C., Roe, S.M., Panaretou, B., Piper, P.W. and Pearl, L.H. (2003) 'Structural and functional analysis of the middle segment of Hsp90: Implications for ATP hydrolysis and client-protein and co-chaperone interactions.', *Mol Cell*, vol. 11, pp. 647-658.
- Minami, Y., Kimura, Y., Kawasaki, H., Suzuki, K. and Yahara, I. (1994) 'The carboxy-terminal region of mammalian HSP90 is required for its dimerization and function in vivo.', *Mol Cell Biol*, vol. 14, no. 2, pp. 1459-1464.
- Montesano Gesualdi, N., Chirico, G., Pirozzi, G., Costantino, E., Landriscina, M. and Esposito, F. (2007) 'Tumor necrosis factor-associated protein 1 (TRAP-1) protects cells from oxidative stress and apoptosis.', *Stress*, vol. 10, no. 9, pp. 342-350.

- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hartl, W.E., Belew, R.K. and Olson, A.J. (1998) 'Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.', *J. Comput. Chem.*, vol. 19, p. 1639.
- Moulin, E., Zoete, V., Barluenga, S., Karplus, M. and Winssinger, N. (2005) 'Design, synthesis, and biological evaluation of HSP90 inhibitors based on conformational analysis of radicicol and its analogues', *J. Am. Chem. Soc.*, vol. 127, pp. 6999-7004.
- Muegge, I. and Oloff, S. (2006) 'Advances in virtual screening.', *Drug Discovery Today*, vol. 4, no. 3, pp. 405-411.
- Nakai, A. (1999) 'New aspects in the vertebrate heat shock factor system: Hsf3 and Hsf4.', *Cell stress & chaperones*, vol. 4, pp. 86-93.
- Nemoto, T. and Sato, N. (1998) 'Oligomeric forms of the 90-kDa heat shock protein.', *Biochem J*, vol. 330, pp. 989-995.
- Nicklaus, M.C., Wang, S., Driscoll, J.S. and Milne, G.W. (1995) 'Conformational changes of small molecules binding to proteins.', *Bioorg. Med. Chem.*, vol. 4, no. 3, pp. 411-428.
- Nover, L., Scharf, K.D., Gagliardi, D., Vergne, P., Czarnecka-Verner, E. and Gurley, W.B. (1996) 'The Hsf world: classification and properties of plant heat stress transcription factors.', *Cell stress & chaperones*, vol. 1, pp. 215-223.
- Onodera, K., Satou, K. and Hirota, H. (2007) 'Evaluations of molecular docking programs for virtual screening.', *J Chem Inf Model.*, vol. 4, no. 47, pp. 1609-1618.
- Oprea, T.I. (2000) 'Property distribution of drug-related chemical databases.', *J Comput Aided Mol Des*, vol. 14, no. 3, pp. 251-264.
- Oprea, T.I. and Matter, H. (2004) 'Integrating virtual screening in lead discovery.', *Current Opinion in Chemical Biology*, vol. 8, no. 4, pp. 349-358.
- Pak, Y.S. and Wang, S. (2000) 'Application of a molecular dynamics simulation method with a generalized effective potential to the flexible molecular docking problems.', *J. Phys. Chem. B*, vol. 104, p. 354.
- Palermo, C.M., Westlake, C.A. and Gasiewicz, T.A. (2005) 'Epigallocatechin gallate inhibits aryl hydrocarbon receptor gene transcription through an indirect mechanism involving binding to a 90 kDa heat shock protein.', *Biochemistry*, vol. 44, pp. 5041-5052.
- Peach, M.L. and Nicklaus, M.C. (2009) 'Combining docking with pharmacophore filtering for improved virtual screening.', *J Cheminform.*, vol. 1, no. 1, p. 6.
- Pearl, L.H. and Prodromou, C. (2006) 'Structure and mechanism of the Hsp90 molecular chaperone machinery.', *Annu. Rev. Biochem.*, vol. 75, pp. 271-294.
- Pearl, L.H., Prodromou, C. and Workman, P. (2008) 'The Hsp90 molecular chaperone: an open and shut case for treatment', *Biochem. J.*, vol. 410, pp. 439-453.
- Perez, J.J. (2005) 'Managing molecular diversity.', *Chemical Society Reviews*, vol. 34, no. 2, pp. 143-152.
- Picard, D., Khursheed, B., Garabedian, M.J., Fortin, M.G., Lindquist, S. and Yamamoto, K.R. (1990) 'Reduced levels of hsp90 compromise steroid receptor action in vivo.', *Nature*, vol. 348, pp. 166-168.

- Plewczynski, D., Lazniewski, M., Augustyniak, R. and Ginalski, K. (2011) 'Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database.', *J Comput Chem.*, vol. 32, no. 4, pp. 742-755.
- Pratt, W.B. (1992) 'Control of steroid receptor function and cytoplasmic-nuclear transport by heat shock proteins.', *Bioessays*, vol. 14, pp. 841-848.
- Pratt, W.B. and Toft, D.O. (2003) 'Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery.', *Exp Biol Med*, vol. 228, no. 2, pp. 111-133.
- Prior, R.L. and Wu, X. (2006) 'Anthocyanins: Structural characteristics that result in unique metabolic patterns and biological activities.', *Free Rad Res*, vol. 40, pp. 1014-1028.
- Prodromou, C. (2012) 'The 'active life' of Hsp90 complexes', *Biochimica et Biophysica Acta*, vol. 1823, pp. 614-623.
- Prodromou, C., Panaretou, B., Chohan, S., Siligard, i.G., O'Brien, R., Ladbury, J.E., Roe, S.M., Piper, P.W. and Pearl, L.H. (2000) 'The ATPase cycle of Hsp90 drives a molecular 'clamp' via transient dimerization of the N-terminal domains.', *The EMBO Journal*, vol. 19, pp. 4383- 4392.
- Prodromou, C., Roe, S.M., O'Brien, R., Ladbury, J.E., Piper, P.W. and Pearl, L.H. (1997) 'Identification and Structural Characterization of the ATP/ADP-Binding Site in the Hsp90 Molecular Chaperone', *Cell*, vol. 90, pp. 65-75.
- Proisy, N., Sharp, S.Y., Boxall, K., Connelly, S., Roe, S.M., Prodromou, C., Slawin, A.M.Z., Pearl, L.H., Workman, P.C. and Moody, J. (2006) 'Inhibition of Hsp90 with synthetic marolactones : synthesis and structural and biological evaluation of ring and conformational analogs of radicicol', *Chem Biol*, vol. 13, p. 1203.
- Reddy, R.K., Dubeau, L., Kleiner, H., Parr, T., Nichols, P., Ko, B. and al. (2002) 'Cancer- inducible transgene expression by the Grp94 promoter: spontaneous activation in tumors of various origins and cancer-associated macrophages.', *Cancer Res*, vol. 62, no. 24, pp. 7207-7212.
- Ren, W., Qiao, Z., Wang, H., Zhu, L. and Zhang, L. (2003) 'Flavonoids : promising anticancer agents.', *Med Res Rev*, vol. 23, no. 4, pp. 519-534.
- Ripphausen, P., Nisius, B., Peltason, L. and Bajorath, J. (2010) 'Quo Vadis, virtual screening? A comprehensive survey of prospective applications.', *J Med Chem*, vol. 53, no. 24, pp. 8461-8467.
- Robert, J., Menoret, A. and Cohen, N. (1999) 'Cell surface expression of the endoplasmic reticular heat shock protein gp96 is phylogenetically conserved.', *J Immunol*, vol. 163, no. 8, pp. 4133-4139.
- Rossi, M., Bosetti, C., Negri, E., Lagioui, P. and La Vecchia, C. (2010) 'Flavonoids, proanthocyanidins, and cancer risk: a network of case–control studies from Italy.', *Nutr Cancer*, vol. 62, no. 7, pp. 871-877.
- Ross, J.A. and Kasum, C.M. (2002) 'Dietary flavonoids: bioavailability, metabolic effects, and safety', *Annu Rev Nutr*, vol. 22, pp. 19-34.

- Sato, S., Fujita, N. and Tsuruo, T. (2000) 'Modulation of Akt kinase activity by binding to Hsp90', *Proc. Natl. Acad. Sci. USA*, vol. 97, no. 20, pp. 10832-10837.
- Scherman, W., Day, T., Jacobson, M.P., Friesner, R.A. and Farid, R. (2006) 'Novel procedure for modeling ligand/receptor induced fit effects.', *J. Med. Chem.*, vol. 49, pp. 534-553.
- Schulte, T.W., Akinaga, S., Soga, S., Sullivan, W., Stensgard, B., Toft, D. and Neckers, L. (1998) 'Antibiotic radicicol binds to the N-terminal domain of Hsp90 and shares important biologic activities with geldanamycin', *Cell Stress Chaperones*, vol. 3, pp. 100-108.
- Schulte, T.W. and Neckers, L.M. (1998) 'The benzoquinone ansamycin 17-allylamino-17-demethoxygeldanamycin binds to Hsp90 and shares important biologic activities with geldanamycin.', *Cancer Chemother Pharmacol*, vol. 42, pp. 273-279.
- Shoichet, B.K. (2004) 'Virtual screening of chemical libraries.', *Nature*, vol. 7019, no. 432, pp. 862-865.
- Siega, I.G., Van Duynhoven, J. and Baldus, M. (1999) 'Biomolecular NMR: recent advances in liquids, solids and screening.', *Curr. Opin. Chem. Biol.*, vol. 3, pp. 530-543.
- Simonson, T., Archontis, G. and Karplus, M. (2002) 'Free energy simulations come of age: protein-ligand recognition.', *Acc. Chem. Res.*, vol. 35, no. 6, pp. 430-437.
- Singh, B.N., Shankar, S. and Srivastava, R. (2011) 'Green tea catechin, epigallocatechin-3-gallate (EGCG): Mechanisms, perspectives and clinical applications', *Biochemical Pharmacology*, vol. 82, pp. 1807-1821.
- Smyth, T., Ramachandran, V.N. and Smyth, W.F. (2009) 'A study of the antimicrobial activity of selected naturally occurring and synthetic coumarins.', *International journal of antimicrobial agents*, vol. 33, pp. 421-426.
- Sorger, P.K. and Pelham, H.R. (1988) 'Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation.', *Cell*, vol. 54, pp. 855-864.
- Soti, C., Racz, A. and Csermely, P. (2002) 'A Nucleotide-dependent molecular switch controls ATP binding at the C-terminal domain of Hsp90. N-terminal nucleotide binding unmasks a C-terminal binding pocket.', *J Biol Chem*, vol. 277, pp. 7066-7075.
- Sousa, S.F., Fernandes, P.A. and Ramos, M.J. (2006) 'Protein-ligand docking: current status and future challenges.', *Proteins : Struct., Funct., Bioinf.*, vol. 65, no. 1, pp. 15-26.
- Stahura, F.L. and Bajorath, J. (2004) 'Virtual screening methods that complement HTS.', *Comb Chem High Throughput Screen*, vol. 7, no. 4, pp. 69-259.
- Stebbins, C.E., Russo, A.A., Schneider, C., Rosen, N., Hartl, F.U. and Pavletich, N.P. (1997) 'Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent.', *Cell*, vol. 89, pp. 239-250.
- Stevenson, D.E., Scheepens, A. and Hurst, R.D. (2009) 'Bioavailability and Metabolism of Dietary Flavonoids – Much Known – Much More to Discover', in Raymond, B.K.

- (ed.) Flavonoids : biosynthesis, biological effects and dietary sources, New York: Nova Science Publishers, Inc.
- Stewart, A.J., Mullen, W. and Crozier, A. (2005) 'On-line high-performance liquid chromatography analysis of the antioxidant activity of phenolic compounds in green and black tea.', *Mol Nutr Food Res*, vol. 49, p. 19.
- Sudbeck, E.A., Mao, C., Venkatachalam, T.K., Thuel-Ahlgren, L. and Uckum, F.M. (1998)'Structure- based design of novel dihydroalkoxybenzyloxypyrimidine derivatives as potent nonnucleoside inhibitors of the human immunodeficiency virus reverse transcriptase', *Antimicrob. Agents Chemother.*, vol. 42, p. 3225.
- Supko, J.G., Hickman, R.L., Grever, M.R. and Malspeis, L. (1995) 'Preclinical pharmacologic evaluation of geldanamycin as an antitumour agent.', *Cancer Chemother, Pharmacol*, vol. 36, pp. 305-315.
- Suriano, R., Ghosh, S.K., Ashok, B.T., Mittelman, A., Chen, Y., Banerjee, A. and al. (2005)'Differences in glycosylation patterns of heat shock protein, gp96: implications for prostate cancer prevention.', *Cancer Res*, vol. 65, no. 24, pp. 6466-6475.
- Sydor, J.R., Normant, E., Pien, C.S., Porter, J.R., Ge, J. and al (2006) 'Development of 17- allylamino-17-demethoxygeldanamycin hydroquinone hydrochloride (IPI-504), an anti- cancer agent directed against Hsp90.', *Proc. Natl. Acad. Sci. USA*, vol. 103, pp. 17408-17413.
- Talele, T.T., Khedkar, S.A. and Rigby, A.C. (2010) 'Successful applications of computer aided drug discovery: moving drugs from concept to the clinic.', *Curr Top Med Chem*, vol. 10, no. 1, pp. 127-141.
- Tang, Y., Zhu, W., Chen, K. and Jiang, H. (2006) 'New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery.', *Drug Discovery Today*, vol. 3, no. 3, pp. 307-313.
- Teague, S.J. (2003) 'Implications of protein flexibility for drug discovery.', *Nat. Rev. DrugDiscovery*, vol. 7, no. 2, pp. 527-541.
- Tissieres, A., Mitchell, H.K. and Tracy, U.M. (1974) 'Protein synthesis in salivary glands of *Drosophila melanogaster*: relation to chromosome puffs.', *J Mol Biol*, vol. 84, pp. 389-398.
- Todeschini, R. and Consonni, V. (2000) *Handbook of Molecular Descriptors.* , New York: Wiley- VCH.
- Tran, P., Kim, S.A., Choi, H.S., Yoon, J.H. and Ahn, S.G. (2010) 'Epigallocatechin-3-gallate suppresses the expression of Hsp70 and Hsp90 and exhibits anti-tumor activity in vitro and in vivo.', *BMC Cancer*, vol. 10, p. 276.
- Travers, J., Sharp, S. and Workman, P. (2012) 'HSP90 inhibition: two-pronged exploitation of cancer dependencies', *Drug Discovery Today*, vol. 17, no. 5-6, pp. 242-252.
- Vieth, M., Hirst, J.D., Kolinski, A. and Brooks, C.L. (1998) ' Assessing energy functions for flexible docking.', *J. Comput. Chem.* , vol. 19, pp. 1612-1622.
- Vilenchik, M., Solit, D., Basso, A., Huezio, H., Lucas, B., He, H., Rosen, N., Spampinato, C., Modrich, P. and Chiosis, G. (2004) 'Targeting wide-range oncogenic transformation via PU24FCI, a specific inhibitor of tumor Hsp90.', *Chem. Biol.*, vol. 11, pp. 787-797.

-
- Voss, A.K., Thomas, T. and Gruss, P. (2000) 'Mice lacking HSP90E fail to develop a placental labyrinth.', *Development*, vol. 127, pp. 1-11.
- Walters, W.P., Stahl, M.T. and Murcko, M.A. (1998) 'Virtual screening - an overview.', *Drug Discov. Today*, vol. 4, no. 3, pp. 160-178.
- Wandinger, S.K., Richter, K. and Buchner, J. (2008) 'The Hsp90 chaperone machinery.', *J Biol Chem*, vol. 283, pp. 18473-18477.
- Wang, R., Lu, Y., Fang, X. and Wang, S. (2004) 'An extensive test of 14 scoring functions using the PDBbind refined set of 800 protein-ligand complexes.', *J. Chem. Inf. Comput. Sci.*, vol. 44, no. 6, pp. 2114-2125.
- Wegele, H., Muller, L. and Buchner, J. (2004) 'Hsp70 and Hsp90--a relay team for protein folding.', *Rev Physiol Biochem Pharmacol*, vol. 151, pp. 1-44.
- Wong, C.F., Kua, J., Zhang, Y., Straatsma, T.P. and McCammon, J.A. (2005) 'Molecular docking of balanol to dynamics snapshots of protein kinase A.', *Proteins: Struct., Funct., Bioinf.*, vol. 61, no. 4, pp. 850-858.
- Wright, L., Barril, X., Dymock, B., Sheridan, L., Surgenor, A., Beswick, M., Drysdale, M., Collier, A., Massey, A., Davies, N., Fink, A., Fromont, C., Aherne, W., Boxall, K., Sharp, S., Workman, P. and Hubbard, R.E. (2004) 'Structure-Activity Relationships in Purine-Based Inhibitor Binding to HSP90 Isoforms.', *Chem. Biol.*, vol. 11, pp. 775-785.
- Wu, C. (1995) 'Heat shock transcription factors: structure and regulation.', *Annu Rev Cell Dev Biol*, vol. 11, pp. 441-469.
- Wu, X.L., Beecher, G.R., Holden, J.M., Haytowitz, D.B., Gebhardt, S.E. and Prior, R.L. (2006) 'Concentrations of anthocyanins in common foods in the United States and estimation of normal consumption. .', *Journal of Agricultural and Food Chemistry*, vol. 54, pp. 4069-4075.
- Xu, Y. and Lindquist, S. (1993) 'Heat-shock protein hsp90 governs the activity of pp60v-src kinase.', *Proc Natl Acad Sci USA*, vol. 90, pp. 7074-7078.
- Yang, Y. and Li, Z. (2005) 'Roles of heat shock protein gp96 in the ER quality control: redundant or unique function?', *Mol Cells*, vol. 20, no. 2, pp. 173-182.
- Yang, Y., Liu, B., Dai, J., Srivastava, P.K., Zammit, D.J., Lefrancois, L. and al. (2007) 'Heat shock protein gp96 is a master chaperone for toll-like receptors and is important in the innate function of macrophages.', *Immunity*, vol. 26, no. 2, pp. 215-226.
- Yin, Z., Henry, E.C. and Gasiewicz, T.A. (2009) '(#)-Epigallocatechin-3-gallate is a novel Hsp90 inhibitor.', *Biochemistry*, vol. 48, no. 2, pp. 336-345.
- Yue, L., Karr, T.L., Nathan, D.F., Swift, H., Srinivasan, S. and Lindquist, S. (1999) 'Genetic analysis of viable Hsp90 alleles reveals a critical role in Drosophila spermatogenesis.', *Genetics*, vol. 151, pp. 1065-1079.
- Zhou, Z., Felts, A.K., Friesner, R.A. and Levy, R.M. (2007) 'Comparative Performance of Several Flexible Docking Programs and Scoring Functions: Enrichment Studies for a Diverse Set of Pharmaceutically Relevant Targets.', *J. Chem. Inf. Model.*, vol. 47, no. 4, pp. 1599- 1608.
-

Annexes

Annexe01

- **#rotor (0 – 15)**: Number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds.
- **mol_MW (130.0 – 725.0)**: Molecular weight of the molecule.
- **dipole (1.0 – 12.5)**: Computed dipole moment of the molecule.
- **SASA (300.0 – 1000.0)**: Total solvent accessible surface area (SASA) in square angstroms using a probe with a 1.4 Å radius.
- **FOSA (0.0 – 750.0)**: Hydrophobic component of the SASA (saturated carbon and attached hydrogen).
- **FISA (7.0 – 330.0)**: Hydrophilic component of the SASA (SASA on N, O, and H on heteroatoms).
- **PISA (0.0 – 450.0)**: π (carbon and attached hydrogen) component of the SASA.
- **WPSA (0.0 – 175.0)**: Weakly polar component of the SASA (halogens, P, and S).
- **volume (500.0 – 2000.0)**: Total solvent-accessible volume in cubic angstroms using a probe with a 1.4 Å radius.
- **donorHB (0.0 – 6.0)**: Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
- **accptHB (2.0 – 20.0)**: Estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
- **glob (0.75 – 0.95)**: Globularity descriptor, where r is the radius of a sphere with a volume equal to the molecular volume. Globularity is 1.0 for a spherical molecule.
- **QPpolrz (13.0 – 70.0)**: Predicted polarizability in cubic angstroms.
- **QPlogPC16 (4.0 – 18.0)**: Predicted hexadecane/gas partition coefficient. QPlogPoct (8.0 – 35.0): Predicted octanol/gas partition coefficient. QPlogPw (4.0 – 45.0): Predicted water/gas partition coefficient. QPlogPo/w (–2.0 – 6.5): Predicted octanol/water partition coefficient.
- **QPlogS (–6.5 – 0.5)**: Predicted aqueous solubility, $\log S$; S in mol dm^{-3} is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid.
- **QPlogBB (–3.0 – 1.2)**: Predicted brain/blood partition coefficient. Note: QikProp predictions are for orally delivered drugs so, for example, dopamine and serotonin are CNS negative because they are too polar to cross the blood-brain barrier
- **QPlogKhsa (–1.5 – 1.5)**: Prediction of binding to human serum albumin.
- **IP(ev) (7.9 – 10.5)**: PM3 calculated ionization potential.
- **EA(eV) (–0.9 – 1.7)** PM3 calculated electron affinity.

- **#metab (1 – 8)**: Number of likely metabolic reactions.