

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mentouri Constantine
Faculté des Sciences de la Nature et de la Vie
Département de Biochimie et Microbiologie

N° d'ordre:

N° de série :

Mémoire en vue de l'obtention du diplôme de magistère en Biochimie

Option : Technologie des Explorations Biochimiques

Thème

**Conception *in silico* de nouveaux inhibiteurs de la
topoisomérase I de cellules cancéreuses**

Présenté par : Teniou Soumia

Devant le jury :

Président:	Mr. Chikhi A.	M.C. Univ. Mentouri Constantine.
Rapporteur:	Mr. Bensegueni A.	M.C. Univ. Mentouri Constantine.
Examineurs:	Mr. Khelifi D.	Prof. Univ. Mentouri Constantine.
	Mr. Nouadri T.	M.C. Univ. Mentouri Constantine.

Année universitaire : 2011 / 2012

Remerciements

J'exprime ma sincère gratitude à Monsieur Bensegueni A -Maitre de conférences- pour sa précieuse et honorable aide dans l'orientation et la direction de ce travail.

J'adresse mes respectueux remerciements à Monsieur Chikhi A - Maitre de conférences- qui ma fait l'honneur d'avoir accepté de présider le jury de ce mémoire.

Mes profonds remerciements à Monsieur le Professeur Khelifi D. d'avoir accepté de faire partie du jury.

Mes sincères remerciements à Monsieur Nouadri T -Maitre de conférences- d'avoir accepté de juger ce travail.

Mes vifs remerciements à Monsieur le Professeur Teniou A. pour sa disponibilité, sa patience et son aide précieux.

Je tiens sincèrement à remercier toute personne ayant collaboré de loin ou de près à la réalisation de ce travail en particulier mon collègue Mokrani Elhacen.

Enfin, je tiens à remercier ma famille pour son soutien sans faille durant ce travail.

Table des matières

Liste des abréviations

Liste des tableaux

Liste des figures

Introduction générale1

ETUDE BIBLIOGRAPHIQUE

CHAPITRE I : Les flavonoïdes

1- Origine	3
2- Biosynthèse des flavonoïdes	3
3- Intérêt biologique	7
4- Propriétés anticancéreuses des flavonoïdes	7
5- Critères de Choix des flavonoïdes.....	8

CHAPITRE II: Généralités sur le cancer

1- Définition et caractéristiques du cancer.....	10
2- Diagnostic et traitements	11

CHAPITRE III: Les topoisomérases d'ADN

1- Définition	14
2- Classification des topoisomérases.....	17
3- Topoisomérase I	18
3-1 Structure	18
3-2 Mécanisme enzymatique	19
3-3 Fonctions biologiques	22

4- Les inhibiteurs de l'ADN topoisomérase I humaine	22
4-1 Les poisons	23
4-2 Les supprimeurs.....	24

CHAPITRE IV : Modélisation des interactions protéine-ligand par docking moléculaire

1- Modélisation moléculaire	25
2- Reconnaissance moléculaire	26
2-1 Interactions de Van Der Waals.....	26
2-2 Ponts hydrogènes.	26
2-3 Interactions électrostatiques	27
3- Amarrage moléculaire (ou docking)	27
3-1 Les différents types de docking moléculaire	28
3-2 Principe de fonctionnement	29
3-3 Le processus de scoring	30
3-3-1. Représentation géométrique.....	31
3-3-2. Représentation atomique.....	31
4- Le programme de docking moléculaire Surflex	33

ETUDE PRATIQUE

I- Matériels et méthodes

1- Préparation de la protéine	35
2- Préparation des flavonoïdes	35
3- Docking	35

II- Résultats et discussions

1-Test de fiabilité du logiciel	36
2-Interactions entre les flavonoïdes et la topoisomérase I.....	45
2-1 Règle de Lipinski.....	45
2-2. L'analyse visuelle des interactions.....	46

3-Substitution de la myricétine.....	53
4- Application de la règle de Lipinski sur les 13 composés.....	55
5-Etude des interactions du composé 28.....	57
Conclusion et perspectives	59
Références bibliographiques	
Annexes	

Liste des abréviations

ADME: Absorption, digestion, métabolisme et 'excrétion

AH: Accepteur hydrogène

AM1: Austin Model 1

CL: COA-ligase

C4H: Cinnamate-4-hydroxylase

DH: Donneur hydrogène

LF: Liaison flexible

PAL: Phenyl ammonia lyase

PDB: Protein Data Bank

RMN: Résonance magnétique nucléaire

RMSD: root-mean-square derivation

4CL: COA-ligase

Liste des tableaux

Tableau 1. Structure chimique des flavonoïdes inhibant l'ADN topoisomérase I.....	9
Tableau 2. Valeurs du RMSD des six complexes protéine-ligand.....	38
Tableau 3. Résultats de la règle de Lipinski sur les quatre flavonoïdes.....	45
Tableau 4. Résultats théoriques de l'affinité obtenue par docking des flavonoïdes dans le site actif du topoI et leurs relations avec les données expérimentales (IC ₅₀)	47
Tableau 5. Résultats de la substitution.....	54
Tableau 6. résultats de la règle de Lipinski sur les 13 composés	56
Tableau 7. Autre propriétés physico-chimiques du composé 28.....	57

Liste des figures

Figure 1. Motif flavane (a) et flavone (b) et numérotation systématique.....	4
Figure 2. Etapes de la biosynthèse des Flavonoïdes.....	5
Figure 3. Structure des différents sous-groupes de flavonoïdes.....	6
Figure 4. Complexité structurale des flavanols.....	6
Figure 5. Représentation schématisée de la compaction de l'ADN dans la cellule.....	14
Figure 6. Représentation des surenroulements qui interviennent après séparation locale de deux brins sur un ADN circulaire.....	15
Figure 7. Réactions de transestérification déterminant la coupure de l'ADN commune à toutes les topoisomérases. La réaction inverse permet la religation de l'ADN coupé.....	17
Figure 8. Structure de l'ADN topoisomérase I humaine.....	19
Figure 9. Schéma du mécanisme d'action de l'ADN topoisomérase I.....	20
Figure 10. Mécanisme de rotation contrôlée de topoI. L'ADN hautement surenroulé de tours négatifs (en rouge) devient à la fin du processus moins surenroulé (en vert).....	21
Figure 11. Mécanisme d'action d'un poison de l'ADN topoisomérase I.....	23
Figure 12. Interactions de Van Der Waals.....	26
Figure 13. Ponts hydrogène.....	27
Figure 14. Interactions électrostatiques.....	27
Figure 15. Représentation schématique du docking de deux molécules. (a) Molécules de départ. (b) Complexe moléculaire.....	28
Figure 16. Résultats en % obtenus par Surfex à deux intervalles de RMSD (Å).....	37
Figure 17. Superposition du ligand EHD donné par rayon-X (coloré en marron) et par docking moléculaire (coloré en vert).	39

Figure 18. Comparaison de la géométrie de la TTC donnée par rayon-X (coloré en marron) avec celle obtenue par docking moléculaire (coloré en vert).....	40
Figure 19. Superposition du ligand TTG donné par rayon-X (coloré en marron) et par docking moléculaire (coloré en vert).....	41
Figure 20. Comparaison des géométries de l'inhibiteur M38 données par rayon-X (coloré en marron) et par simulation avec Surfex (coloré en vert).....	42
Figure 21. Superposition du ligand AI3 donné par rayon-X (coloré en marron) et par docking moléculaire (coloré en vert).....	43
Figure 22. Comparaison des géométries de l'inhibiteur SA3 données par rayon-X (coloré en marron) et par simulation avec Surfex (coloré en vert).....	44
Figure 23. Positionnement du ligand EHD (coloré en vert) entre les bases T10 et G1 de l'ADN.....	46
Figure 24. Mode de liaison de la myricétine avec le site actif de l'ADN topoisomérase I.....	49
Figure 25. Mode de liaison de la fisétine avec le site actif de l'ADN topoisomérase I.....	50
Figure 26. Mode de liaison de la quercétine avec le site actif de l'ADN topoisomérase I.....	51
Figure 27. Mode de liaison de l'apigénine avec le site actif de l'ADN topoisomérase I.....	52
Figure 28. Structure de la myricétine.....	53
Figure 29. Mode de liaison du composé 28 avec le site actif de l'ADN topoisomérase I.....	58

CHAPITRE 1
Les flavonoïdes

Les flavonoïdes

1. Origine

L'organisme humain ne synthétise pas de flavonoïdes mais, de manière générale, elles sont largement rencontrées dans le règne végétal. Elles sont cependant rares chez les végétaux inférieurs puisque ce sont des métabolites secondaires -composés non issus de la photosynthèse mais de réactions chimiques ultérieures. Nous retrouvons les flavonoïdes dans notre alimentation quotidienne [8, 9, 10,11]. Les fruits (orange, raisin, etc.), les légumes (oignon, laitue, etc.), les graines (fève, cacao) ou encore les racines des plantes sont, pour l'homme, des sources importantes de flavonoïdes. Les feuilles de thé sont également connues pour être riches en flavonoïdes.

La consommation journalière en flavonoïdes n'a jamais été étudiée de manière systématique au vu du très grand nombre de composés et de matrices naturelles à prendre en considération. Les seules estimations précises sur la consommation en flavonoïdes concernent essentiellement les flavonols et les flavones [12,13]. Ainsi la consommation journalière en flavones et flavonols est de l'ordre de quelques dizaines à une centaine de mg.

2. Biosynthèse des flavonoïdes

Les flavonoïdes ont tous une origine biosynthétique commune et par conséquent, possèdent tous un même squelette de base de quinze atomes de carbones constitué de deux unités aromatiques, deux cycles en C6 (A et B) reliés par une chaîne en C3 [14] (fig1).

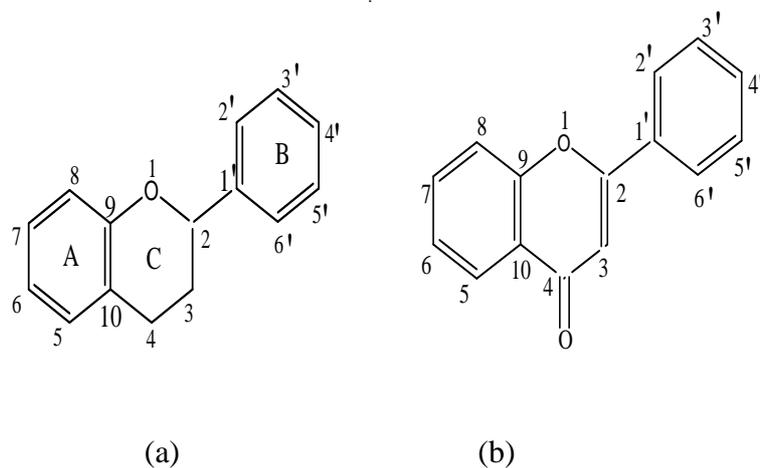


Figure 1 : Motif flavane (a) et flavone (b) et numérotation systématique

A l'origine, les flavonoïdes proviennent de la désamination d'un acide aminé essentiel, la phénylalanine. La réaction de désamination est catalysée par la phényl ammonia lyase (PAL) et conduit à la formation de cinnamate. Ce dernier est ensuite transformé en acide coumarique puis en 4-coumaroyl-coenzyme A, respectivement par l'enzyme cinnamate-4- hydroxylase (C4H) et la CoA-ligase (4CL). Cette réaction préliminaire précède la biosynthèse de tous les flavonoïdes. Le coumaroyl-CoA est ensuite transformé en chalcone en faisant intervenir le malonyl-CoA et la chalcone synthase. Le motif chalcone est ainsi le point de départ de la synthèse des différents groupes de flavonoïdes (Fig. 2) [15]

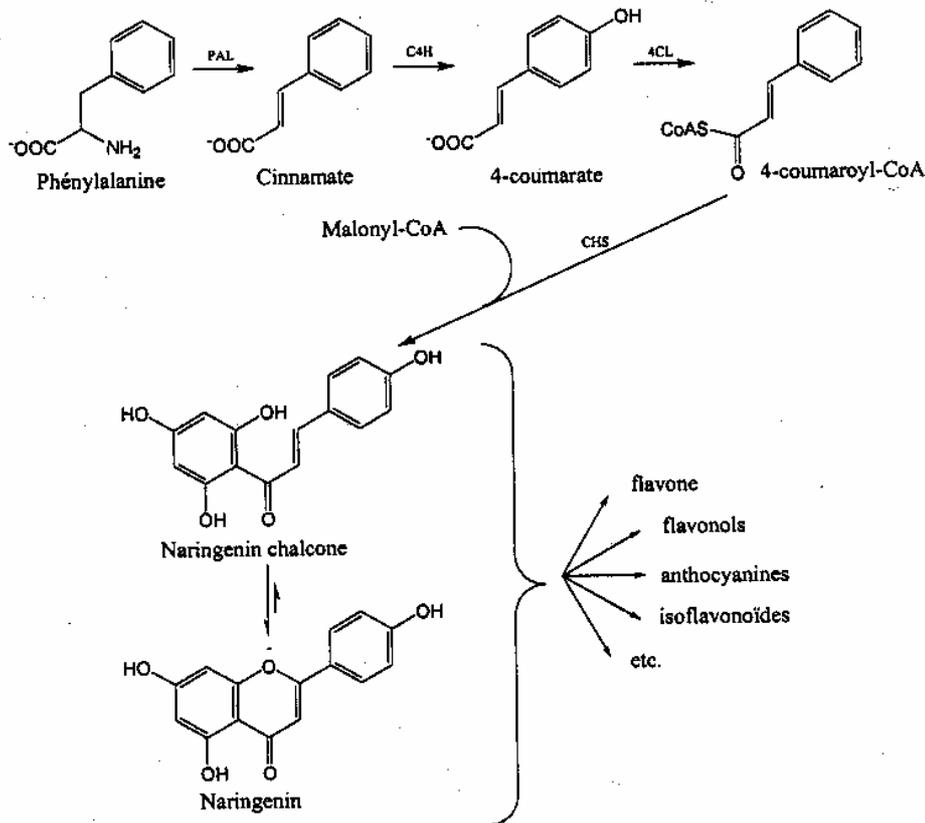


Figure 2 : Etapes de la biosynthèse des Flavonoïdes

Les flavonoïdes sont classés selon plusieurs critères : présence ou non d'une double liaison en position 2, et présence ou non d'un groupement hydroxyle en position. Des groupements hydroxyles (et méthoxyles) peuvent le plus souvent se situer en position 2', 3', 4' et 5' ainsi qu'en position 5 et 7. Les isoflavones proviennent d'une transposition du noyau aromatique du carbone 2 vers le carbone 3 ; les flavones et flavonols d'une oxydation (formation de la double liaison sur le cycle C) respectivement des flavanones et des dihydroflavonols (Fig. 3).

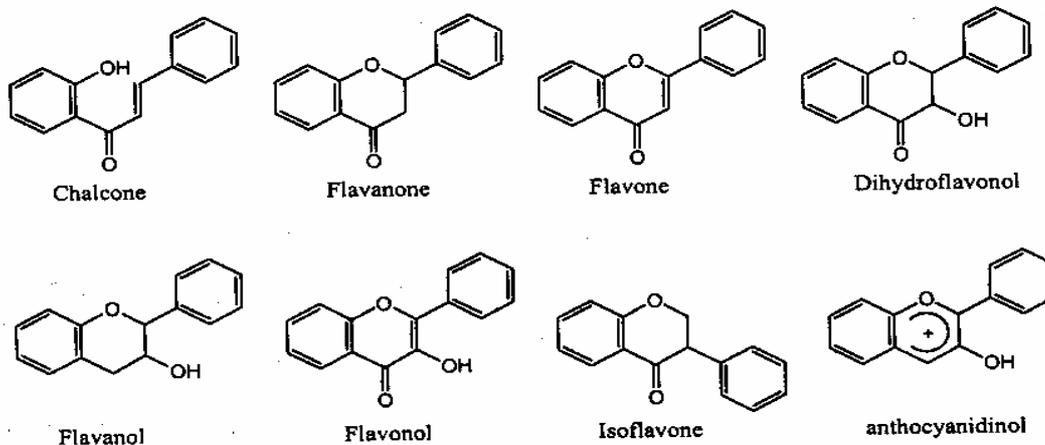


Figure 3 : Structure des différents sous-groupes de flavonoïdes

Les flavonoïdes se rencontrent soit sous forme glycosylée (O et C-hétérosides) soit sous forme d'aglycone ou forme non substituée par un ose. La substitution par les sucres peut avoir lieu soit en position 3, 5 ou 7 des flavones et flavonols. Le cas des flavanols est singulier car en plus des formes glycosylées, il est possible de les rencontrer substituées par l'acide gallique ou encore sous formes polymériques. La configuration des carbones asymétriques augmente la complexité de structure des flavonoïdes (Fig.4).

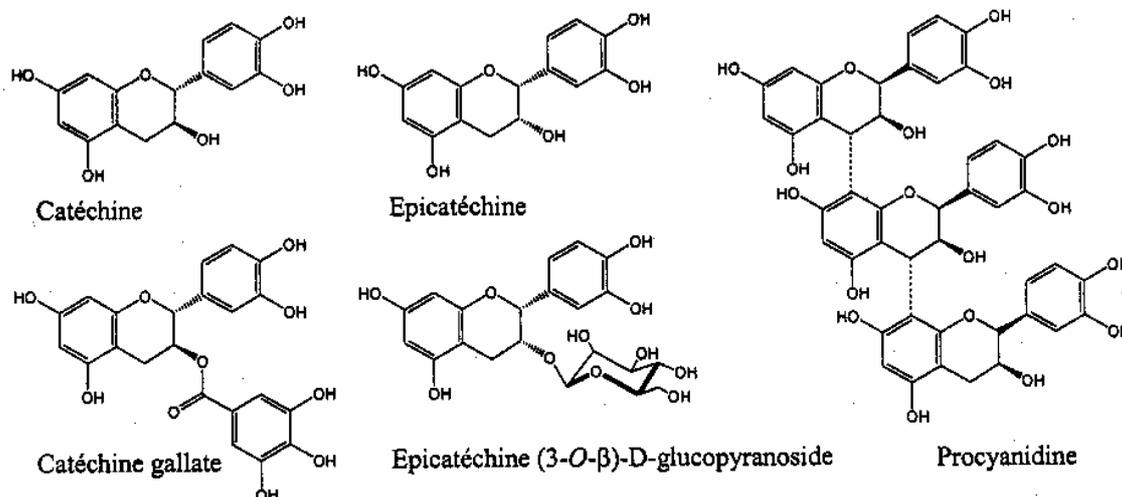


Figure 4 : Complexité structurale des flavanols

Cette grande diversité de structures fait des flavonoïdes l'une des plus grandes familles de substances naturelles connues. En effet, les techniques d'analyse récentes ont permis d'identifier plus de 8000 flavonoïdes [16].

3. Intérêt biologique des flavonoïdes

La recherche sur les flavonoïdes a reçu un écho particulier au sein de la communauté scientifique lors de la découverte du paradoxe français [17,18]. En effet, des études épidémiologiques ont montré qu'une population ayant une alimentation riche en acides gras polyinsaturés présentait un taux de mortalité accru dû aux maladies cardiovasculaires. En revanche, les populations méditerranéennes, bien que présentant ce même type d'alimentation, étaient moins sujettes à ces maladies. L'explication pourrait provenir en partie de la consommation quotidienne de flavonoïdes présents dans les jus de raisin et dans l'huile d'olive. Les flavonoïdes présentent de nombreuses activités : antioxydantes, anti-inflammatoires, inhibitrices d'enzymes, et prévention des maladies cardiovasculaires. Pharmacologiquement, les aglycones sont particulièrement efficaces. Certains ont des activités hépatoprotectrices, diurétiques, vasodilatatrices, antibactériennes, chimoprotectrices, anti-inflammatoires,

antidiabétiques, inhibitrices de l'aldolase réductase et antiallergiques [19].

La majorité des activités biologiques des flavonoïdes est due à leur pouvoir antioxydant et chélateur. Plusieurs études ont montré qu'un régime alimentaire riche en flavonoïdes peut avoir des effets bénéfiques sur la santé [19].

4. Propriétés anticancéreuses des flavonoïdes

Parmi les nombreuses propriétés biologiques des flavonoïdes, nous nous intéressons dans ce travail à leurs propriétés anticancéreuses notamment.

Selon les dernières observations du Fonds Mondial de la Recherche contre le Cancer publiées en 2007, on estime que 30% de tous les cancers sont directement reliés à la nature du régime alimentaire des individus. Ce pourcentage pouvant même atteindre jusqu'à 70% dans le cas des cancers du système gastro-intestinal : œsophage, estomac et colon. De nombreuses études épidémiologiques ont montré que la consommation de fruits et légumes était associée à une baisse importante du risque de développer cette maladie. Dans les études cliniques, les personnes consommant le moins de fruits et légumes sont environ deux fois plus susceptibles de développer certains cancers que celles ayant la plus forte consommation de ces aliments.

Des études sur la composition chimique de certains fruits et légumes ont montré que la richesse de ces végétaux en certains métabolites tels que les flavonoïdes est à l'origine de leur efficacité dans la prévention et le traitement du cancer.

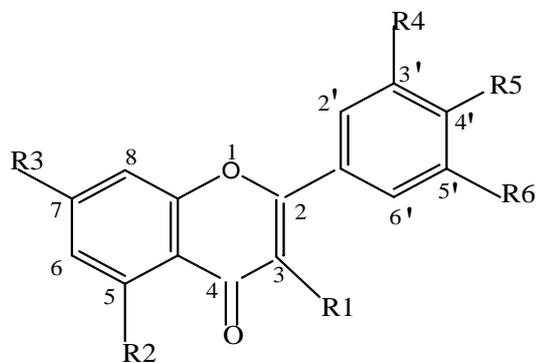
In vitro, les flavonoïdes sont considérés comme des agents antiprolifératifs et cytotoxique vis-à-vis de plusieurs lignées cellulaires cancéreuses. D'autres études menées sur des animaux se sont intéressées à évaluer le pouvoir anticancéreux des flavonoïdes *in vivo* [20].

5. Critères de Choix des flavonoïdes

Dans cette étude, nous nous sommes limités à deux types de molécules les plus représentatifs de la famille des flavonoïdes : les flavones et flavonols. Des études récentes ont montré que la flavone apigénine, et les flavonols fisétine, quercétine et myricétine (Tableau 1) sont des inhibiteurs de l'ADN topoisomérase I humaine [21]. La structure particulière de leur squelette de base leur confère les caractéristiques les plus souvent mises en avant dans l'activité d'un flavonoïde : le cycle catéchol et l'insaturation en position 2 conjuguée avec la fonction carbonyle

en position 4. Pour toutes ces raisons, ces quatre flavonoïdes sont les molécules modèles par excellence pour aborder l'étude des interactions enzyme-inhibiteur à l'aide des méthodes de modélisation par docking moléculaire.

Tableau 1. Structure chimique des flavonoïdes inhibant l'ADN topoisomérase I



Squelette de base des flavones et flavonols

Composés	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
Myricétine	OH	OH	OH	OH	OH	OH
Fisétine	OH	H	OH	OH	OH	H
Quercétine	OH	OH	OH	OH	OH	H
Apigénine	H	OH	OH	H	OH	H

CHAPITRE 2
Généralités sur le cancer

1. Définition et caractéristiques du cancer

Le cancer est une maladie caractérisée par la croissance anarchique et ininterrompue de cellules anormales. Ces cellules dérivent toutes de la même cellule, initiatrice du cancer, qui a acquis certaines caractéristiques lui permettant de se diviser indéfiniment. Au cours de l'évolution de la maladie, certaines cellules prennent l'apparence d'une excroissance locale : une tumeur primitive qui, dans un premier temps, grossit et dissocie les tissus voisins. D'autres cellules forment des cancers « non solides » sous forme de cellules sanguines anormales qui ne sont pas associées à un organe en particulier.

La cellule cancéreuse va au cours de son développement acquérir une série de capacités que ne possèdent pas les cellules normales. En 2000, R. Weinberg et D. Hanahan ont suggéré que l'apparition de cellules cancéreuses passe par l'acquisition d'au moins six propriétés [21] :

- Indépendance vis-à-vis des signaux stimulant la prolifération : Normalement, les cellules ne se divisent que lorsqu'elles reçoivent un stimulus particulier mais dans les tumeurs, les cellules n'ont plus besoin de ce signal à cause de l'expression accrue et continue des gènes de la prolifération cellulaire.
- Insensibilité aux signaux inhibiteurs de la prolifération cellulaire.
- Abolition de l'apoptose ou mort cellulaire programmée. En cas de stress ou d'anomalie ne pouvant pas être éliminée, une cellule normale se suicide en utilisant l'apoptose. Les cellules tumorales ont inactivé tous ces mécanismes pour mieux survivre.
- Capacité proliférative illimitée : le nombre usuel de divisions cellulaires pour une cellule humaine est de 50 à 60, après quoi elle cesse de pouvoir se diviser. Les cellules tumorales continuent de se diviser sans limite visible grâce à l'activité de la télomérase qui est fortement active dans toutes les cellules cancéreuses. Les cellules deviennent immortelles.
- Capacité de susciter l'angiogenèse. Les cellules tumorales (et la tumeur) ont un besoin important en oxygène et en nutriments pour survivre. Elles vont donc stimuler la formation de nouveaux vaisseaux sanguins afin d'alimenter la tumeur.
- Acquisition d'un pouvoir invasif. Les cellules tumorales sont capables de passer à l'intérieur d'un vaisseau sanguin afin d'être transportées dans un autre organe où elles vont générer une seconde tumeur.

De manière générale, les cancers résultent de mutations de l'ADN de certaines cellules, mais des facteurs génétiques et des agents extérieurs comme les cancérogènes physiques (UV et les rayonnements ionisants), chimiques et biologiques (virus du papillome humain, virus de l'immunodéficience humaine, schistosomiase, Hépatite C et Hépatite B) tendent à accroître de manière importante l'apparition des cancers. Le tabagisme, la consommation d'alcool, la consommation insuffisante de fruits et de légumes ainsi que le vieillissement ont une incidence directe avec l'apparition des cancers. De plus, on estime que 5 % à 10 % des mutations génétiques à l'origine d'un cancer sont héréditaires. Plus globalement, c'est la conjugaison de ces différents facteurs qui entraînent l'apparition de cancers.

Les principaux types de cancers sont les carcinomes (touchant l'épithélium, c'est-à-dire une surface composée uniquement de cellules), les sarcomes (touchant les tissus conjonctifs comme les os) et les cancers hématopoïétiques qui sont des cancers des cellules sanguines.

Il existe presque autant de sources de cancer que de tissus dans l'organisme. A ceci s'ajoute le fait que certains cancers sont plus fréquents que d'autres. Ainsi les cancers de la prostate (33 % des cancers chez l'homme) et du sein (30 % des cancers chez la femme) sont parmi les plus fréquents tandis que les cancers de la thyroïde (1 à 5 % des cancers) sont parmi les plus rares [22].

2. Diagnostic et traitements

Aujourd'hui encore, la majorité des cancers sont diagnostiqués à un stade avancé et sont incurables, d'où la nécessité de développer de nouvelles techniques de plus en plus précises pour détecter les cancers le plus précocement possible. Les recherches actuelles s'orientent vers l'imagerie moléculaire avec le développement de sondes capables de détecter les stades précoces de la progression tumorale et de visualiser les métastases de taille inférieure

au millimètre. Ces sondes reconnaissent des processus biologiques spécifiques des cancers et permettent de les imager au moyen de techniques d'imagerie médicale de plus en plus performantes [23].

Les principaux moyens de traiter un cancer sont [24] :

- la chirurgie : elle permet l'ablation d'une tumeur ou des tissus lésés. C'est le principale mode de traitement des tumeurs solides, localisées.

- La radiothérapie : la tumeur est irradiée par des rayons X ou γ afin de détruire les cellules tumorales. Elle est souvent utilisée en association avec d'autres thérapies.
- Les traitements biologiques : ils utilisent des substances endogènes. Par exemple, l'immunothérapie renforce les défenses immunitaires contre les cellules tumorales.
- Les traitements hormonaux : ils agissent sur la vitesse à laquelle les cellules tumorales grossissent, se multiplient ou meurent. C'est une forme particulière de chimiothérapie qui agit sur le système endocrinien et qui est utilisée pour le traitement de cancers hormono-dépendants.
- Les greffes de la moelle osseuse : il ne s'agit pas d'un traitement direct mais elles sont parfois utilisées pour renforcer le système de production des cellules sanguines.
- La chimiothérapie : il s'agit de l'utilisation de substances pharmacologiques pour tuer les cellules tumorales. L'utilisation de la chimiothérapie anticancéreuse dépend du type de cancer à traiter. Plusieurs agents anticancéreux avec des mécanismes d'action différents sont généralement combinés pour augmenter l'efficacité du traitement, et éviter des phénomènes de résistance. L'inconvénient principal des composés utilisés est le manque de sélectivité qui entraîne de nombreux effets secondaires [25].

Les principaux agents chimiothérapeutiques utilisés actuellement peuvent être regroupés en trois catégories selon leur mode d'action [25] :

- les premiers perturbent la synthèse d'ADN, comme par exemple les inhibiteurs des topoisomérases. Ces derniers sont considérés comme les agents chimiothérapeutiques les plus efficaces utilisés en clinique.
- les seconds introduisent des dommages à l'ADN. Citons à titre d'exemple les anthracyclines et la bléomycine qui sont issus de différents champignons du genre *Streptomyces*.
- les troisièmes inhibent les fonctions du fuseau mitotique. Les inhibiteurs des microtubules agissent en se liant à la tubuline, perturbant ainsi la mise en place correcte d'un fuseau mitotique fonctionnel.

L'ADN et les microtubules sont donc les deux cibles essentielles des agents chimiothérapeutiques. Plusieurs dizaines de composés naturels ou synthétiques sont utilisés en chimiothérapie anticancéreuse. Parmi ces agents, plus de 60% sont soit directement isolés de sources naturelles, soit dérivés de produits naturels après modification chimique ou encore entièrement synthétiques mais dont la structure est inspirée de produits naturels.

La diversité des agents utilisés en chimiothérapie anticancéreuse a permis d'augmenter le nombre de rémission et également leur durée mais les traitements restent lourds, difficiles à supporter et présentent de nombreux effets secondaires. C'est pourquoi il est important de trouver de nouvelles molécules plus sélectives sur certains types de cancers.

Ces dernières années, l'évolution des connaissances a permis de comprendre de manière plus approfondie le mécanisme d'action de médicaments déjà utilisés et ainsi, d'envisager de nouvelles perspectives thérapeutiques. [24]

Parmi les différents agents thérapeutiques cités précédemment les inhibiteurs des topoisomérases retiennent d'avantage notre attention. En effet les nombreux rôles exécutés par les topoisomérases en font des cibles privilégiées de certains médicaments déjà utilisés comme antibiotiques ou comme antitumoraux et qui agissent en inhibant l'activité de ces enzymes [26]. Ceci révèle l'importance d'approfondir l'étude biochimique de ces topoisomérases en passant par une meilleure connaissance de leur mécanisme d'action afin de contribuer au développement *in silico* de nouveaux inhibiteurs de cette enzyme.

CHAPITRE 3

Les topoisomérases d'ADN

Les topoisomérases d'ADN

1. Définition

La molécule d'ADN contient les informations génétiques transmises d'une cellule à l'autre, et de génération en génération, grâce aux différentes étapes qui constituent son métabolisme : réplication, transcription et traduction, étapes essentielles du cycle cellulaire. Depuis la découverte de la structure en double-hélice de l'ADN, de nombreuses questions ont été soulevées sur les mécanismes avec lesquels ces fonctions sont rendues possibles.

L'extraordinaire longueur - plusieurs millions de kilo bases chez l'homme ($3 \cdot 10^9$) - la compaction dans les chromosomes ainsi que la structure en double-hélice de la molécule d'ADN conduisent à une entité fortement enchevêtrée présentant un grand nombre de contraintes topologiques : entrecroisements, surenroulements, nœuds (Fig 5).

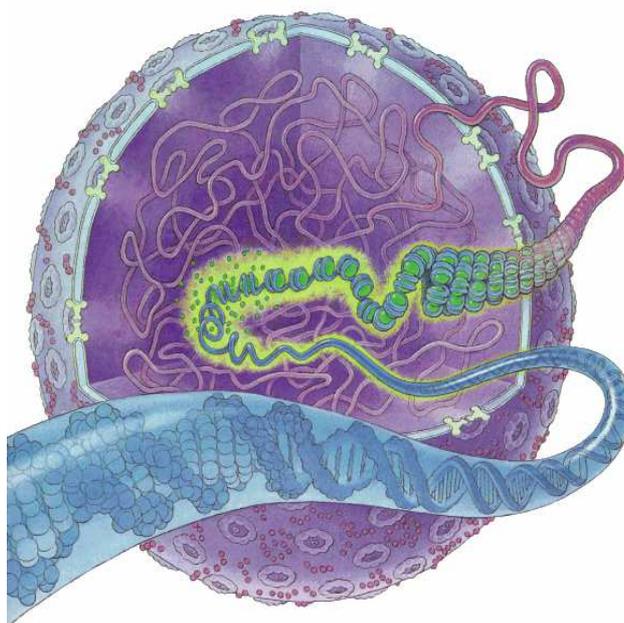


Figure 5. Représentation schématisée de la compaction de l'ADN dans la cellule.

En fait, dans la plupart des génomes, l'ADN se présente soit sous forme circulaire (virus, mitochondries et chloroplastes) soit sous forme de boucles de 103 à 105 paires de bases (pb), attachées par leur extrémités en des points fixes de la matrice nucléaire chez les eucaryotes, ou

du nucléoïde chez les procaryotes. Topologiquement, boucles et molécules circulaires sont équivalentes ; ce sont des molécules fermées, sans extrémités libres et soumises à de fortes contraintes topologiques. Ainsi, les processus qui nécessitent la séparation des deux brins d'ADN demanderaient une rotation sur une longueur impensable de matériel génétique. Cela implique que les deux brins de la double-hélice sont inséparables à moins de générer une cassure sur un ou sur les deux brins [28].

Lorsque deux brins entremêlés d'une double-hélice d'ADN doivent être séparés, au cours d'un des processus biologique, cela entraîne l'augmentation de leur enroulement l'un autour de l'autre plus loin dans la molécule (Fig 6). Ainsi, le mouvement de la fourche de réplication produit un superenroulement positif devant elle, entraînant rapidement une résistance insurmontable pour tout mouvement supplémentaire. Tout changement du nombre d'enlacements nécessite au moins la coupure d'un brin. Grâce à son extrémité libre, un brin peut tourner autour de l'autre, après quoi la coupure est ressoudée. Une telle réaction convertit un topoisomère en un autre.

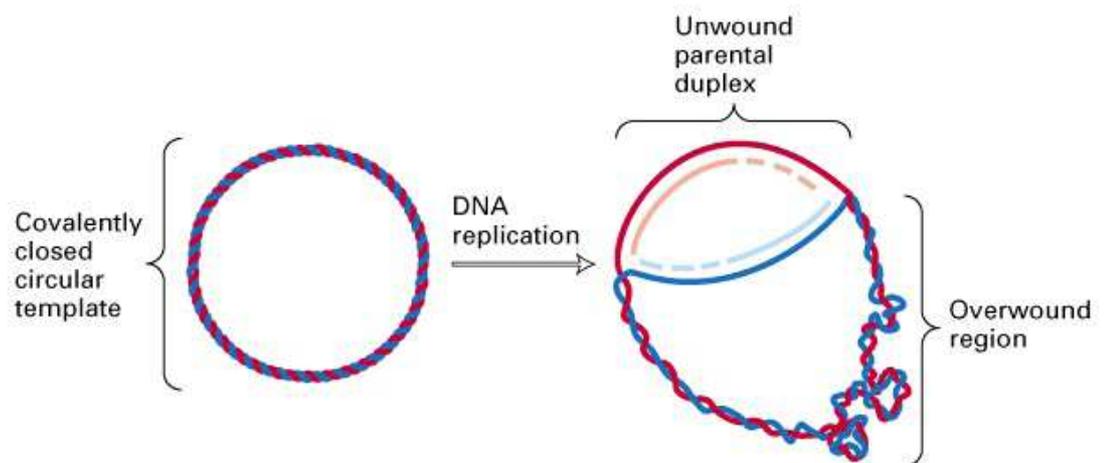


Figure 6. Représentation des surenroulements qui interviennent après séparation locale de deux brins sur un ADN circulaire.

Sur la base de ces considérations, il apparaît évident que la régulation de la topologie de l'ADN est indispensable à la survie des cellules et qu'elle nécessite la présence d'un système enzymatique qui permette de résoudre les contraintes existantes sur l'ADN au cours de son métabolisme. Cette action de contrôle topologique est effectuée par des enzymes nucléaires, les topoisomérases d'ADN, qui jouent un rôle essentiel dans les processus de croissance et de division cellulaire. Ces enzymes catalysent l'interpénétration entre brins d'ADN ou double-

hélices, en générant des coupures transitoires sur l'ADN [29]. De cette façon les surenroulements et les nœuds présents dans la molécule d'ADN peuvent être éliminés et ainsi la réplication, la transcription ou les autres processus du métabolisme de l'ADN, avoir lieu. Grâce à leur capacité de résoudre tous les problèmes topologiques, ces enzymes sont aussi indiquées comme des véritables « magiciens » de l'ADN [30].

Ce fut en 1971 que James Wang découvrit le premier représentant des topoisomérases en purifiant une enzyme à partir d'extraits d'*Escherichia coli* [31]. Wang identifia cette enzyme, appelée protéine ω , comme capable de relaxer les surenroulements négatifs d'une molécule d'ADN circulaire. Suite à cette découverte, une activité de relaxation de l'ADN surenroulé, soit positivement, soit négativement, fut identifiée dans des extraits de cellules de souris [32]. Puis une enzyme ATP-dépendante, capable d'introduire des surenroulements négatifs sur un ADN relaxé, fut identifiée chez *E. coli* par Gellert et al [33]. Ces trois topoisomérases, les premières en termes historiques, sont des représentantes de trois différentes sous-classes de topoisomérases.

Les études successives ont élargi la famille des topoisomérases grâce à la découverte de nouveaux membres, qui ont permis de classifier les différentes topoisomérases et de conclure que ces enzymes participent à pratiquement chaque processus cellulaire impliquant l'ADN. En effet, au cours de la réplication et de la transcription, les topoisomérases permettent d'éliminer les surenroulements, qui peuvent s'accumuler en amont par rapport à l'avancement de la fourche de réplication sur la double-hélice d'ADN. Au cours de l'assemblage, de la condensation et de la ségrégation des chromatides, les topoisomérases sont essentielles afin de permettre le correct déroulement de ces procès [34, 35]. Le rôle des topoisomérases dans la recombinaison et la réparation de dommages sur l'ADN a aussi été démontré [36].

Dans ce contexte, il est évident que la régulation de la topologie, et ainsi de la superhélicité, est indispensable au cours de chaque étape du métabolisme de l'ADN [28].

2. Classification des topoisomérases

Ces enzymes contrôlent et modifient la topologie de l'ADN d'une façon simple, et toutefois élégante, grâce à une réaction de transestérification. Ce mécanisme, postulé la première fois par Wang en 1971 [31] (Fig7), permet la coupure et la religation de l'ADN par attaque d'un hydroxyle situé sur un résidu de tyrosine de l'enzyme (la Tyr₈₀₅ pour la topoisomérase II humaine et la Tyr₇₂₃ pour la topoisomérase I humaine) sur un atome de phosphore du squelette

phosphodiester de l'ADN, ainsi générant une nouvelle liaison phosphodityrosyl et une coupure sur l'ADN. La religation se réalise grâce à une deuxième transestérification qui est essentiellement la réaction inverse de la première. Ainsi, l'oxygène de l'hydroxyle généré sur l'ADN au cours de la première réaction attaque le phosphore de la liaison phosphotyrosyl et cause la coupure de la liaison covalente entre l'enzyme et l'ADN. Le squelette de l'ADN est ainsi reconstitué [30].

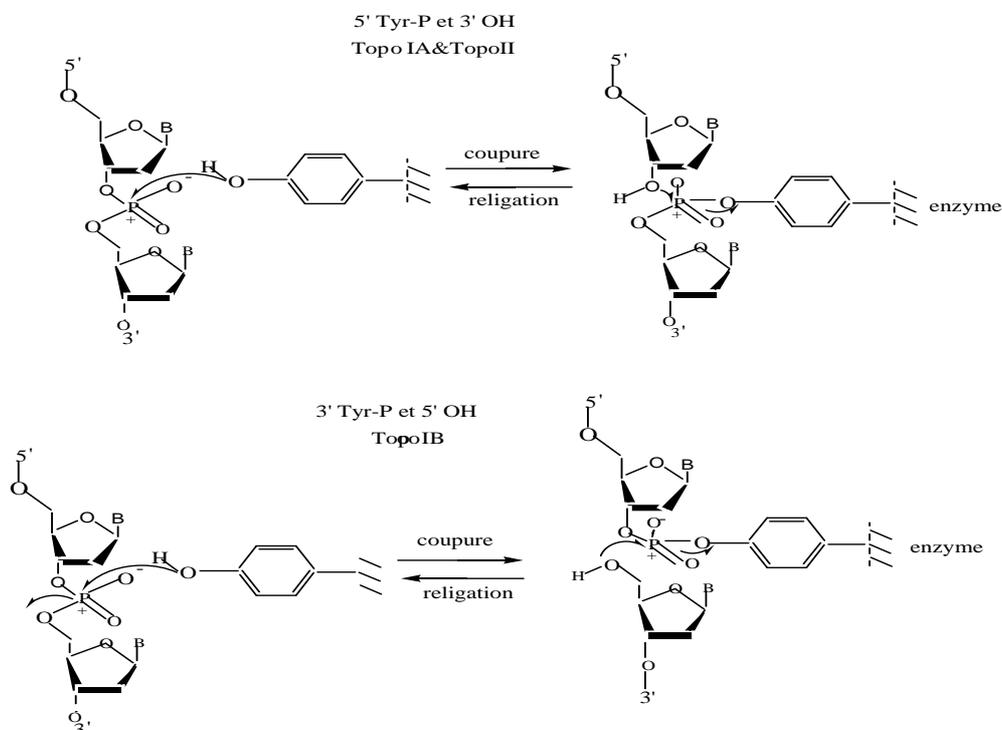


Figure 7. Réactions de transestérification déterminant la coupure de l'ADN commune à toutes les topoisomérases. La réaction inverse permet la religation de l'ADN coupé.

Les topoisomérases se divisent en deux classes principales, selon la nature des mécanismes qu'elles mettent en œuvre. Les enzymes de type I introduisent une coupure transitoire dans un des brins d'ADN. Les enzymes de type II introduisent une coupure transitoire double-brin. Les topoisomérases I et II ont été ultérieurement classifiées en sous-classes sur la base des différences structurales et de leur mécanisme d'action [29,37]. Les topoisomérases de type IA, découvertes par Wang en 1971 [31], déterminent la formation d'un hydroxyle libre en 3' du squelette phosphodiester de l'ADN par formation d'une liaison phosphotyrosyl avec le phosphate en position 5' (Fig 7, haut). Inversement, les topoisomérases de type IB identifiées pour la première fois par Champoux et Dulbecco [32] génèrent un hydroxyle libre en position 5' par formation de la liaison phosphotyrosyl avec le phosphate en position 3' du squelette phosphodiester (Fig 7, bas). Enfin, les topoisomérases de type II, isolées pour la première fois par Gellert et *al* [33] comme les topoisomérases IA, catalysent la formation d'une liaison

phosphotyrosyl avec le phosphate en position 5', mais cette fois sur les deux brins d'ADN.

Dans ce présent travail nous nous sommes focalisés sur les topoisomérases I qui se présente comme cibles privilégiées dans les traitements anticancéreux.

3. Topoisomérases I

3-1. Structure

L'ADN topoisomérase I humaine (topoI) est une protéine monomérique de 765 acides aminés [38] Elle est constituée de quatre régions majeures : le domaine terminal NH₂, le cœur, le domaine de liaison et le domaine terminal COOH (fig8). Le domaine terminal NH₂, d'environ 210 résidus, est hautement chargé, contient très peu d'acides aminés hydrophobes et est grandement désordonné [39]. Le domaine terminal COOH, d'environ 50 résidus, contient l'essentiel Tyrosine 723 qui forme une liaison phosphoester avec le phosphate 3' du brin d'ADN coupé du site de clivage [40]. Le cœur est constitué d'environ 430 résidus et le domaine de liaison, qui n'est pas requis pour l'activité de l'enzyme, de 70 résidus [41].Topo I apparait donc comme une enzyme multi-domaines comportant :

- Deux domaines globulaires hautement conservés (le cœur et le domaine terminal COOH) qui sont cruciaux pour l'activité catalytique.
- Deux régions (le domaine terminal NH₂ et le domaine de liaison) qui ne sont pas indispensable aux fonctions catalytiques et de relaxation de l'enzyme.

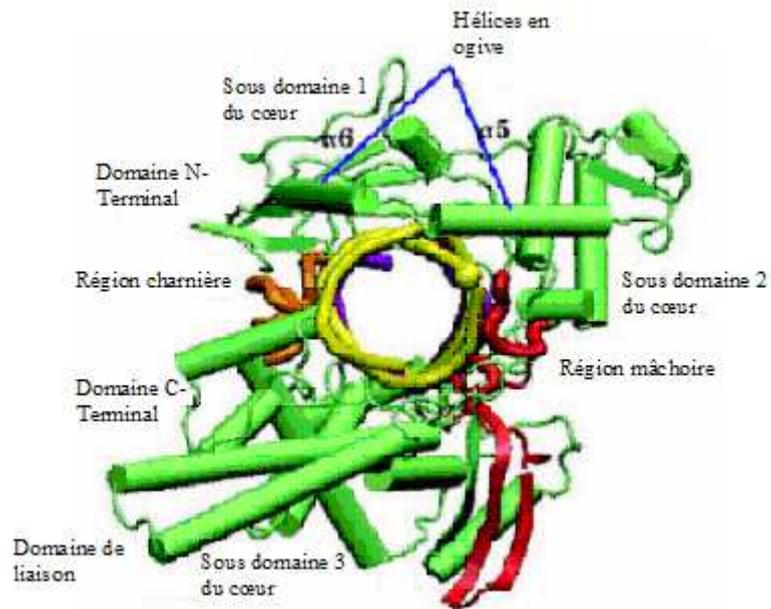


Figure 8 : Structure de l'ADN topoisomérase I humaine

3-2. Mécanisme enzymatique

Le cycle catalytique de la topoisomérase I humaine est divisé en quatre étapes : une étape de liaison de l'enzyme à l'ADN, une étape de clivage introduisant la coupure, une étape de rotation du brin coupé autour du brin non coupé, et une étape de religation permettant de restaurer la continuité du brin d'ADN [42, 43] (Fig 9).

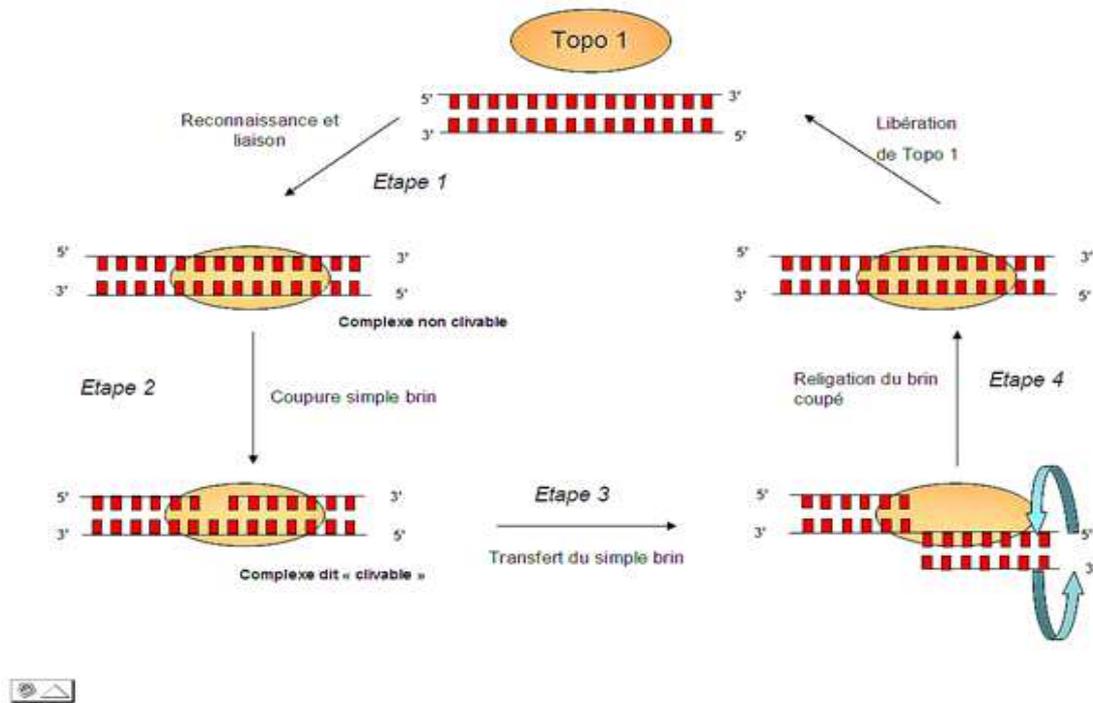


Figure 9. Schéma du mécanisme d'action de l'ADN topoisomérase I

Liaison de l'enzyme à l'ADN

Topo I se lie uniquement à l'ADN double brin de manière non covalente. Cette liaison apparaît comme étant plus efficace au niveau des régions coudées ou sur-enroulées de l'ADN. [44, 45]

Clivage de l'ADN

Le clivage de l'un des brins d'ADN, suivi de l'attachement de manière covalente de topo I à l'extrémité 3' de l'ADN coupé, conduit à la formation d'un complexe covalent : le complexe dit "clivable" [46, 47].

Relaxation de l'ADN

Plusieurs hypothèses sont émises quant aux mécanismes de relaxation de l'ADN, après la formation du complexe covalent et avant la religation de l'ADN. Parmi ceux-ci figurent "le passage simple-brin" et "le modèle de rotation libre".

Le mécanisme de passage simple brin propose que le brin non coupé passe à travers un espace formé entre la liaison covalente de l'extrémité 3' et la liaison non covalente de l'extrémité 5' du brin coupé avec l'enzyme. Dans ce cas, un seul événement de désenroulement peut être réalisé lors d'un cycle de clivage/reliation [48].

Le modèle de rotation libre suppose que l'extrémité 5' du brin coupé soit libérée du site actif de l'enzyme et puisse tourner librement autour du brin complémentaire non coupé. Ceci sous-entend que de multiples évènements de désenroulement peuvent être réalisés pour chaque cycle de clivage/religation [49].

Cependant, un mécanisme intermédiaire a été proposé par Stewart et coll en 1998, sous le nom de " rotation contrôlée " (fig 10), dans lequel les interactions ioniques entre l'ADN et certaines régions de l'enzyme régularaient le processus de désenroulement [43]. Dans ce cas, les surtensions de l'ADN seraient libérées par un ou plusieurs cycles de rotation contrôlée.

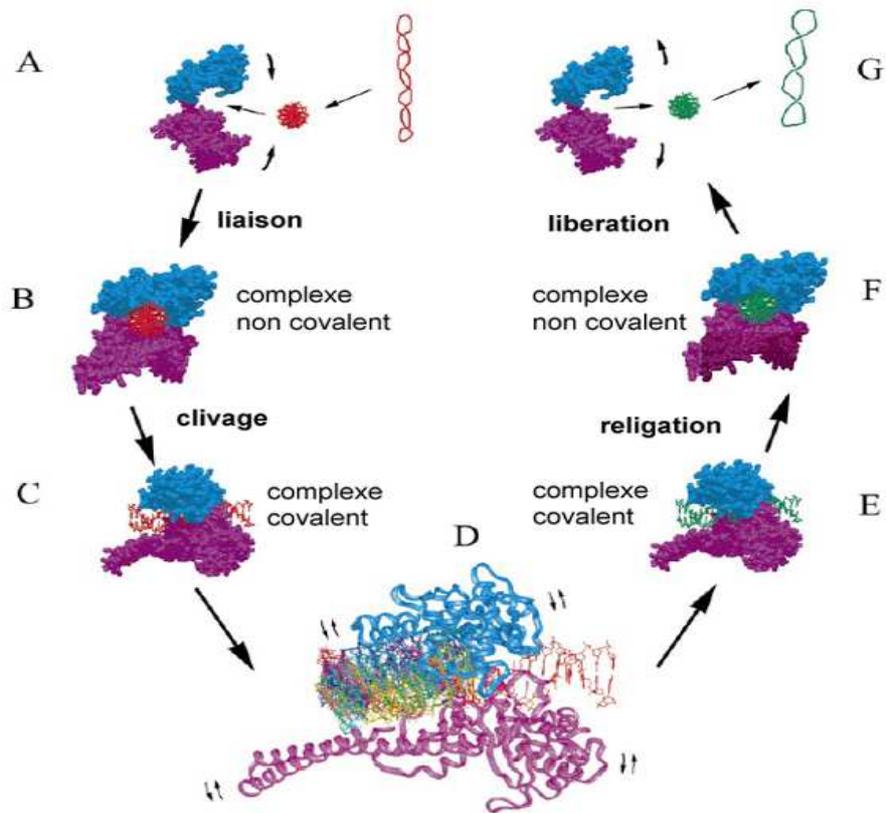


Figure 10. Mécanisme de rotation contrôlée de topoI. L'ADN hautement surenroulé de tours négatifs (en rouge) devient à la fin du processus moins surenroulé (en vert).

Religation et libération de topoI

La religation se fait par une deuxième réaction de transestérification où les deux brins clivés de l'ADN sont reliés et topoI libérée.

Sélectivité de séquence

Il est connu que les sites de clivage de l'ADN par topoI ne sont pas distribués au hasard au

sein des fragments d'ADN, signifiant que topoI reconnaît des séquences spécifiques.

Neufs nucléotides sur le brin coupé et cinq sur le brin non coupé s'avèrent être essentiels pour une forte liaison de topoI à l'ADN [50, 51].

3-3. Fonctions biologiques

Elles sont diverses et variées au niveau de l'ADN. Durant la réplication, des super-tours positifs s'accumulent dans la portion non répliquée de l'ADN, créant ainsi des forces de torsion. En absence de topoI, le mouvement de la fourche de réplication est alors arrêté. Cet effet est pallié par l'activité de coupure/reliation de topoI.

TopoI est impliqué dans la transcription de certains gènes. Les sites de liaison et de clivage de topoI ont été cartographiés le long de régions d'ADN activement transcrites comme les gènes des ARN polymérase I et II [52, 53]. Topo I agit comme un pivot pour soulager les contraintes de torsion causées par la génération de super-tours positifs en amont et de super-tours négatifs en aval du mouvement de l'ARN polymérase [54, 55].

4. Les inhibiteurs de l'ADN topoisomérase I humaine

La topoisomérase I est la cible de nombreux inhibiteurs, qui empêchent l'enzyme d'accomplir entièrement son cycle catalytique et provoquent des dommages irréversibles de l'ADN, ainsi qu'un arrêt du cycle cellulaire. Parmi les inhibiteurs de topoisomérases, on retrouve certains médicaments les plus actifs contre le cancer utilisés en clinique.

Les inhibiteurs de la topoisomérase I sont divisés en deux groupes principaux : les poisons et les supprimeurs. Ils inhibent tous les deux l'activité catalytique de l'enzyme, mais selon des modes d'actions différents [56].

4-1. Les poisons

Ils tuent les cellules en stabilisant les complexes de clivage ADN-topoI, ceci empêchant la reliation des brins coupés d'ADN et la libération de topo I. (Fig 11)

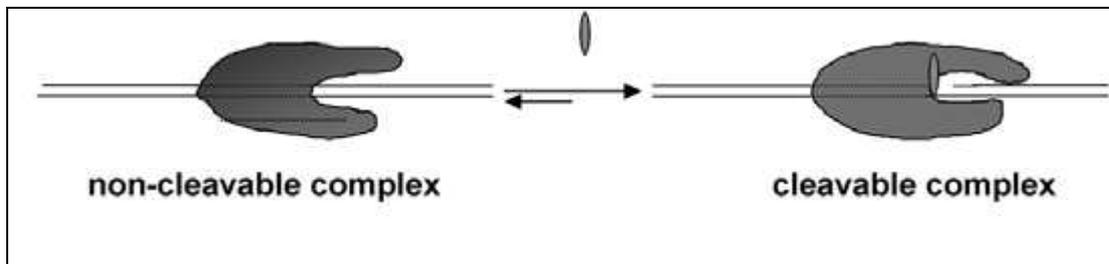


Figure 11. Mécanisme d'action d'un poison de l'ADN topoisomérase I.

Il a été montré que la sensibilité des cellules aux poisons augmentait avec la surexpression de topoI [57] et qu'une réduction de l'activité de topoI est une cause de résistance des cellules aux poisons [58, 59].

Les grandes classes de poisons sont les camptothécines et ses dérivés " irinotecan, topotecan " et les benzo[c]phénanthridines, les benzoanthracènes, les protoberbérines, des composés hétérocycliques comme les indoloquinolinediones, les indolacarbazoles, les benzimidazoles [60, 61]. Dans des études récentes, des nouveaux composés ont été prouvés comme des inhibiteurs de l'ADN topoisomérase I humaine : les dérivés du 6-Arylamino-7-chloro-quinazoline-5,8-diones [62] et le 2,6-Dithienyl-4-furyl pyridines [63].

Camptothécine, l'irinotecan, topotecan, sont largement utilisées dans le domaine clinique pour le traitement des cancers, et particulièrement dans le cancer du poumon, le cancer de l'ovaire et le cancer colorectal [64].

4-2. Les supprimeurs

Ils agissent en empêchant la reconnaissance de topoI à son site de fixation, inhibant ainsi la formation des complexes de clivage.

En générale, les ligands de l'ADN, incluant les intercalants comme la doxorubicine, l'aclacinimycine A, et la distamycine A peuvent supprimer la liaison de topoI à l'ADN. Cet effet est observé à des concentrations de drogue relativement élevées et apparaît directement lié à l'affinité de liaison à l'ADN. De ce fait, la drogue se liant à l'ADN empêche l'accès à topoI et/ou distord la structure de l'ADN [56].

CHAPITRE 4

Modélisation des interactions Protéine-Ligand par docking moléculaire

1. Modélisation moléculaire

La modélisation moléculaire est le terme générique regroupant les méthodes qui permettent de simuler le comportement d'un système de particules. La taille du système étudié peut aller d'une simple molécule diatomique jusqu'à des macromolécules biologiques de plusieurs dizaines

de milliers d'atomes.

La modélisation moléculaire peut ainsi permettre de décrire le comportement électronique des atomes et des molécules pour expliquer leurs réactivités, comprendre les processus de repliement d'une protéine ou encore expliquer l'importance de certains acides amines d'un site catalytique enzymatique. Pour réaliser ce type d'étude, il est nécessaire de déterminer une expression de l'énergie d'interaction des atomes du système moléculaire en fonction de leurs positions relatives.

L'association non covalente et réversible d'un récepteur (R) et d'un ligand (L) pour former un complexe récepteur-ligand (RL), dans un milieu aqueux contenant des électrolytes peut être décrit par l'équation 1 [65] :



Equation 1

À l'équilibre, cette réaction est déterminée par l'énergie libre de liaison ΔG° . Cette quantité est liée (Equation 2) aux constantes d'association (K_A) et de dissociation (K_D), qui peuvent être déterminées expérimentalement (Equation 3). Elle contient une composante enthalpique (ΔH°) et entropique ($T\Delta S^\circ$) (T étant la température absolue) [65].

$$\Delta G^\circ = -RT \ln K_A = \Delta H^\circ - T \Delta S^\circ$$

Equation 2

$$K_A = K_D^{-1} = [RL] / [R][L]$$

Equation 3

3. Reconnaissance moléculaire

Au niveau du site actif, le processus d'interaction se fait par liaison complémentaire du ligand avec l'architecture protéique. Les forces intervenant dans la formation du complexe peuvent être de nature stérique, électronique ou les deux avec des contributions différentes.

Les états conformationnels choisis lors de l'interaction ne sont pas nécessairement ceux de plus basse énergie en solution, mais des états transitoires stabilisés par celle-ci. Les complexes

formés sont alors plus stables que les formes dissociées. Cette association intermoléculaire est stabilisée par plusieurs effets et des liaisons non covalentes. Au niveau du processus biochimique, elles s'expriment par différentes interactions : électrostatiques, hydrophobiques, de Van Der Waals et des ponts hydrogène [15].

2-1. Interactions de Van Der Waals

Les interactions de Van Der Waals sont des dipôles non permanents de faible rayon d'action (fig 12). Elles sont nombreuses et contribuent essentiellement dans la recherche de la concordance stérique entre le ligand et la protéine réceptrice.

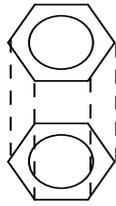


Figure 12. Interactions de Van Der Walls

2-2. Ponts hydrogène

Les ponts hydrogène avec une force de l'ordre de 2 à 4 Kcal/mol agissent dans les systèmes biologiques, à très courte distance (0,8 à 2,8 Å) entre un H lié à un O ou un N avec un O ou un N. Ils sont peu nombreux et s'adaptent très bien à la flexibilité (l'angle peut varier de 120° à 180°). (fig13)

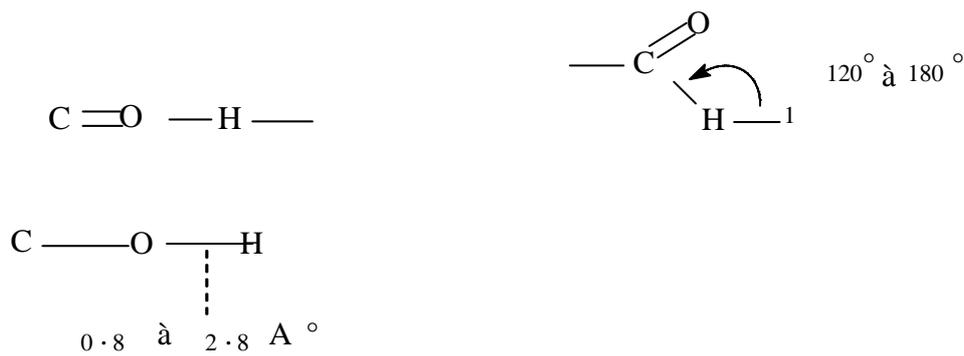


Figure 13 : ponts hydrogène

2-3. Interactions électrostatiques

Les interactions électrostatiques agissent à plus grande distance entre dipôles chargés (fig14). La distance pour une liaison saline, par exemple, est de 3,7 à 4,5 Å, et sa force est de

l'ordre de 8 Kcal/mol. En présence de solvant, ces interactions sont favorisées du fait de la polarité et de la petite taille de la molécule d'eau.

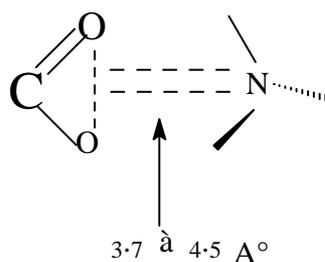


Figure 14 : Interactions électrostatiques

3. Amarrage moléculaire (ou docking)

Les interactions protéine-protéine et protéine-ligand jouent un rôle clé dans l'organisation des systèmes biologiques. Elles permettent la régulation de certains processus biologiques, la transmission des signaux ou encore la catalyse de diverses réactions biochimiques. Connaître la façon dont les protéines interagissent avec d'autres entités biochimiques est une étape essentielle pour comprendre les processus biologiques dans lesquels elles sont impliquées [66].

Différentes méthodes d'étude expérimentale telle que la cristallographie aux rayons- X et la résonance magnétique nucléaire (RMN) permettent généralement de déterminer la réponse à cette question. Cependant, d'une part ces méthodes ne sont pas toujours simples à mettre en œuvre, et d'autre part certaines molécules ont une durée de vie si courte qu'une étude *in vitro* est tout bonnement impossible. Or, pour comprendre les mécanismes qui permettent aux protéines d'assurer leurs fonctions, il faut tout d'abord comprendre comment elles interagissent entre elles ou avec d'autres molécules. D'où l'intérêt d'avoir des méthodes prédictives procédant par simulation, complémentaires à ces approches expérimentales [67]. A l'heure actuelle la simulation par amarrage moléculaire plus connue sous l'appellation de docking moléculaire représente l'approche prédictive la plus utilisée dans ce domaine de la recherche médicale et pharmaceutique.

3-1. Les différents types de docking moléculaire

Le docking moléculaire est l'étude *in silico* des différents mécanismes et interactions qui interviennent lorsque plusieurs structures moléculaires s'assemblent, l'objectif étant de parvenir à déterminer comment ces molécules vont s'agencer les unes par rapport aux autres (fig 15).

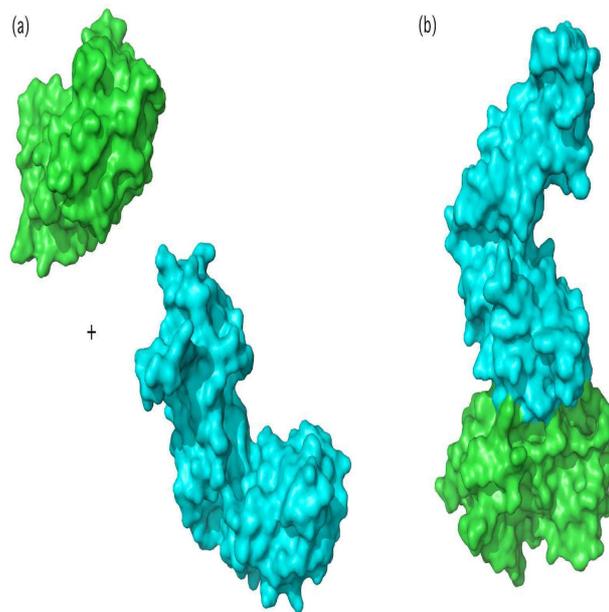


Fig. 15 – Représentation schématique du docking de deux molécules. (a) Molécules de départ. (b) Complexe moléculaire

Il existe deux grands types de docking moléculaire. Le docking rigide consiste à obtenir la conformation préférentielle d'un système protéine-ligand en considérant que chacune des deux molécules conserve une géométrie interne fixe. Dans ce cas, la relaxation de la géométrie interne de chaque entité, en interaction dans le complexe, n'est pas prise en compte. Cependant, il est tout à fait concevable que les structures de la protéine et du ligand soient modifiées durant le processus de docking moléculaire afin d'optimiser au mieux l'interaction entre les deux entités. On parle, dans ce cas, de docking flexible [66].

3-2. Principe de fonctionnement

Le docking moléculaire peut se rapporter à un problème d'optimisation selon une valeur numérique qui rendrait compte de la conformation plus ou moins favorable de deux entités (e.g. l'énergie libre totale du système protéine-ligand). Pour cela le docking moléculaire s'accomplit au deux étapes complémentaires durant la première étape, le programme tend à rechercher les conformations du ligand aptes à établir des interactions idéales avec le récepteur. La deuxième étape, appelée score de docking permet d'évaluer ces conformations par un calcul rapide de l'énergie libre totale du système protéine-ligand formé [66].

La capacité d'un programme à réussir ce travail est habituellement jugé au moyen de la déviation quadratique moyenne ou RMSD (root-mean-square derivation) du modèle conçu par le

logiciel vis-à-vis de la structure du cristal. Le ratio admis est une différence maximum de 2 angströms au-delà de laquelle la prédiction est considérée comme non pertinente [15].

Calcul du RMSD

Les différentes conformations issues de la dynamique sont analysées et comparées entre elles par des superpositions et des calculs de déviations (RMSD) (Equation 4) [68].

$$RMSD = \sqrt{\frac{1}{N} \sum (\overline{r_{i1}} - \overline{r_{i2}})^2}$$

Equation 4. Expression du RMSD

Avec: N le nombre d'atomes du système étudié

$\overline{r_{i1}}$ le vecteur position de l'atome i dans la structure 1

$\overline{r_{i2}}$ le vecteur position de l'atome i dans la structure 2

3-3. Le processus de scoring

Le score est une donnée numérique utile pour quantifier le degré avec lequel un ligand se complexe à un récepteur. C'est globalement une approximation de l'énergie libre résultant du passage de la forme libre de la protéine et du ligand à l'association sous forme de complexe. Le principe thermodynamique est le suivant [68] :

$$\Delta G = \Delta G_{\text{complexe}} - \Delta G_{\text{ligand}} - \Delta G_{\text{protéine}}$$

Concrètement, le score est une estimation de l'affinité entre la macromolécule et la petite molécule organique. Un score ne prédit donc en rien une activité mais bien une affinité.

Un grand nombre de programmes (commerciaux ou non) de docking moléculaire sont d'ores et déjà disponibles. Parmi ceux-ci, nous citerons par exemple AUTODOCK [69], FLEXX [70], GOLD [71], DOCK [72] et SURFLEX [73] etc. Ils diffèrent les uns des autres sur

la manière de représenter le système moléculaire et la manière de déterminer le score de docking (fonction de score). Deux approches sont principalement employées pour la modélisation du système protéine-ligand.

3-3-1. Représentation géométrique

Cette première approche a conduit au développement de méthodes géométriques dites de complémentarité de forme (*shape complementarity*) où la protéine et le ligand sont décrits par des surfaces de descripteurs qu'il suffit d'apparier (*matching*). La construction de la surface de descripteurs de chacune des deux entités prend en compte différentes caractéristiques telles que la surface moléculaire, la surface accessible au solvant, les propriétés hydrophobes, la formation éventuelle de liaisons hydrogène, etc. Dans cette approche, la fonction de score mesure un taux d'appariement correct des descripteurs entre la protéine et le ligand pour chaque conformation à évaluer. Plus ce taux est grand, plus le score de docking est élevé et plus la conformation du système protéine-ligand est optimale. Ces méthodes sont typiquement rapides et robustes. Elles sont choisies préférentiellement, par exemple, dans le cas d'un scan de plusieurs milliers de ligands pharmacophores [66].

3-3-2. Représentation atomique

La seconde approche repose, quant à elle, sur une simulation du processus de docking avec une description atomique du système protéine-ligand. La protéine et le ligand sont, au début, éloignés à une certaine distance et le ligand doit, au fil de la simulation, trouver sa position optimale dans le site actif de la protéine après un certain nombre de déplacements dans l'espace configurationnel. Ces déplacements correspondent aux translations et aux rotations du ligand, auxquelles on ajoute les changements de géométrie interne du ligand et de la protéine (principalement des angles de torsion) si on considère un docking flexible. Le traitement explicite de la flexibilité est un avantage indéniable de cette approche par rapport à la précédente. Cette approche s'apparente aussi d'avantage à un problème de reconnaissance moléculaire.

Signalons qu'il existe des fonctions de score liées au traitement spécifique d'un type de docking tel que protéine-ligand, protéine-protéine, ou protéine-ADN. Ces fonctions calculent généralement le score de docking à partir d'une méthode paramétrée ou empirique. Nous avons regroupé ces méthodes selon trois grandes catégories [66] :

- Champ de forces : Le score est calculé en faisant la somme des forces intermoléculaires de van der Waals et électrostatiques entre tous les atomes du système protéine-ligand. Les

énergies intramoléculaires des deux entités sont parfois incluses dans le score final.

- Empirique : Ces méthodes se basent sur un décompte du nombre d'interactions dans le système protéine-ligand comme par exemple le nombre de liaisons hydrogène, les contacts hydrophobes, les contacts hydrophiles, les contraintes de torsion dues à la formation du complexe protéine-ligand, etc. Le score est finalement obtenu, en général, en utilisant des méthodes de régression linéaire multiple.

- Connaissance : Ces méthodes sont fondées sur des observations statistiques de contacts intermoléculaires proches dans des bases de données 3D (*Cambridge Structural Database* ou *Protein Data Bank*). Il est supposé que les contacts intermoléculaires proches entre certains types d'atome (ou entre certains groupes fonctionnels) plus fréquents qu'une distribution aléatoire conduisent à des interactions énergétiquement favorables, permettant alors d'estimer une affinité de liaison reliée à un potentiel de force moyenne. Cette valeur estimée de l'affinité de liaison est le score de docking.

En plus de la conformation optimale, l'objectif d'un programme de docking moléculaire est idéalement de déterminer l'enthalpie libre (ΔG) de formation en solution du complexe [P : L] entre la protéine P et le ligand L, selon:



Quelle que soit l'approche utilisée, les fonctions précédentes nous donnent le score de docking sous la forme d'une valeur, souvent abstraite, représentative de la conformation obtenue du système protéine-ligand. Cette valeur n'est pas reliée directement à une grandeur énergétique du système, sauf dans le cas de l'utilisation d'un champ de forces où le score est donné par les énergies d'interaction intermoléculaire. Des calculs supplémentaires sont donc nécessaires, à partir des conformations optimales, pour déterminer la valeur de l'enthalpie libre de formation[74].

4- Le programme de docking moléculaire Surflex

Pour réaliser le présent travail de recherche, nous avons utilisé le programme Surflex. Dans ce programme le docking se base sur la construction d'une pseudo-molécule comme cible sur laquelle doit l'aligner le ligand. Il s'agit de fragments de molécules placés dans le site actif de manière à occuper idéalement et de manière redondante celui-ci, à partir de critères

morphologiques. Il est important de préciser que cette pseudo molécule peut être construite à partir du ligand dans le site à condition de bénéficier d'une structure, ou bien à partir du récepteur sans aucun ligand : il utilise dans ce cas 3 types de fragments CH₄, C=O et N-H mais il est nécessaire de connaître l'emplacement du site actif.

L'étape de docking s'effectue suivant deux méthodes :

- la première est une méthode incrémentale dite Hammerhead. Dans ce cas le programme fractionne le ligand et cherche à trouver le meilleur appariement pour chaque fragment, il essaye d'adapter les fragments à ceux de la pseudo-molécule. Les fragments ayant la plus haute affinité sont utilisés pour reconstruire le ligand. À la différence de la recherche incrémentale de Dock il n'y a pas de recherche systématique des angles de torsion mais une étape de minimisation entre chaque ajout de fragment.

- la deuxième méthode dite "molécule entière" reprend la même étape de fractionnement que précédemment. Le changement se fait au niveau des fragments du ligand conservés. Dans la méthode Hammerhead ne sont conservés pour l'étape de docking que les meilleurs fragments qui s'alignent avec ceux de la pseudo-molécule. Ici tous les fragments sont pris en compte pour la recherche de configuration optimale du ligand.

Ainsi le résultat des configurations retenues pour s'apparier avec le ligand contient des molécules proches de l'alignement initial du ligand qui ont été localement modifiées par le processus de fragmentation. Les fragments de ces configurations sont évalués puis une étape de sélection et de fusion des meilleurs fragments entre configurations permet d'obtenir un « conformère » global. Cette deuxième méthode se montre beaucoup plus rapide que la méthode Hammerhead pour des résultats au moins équivalents.

Dans la version de Surflex que nous avons utilisée (version 1.3), il s'agit de la méthode par défaut.

Pour l'utilisateur, l'emploi du logiciel passe par trois étapes :

- Choisir de quelle manière définir le site actif, soit à partir d'un ligand soit à partir du récepteur.
- Construire la « pseudo-molécule » qui sera la cible de la recherche de similarité.
- Enfin lancer le processus de docking.

Le résultat est donné sous la forme des 10 meilleurs « conformères » au format mol2.

Chaque fichier possède trois scores : le premier correspond à l'affinité ; le second correspond au degré de pénétration impropre du ligand dans la protéine appelé «crash score», plus ce score est proche de 0 plus l'interaction est favorable et le troisième ou «polar score» correspond au niveau de contribution des interactions polaires. Plusieurs options permettent d'adapter le calcul suivant le nombre de torsions ou suivant l'espace conformationnel à représenter avec la «pseudo-molécule » [75].

Etude pratique

I.

II.

I. Matériels et méthodes

1. Préparation de la protéine

protein Data Bank ou PDB est une collection mondiale de données sur la structure tridimensionnelle (structure 3D) de macromolécules biologiques : protéines essentiellement, et acides nucléiques. Ces structures sont essentiellement déterminées par cristallographie aux rayons X ou par spectroscopie RMN. Dans notre étude nous avons utilisé le complexe 1T8I disponible dans la PDB contenant l'ADN topoisomérase I humaine et le ligand [76], le complexe

a été téléchargé au format pdb après l'introduction de son code. Les deux molécules du complexe ont été séparées dans des fichiers différents. Le programme Surfex exige le format.mol2, les molécules sont donc protonées, et transformées dans le format.mol2 par le programme OpenBabel 2.3.0.

2. Préparation des flavonoïdes

Les structures des quatre molécules retenues pour le docking "quercétine, myricétine, fisétine et apigénine" ont été construites avec le programme de construction moléculaire Titan, ils sont enregistrés sous format pdb et en suite sont convertis en format.mol2. Le programme Titan dispose d'une banque d'atomes sous différents états d'hybridation permettant de construire tous les groupements chimiques éventuels. Ce type de construction ne peut se faire sans une étape de minimisation importante de l'énergie potentielle, durant laquelle le choix du champ de force est primordial. Pour les flavonoïdes, l'optimisation a été réalisée par la méthode AM1 ou Austin Model1, c'est une méthode de calcul de chimie quantique développée par M. Dewar en 1985 [77]. Le modèle AM1 est basé sur une approche semi-empirique, dont l'objectif est d'optimiser la géométrie de la molécule et de déterminer l'énergie et la distribution électronique.

3. Docking

Le docking a été réalisé en utilisant le programme Surfex 1.3 [73]. Trois étapes sont nécessaires :

Identification du site actif par construction d'une pseudo-molécule appelée "protomol". Cette étape est réalisée grâce à la commande suivante :

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

Docking du ligand

L'étape de docking débute par une fragmentation préalable du ligand. L'assemblage des fragments obtenus permet de rechercher les meilleures conformations du ligand dans le site actif de l'enzyme.

La réalisation de cette étape nécessite la commande suivante :

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

Traitement des résultats

Les résultats pour les dix meilleurs conformères sont fournis dans un même fichier. Chaque conformère possède trois scores : le premier correspond à l'affinité, le second au degré de pénétration impropre du ligand dans la protéine et enfin le troisième correspond au niveau de contribution des interactions polaires.

Le programme Surfex est dépourvu de la visualisation c'est pour cette raison nous avons utilisé le programme ViewerLite 4.2 afin de visualiser les différentes interactions formées entre les flavonoïdes étudiés et le site actif de l'ADN topoisomérase I.

II. Résultats et discussions

1. Test de fiabilité du programme Surfex

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 jugé au moyen de la déviation quadratique moyenne ou RMSD de la position du ligand calculée par le logiciel par rapport au ligand de référence existant au niveau de la PDB. La prédiction est acceptable si sa valeur ne dépasse pas 2 angströms [78, 79].

Le test de fiabilité du programme Surfex par le RMSD, a été réalisé en utilisant 134 complexes protéines-ligands pris de manière arbitraire de la PDB. Les complexes sont téléchargés au format pdb après introduction dans la PDB le code correspondant à chaque complexe.

Après séparation de la protéine et du ligand dans des fichiers différents, ces derniers sont convertis du format pdb au format mol2 grâce au programme OpenBabel.

Une fois la reconstruction du complexe protéine-ligand est réalisée par docking moléculaire, Surfex calcule le RMSD en comparant la géométrie du meilleur conformère (final-0) au ligand d'origine issu directement de la PDB.

Le RMSD est calculé en utilisant la commande suivante :

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

La liste des complexes protéine-ligand étudiés sera retrouvée dans l'annexe 3.

Dans la figure suivante, les résultats sont donnés en pourcent (%), à deux intervalles de RMSD.

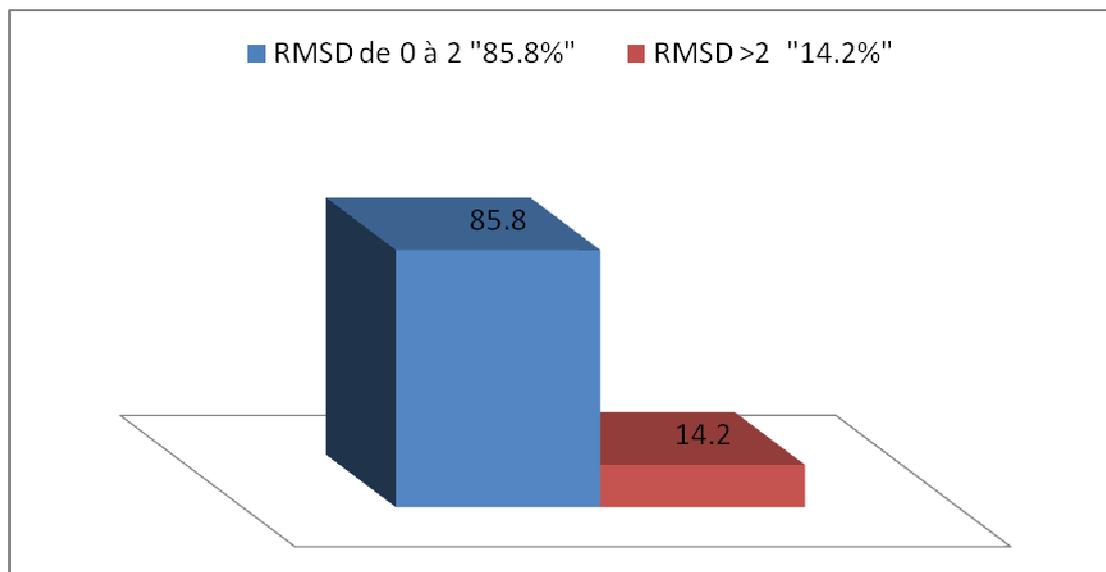


Figure 16. Résultats en % obtenus par Surfex à deux intervalles de RMSD (Å).

La figure montre que 85,8 % des valeurs RMSD sont inférieures à 2 Å et que seulement 14.2 % des résultats sont supérieurs à 2 Å. Conformément aux travaux de vieth M. et *al* (1997) et Chikhi A., Bensegueni A. (2008) [78,79], ces résultats montrent que le programme Surfex est un outil de modélisation par docking moléculaire hautement performant.

Afin de nous assurer que le programme Surfex peut être utilisé sans trop de risque d'erreurs dans notre travail sur la topoisomérase I, nous avons appliqué le test RMSD sur six complexes de la topoI humaine disponibles dans la PDB. Les résultats sont représentés dans le tableau 2.

Tableau 2: Valeurs du RMSD des six complexes protéine-ligand

Code PDB du complexe	Code PDB du ligand	RMSD (Å)
1T8I	EHD	0.411

1K4T	TTC	0.430
1RR8	TTG	0.481
1SC7	M38	0.850
1TL8	AI3	1.045
1SEU	SA3	1.679

Le tableau montre que dans tous les cas, les valeurs de RMSD sont inférieures à 2 Å. Ces résultats sont fortement appuyés par l'analyse visuelle qui montre à chaque fois une superposition maximale de la conformation optimale du ligand calculée par Surflex avec la géométrie du même ligand donnée par diffraction des rayons-X (Fig. 17, 18, 19, 20, 21 et 22). Dans toutes les figures les hydrogènes ne sont pas représentés pour une meilleure clarté de l'image.

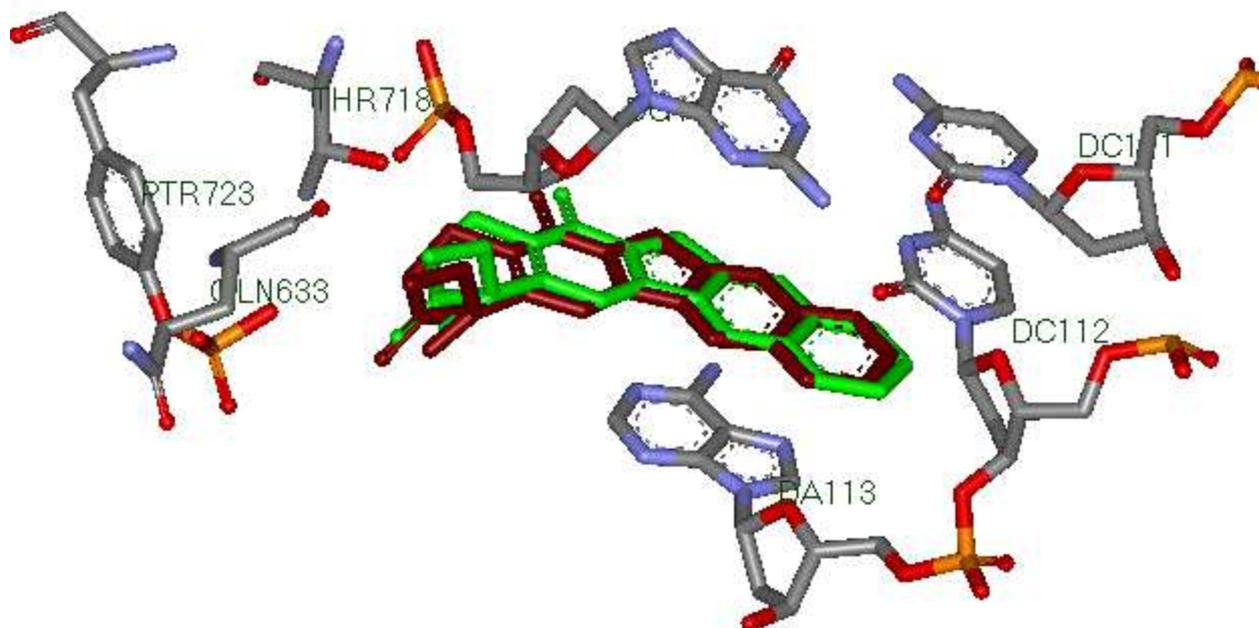


Figure 17. Superposition de la géométrie du ligand EHD obtenu par rayon-X (coloré en marron) et celle calculée par docking moléculaire avec Surflex (coloré en vert).

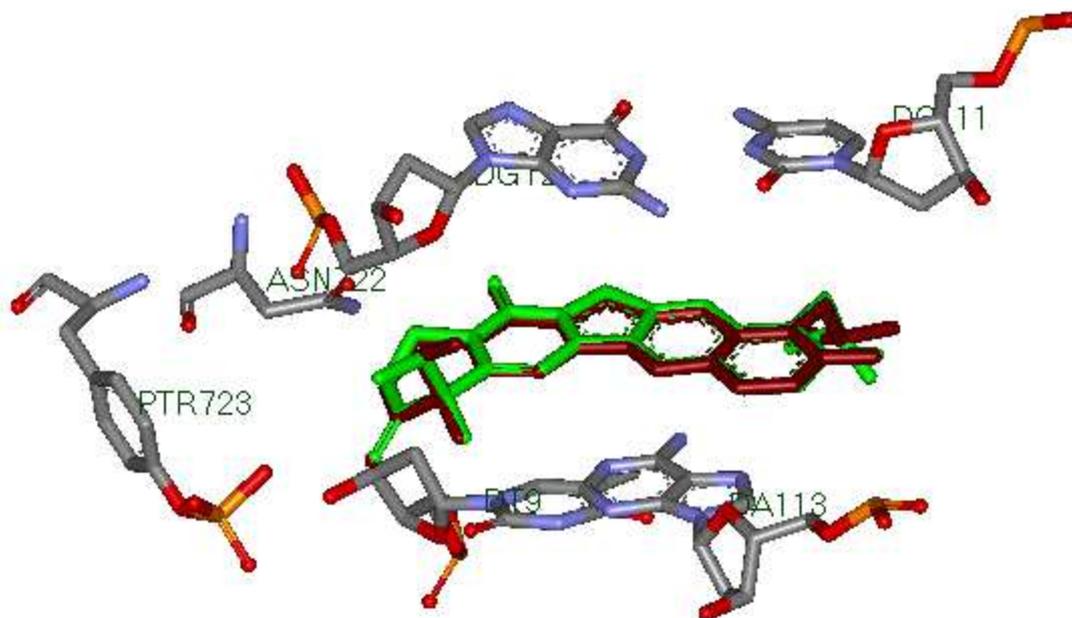


Figure 18. Comparaison de la géométrie de la TTC donnée par rayon-X (coloré en marron) avec celle obtenue par docking moléculaire (coloré en vert).

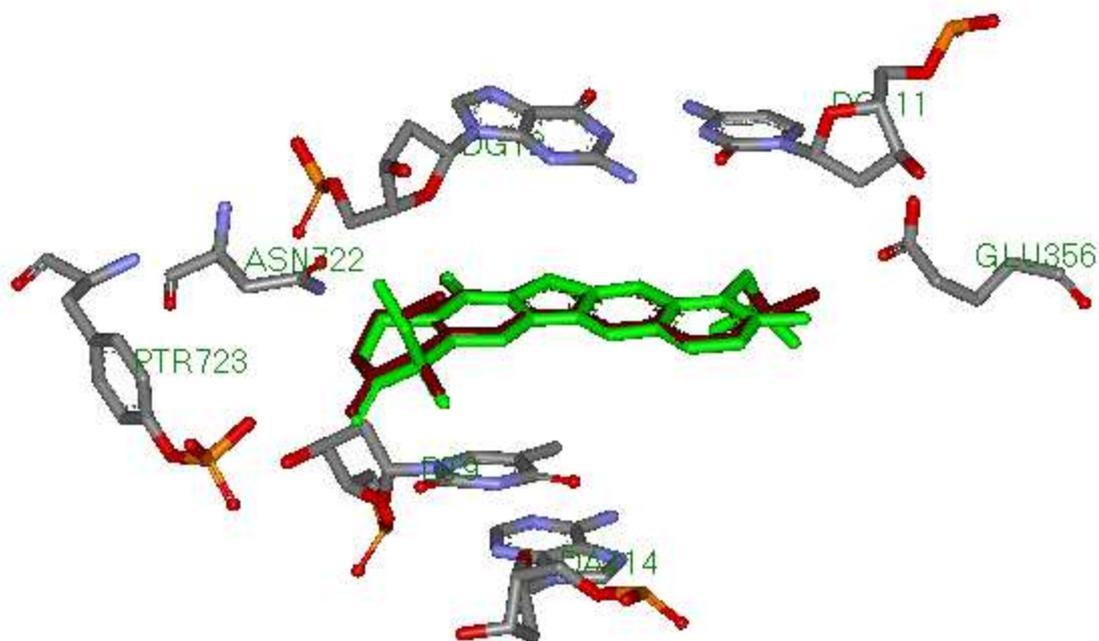


Figure 19. Superposition de la conformation du ligand TTG donnée par rayon-X (coloré en marron) et celle simulée par Surfex (coloré en vert).

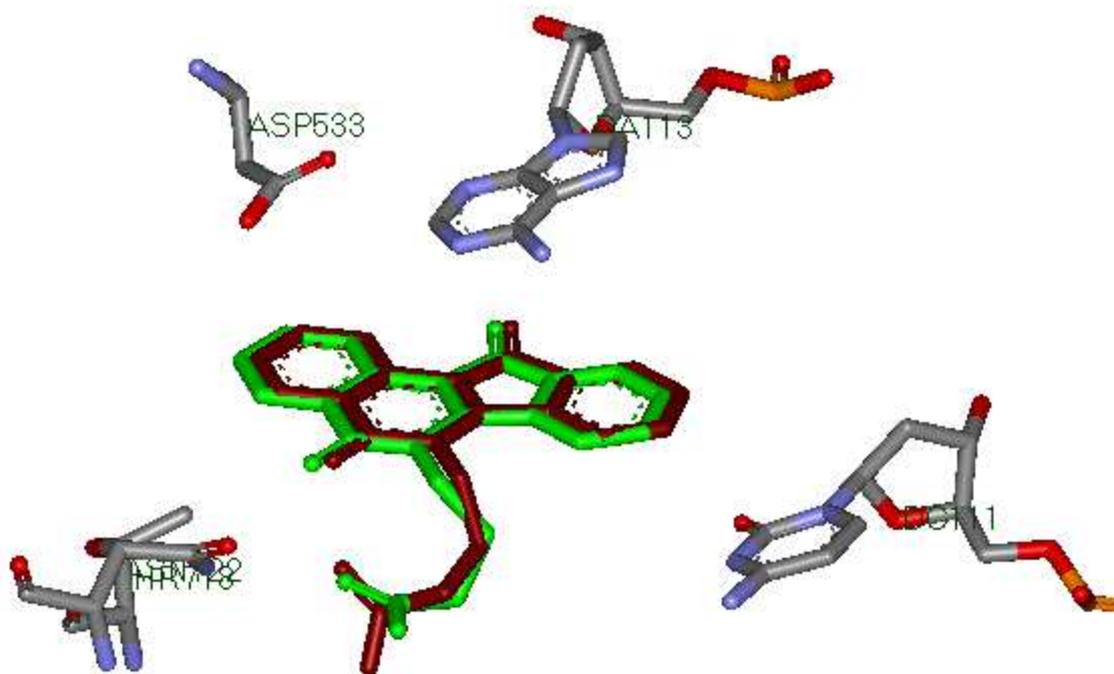


Figure 20. Comparaison des géométries de l'inhibiteur M38 données par rayon-X (coloré en marron) et par simulation avec Surfex (coloré en vert).

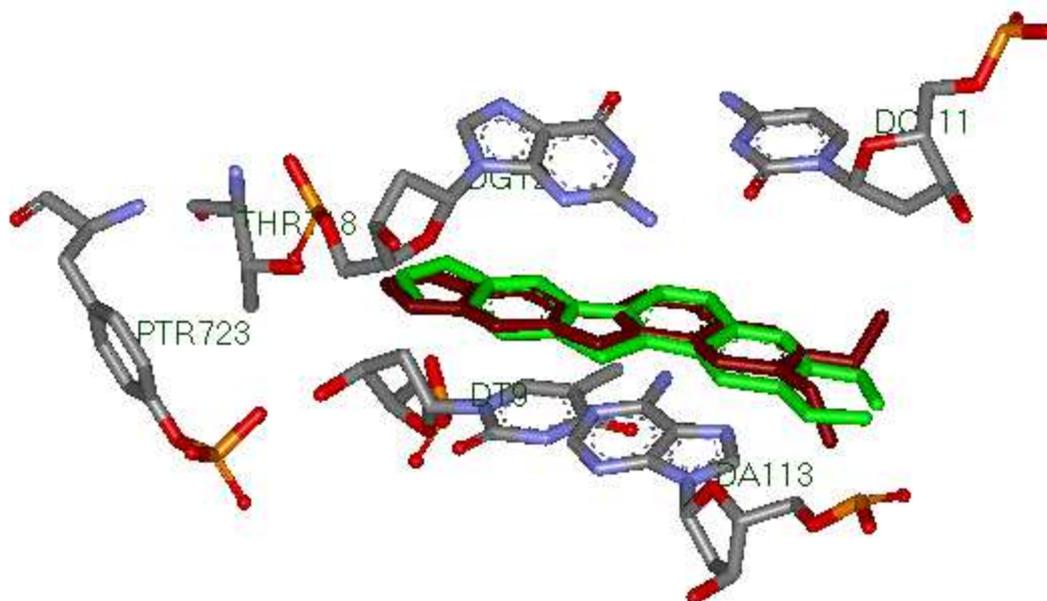


Figure 21. Superposition des géométries du ligand AI3 données par rayon-X (coloré en marron) et par modélisation avec Surflex (coloré en vert).

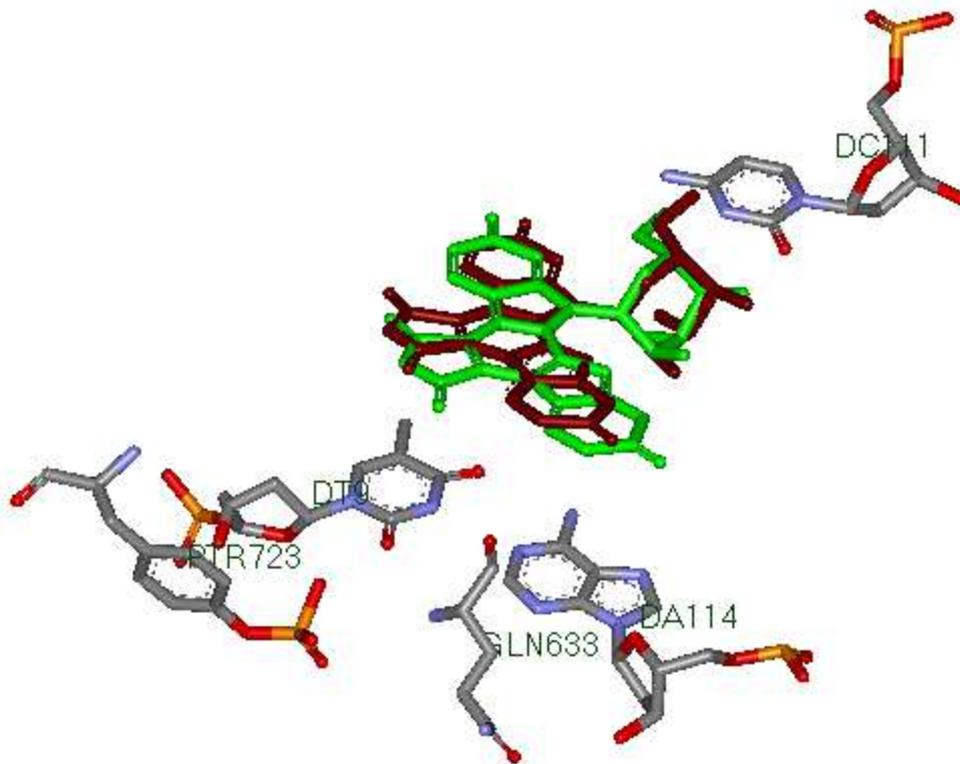


Figure 22. Comparaison des géométries de l'inhibiteur SA3 données par rayon-X (coloré en marron) et par simulation avec Surfex (coloré en vert).

2. Etude des interactions entre les flavonoïdes et la topoisomérase I

L'étude fut menée sur quatre composés naturels appartenant à la classe des flavonoïdes dont trois sont des flavonols : la myricétine, la quercétine, la fisétine et une flavone : l'apigénine.

2-1 Règle de Lipinski

La règle de 5 de Lipinski [80] est en relation directe avec les propriétés biologiques d'absorption, de digestion, de métabolisme et d'excrétion (ADME) éventuelles des molécules étudiées, repose sur cinq propriétés physico-chimiques à savoir :

- Un poids moléculaire (PM) maximum de 500g/mol
- Un maximum de 5 liaisons H-donneurs (DH) (OH ou NH)
 - Un maximum de 10 liaisons H-accepteurs (AH) (O ou N)
- nombre de liaisons retables (LF) ≤ 15
- Un coefficient de partition octanol / eau "logP" ou une lipophilie ≤ 5 .

Le tableau suivant présente les résultats de la règle de Lipinski appliquée aux flavonoïdes étudiés.

Tableau 3 : Résultats de la règle de Lipinski sur les quatre flavonoïdes

Flavonoïdes \ Descripteur	PM	DH	AH	LF	LogP
Myricétine	318.2	6	8	1	2.63
Fiséatine	286.2	4	6	1	3.23
Quercétine	302.2	5	7	1	2.82
Apégénine	270.2	3	5	1	2.77

Le tableau montre que parmi les cinq conditions de la règle de Lipinski, une seule n'est pas respectée pour la myricétine. Il s'agit du donneur d'hydrogène dont le nombre est égal à 6 alors que le maximum est fixé à 5, cependant les quatre flavonoïdes sont capables de présenter une activité biologique très importante conformément aux résultats de Lipinski., et *al* (1997) [80] qui montre que toute molécule validant au moins trois conditions de la règle de Lipinski est considérée comme candidat potentiel.

2-2. L'analyse visuelle des interactions

Pour effectuer le docking de ces molécules, nous avons utilisé le complexe 1T8I. L'analyse visuelle de ce complexe montre que l'enzyme topoisomérase I est liée à une séquence d'ADN de 22 bases, par une liaison phosphotyrosyl qu'on retrouve entre le résidu Tyr 723 du site actif de l'enzyme et la thymine T10 du brin d'ADN. Dans ce complexe, le ligand est la camptothécine portant le code "EHD". Il s'agit d'un inhibiteur de l'ADN topoisomérase I humaine se trouvant en position intercalée entre la thymine T10 et la guanine G11, empêchant ainsi la religation de l'ADN par cette enzyme. (Voir figure 23)

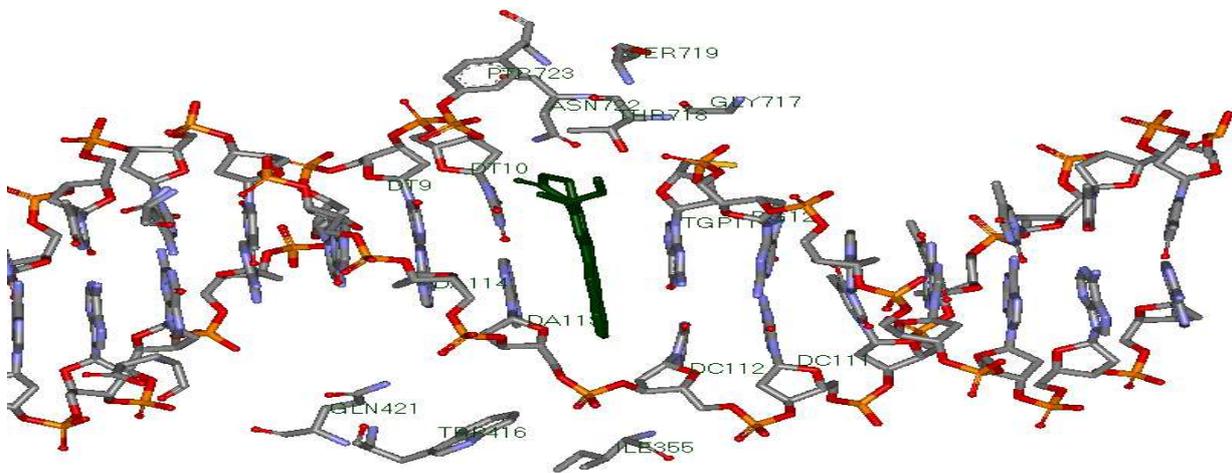
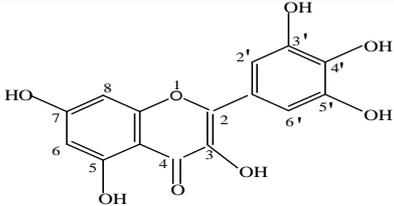
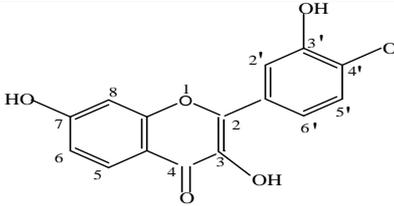
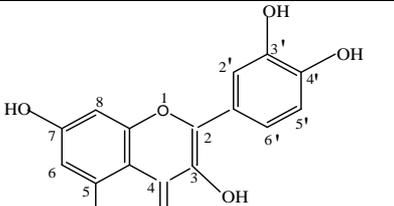
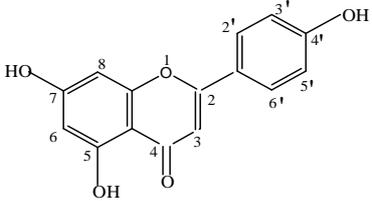


Figure 23 : Positionnement du ligand EHD (coloré en vert) entre les bases T10 et G11 de l'ADN (Pour la clarté de l'image nous avons représenté seulement quelques résidus de la topo I)

Afin de réaliser correctement notre docking, nous avons éliminé au préalable le ligand EHD du site de l'enzyme. Les résultats du docking des flavonoïdes, dans le site actif de l'ADN topoisomérase I sont représentés dans le tableau n°4. Dans le but d'établir une relation entre les résultats du docking moléculaire et les données expérimentales, nous avons reporté dans un même tableau les valeurs théoriques des affinités et les IC_{50} obtenus *in vitro* à partir d'études très récentes [21].

Tableau 4. Résultats théoriques de l'affinité obtenue par docking des flavonoïdes dans le site actif du topoI et leurs relations avec les données expérimentales (IC₅₀).

Composés	Structures	IC ₅₀ (μM)	Affinité (M ⁻¹)
Myricétine		11.1 ± 2.0	7.01
Fisétine		17.4 ± 5.2	6.09
Quercétine		43.9 ± 7.0	5.53
Apigénine		45 ± 10.8	4.16

Les résultats de l'affinité obtenue par surflex seront trouvés dans l'annexe 4.

Le tableau montre que le score d'affinité des 3 flavonols myricétine, fisétine et quercétine est meilleur que celui de la flavone apigénine, dû vraisemblablement au nombre plus élevé d'hydroxyles qu'on retrouve sur leur squelette de base en position 3, 3' et 5' notamment.

Il ressort également de ce tableau que parmi ces trois flavonols, la myricétine présente le meilleur score d'affinité (7.01 M^{-1}), suivie de la fisétine (6.09 M^{-1}) puis de la quercétine (5.53 M^{-1}). Enfin, le tableau révèle que les scores d'affinité des quatre flavonoïdes sont en parfait accord avec les résultats expérimentaux des IC_{50} obtenus dans des tests *in vitro* par Lopez-Lazaro M. et *al* (2010) [21]. D'autres travaux ont montrés que ces trois flavonols présentent une activité inhibitrice vis-à-vis de la topoisomérase I [81,82,83]. Certains chercheurs ont même prouvé que la flavone apigénine peut inhiber la topo I [84].

Nos résultats sont complétés par une analyse visuelle des interactions de chaque molécule avec l'ADN topoisomérase I.

a/ Myricétine

Il est utile de rappeler que parmi les quatre flavonoïdes étudiés, la myricétine avec les meilleurs résultats sur l'affinité d'interaction (7.01 M^{-1}) et sur l' IC_{50} ($11.1 \pm 2.0 \mu\text{M}$) se présente comme l'inhibiteur le plus puissant de la topo I.

Ces résultats s'expriment visuellement par la mise en place de nombreuses interactions de type hydrophobe et hydrophile entre l'enzyme, l'ADN et l'inhibiteur. Comme dans le cas de la camptothécine, la myricétine occupe une position intercalante entre les bases T10 et G11.

On observe la présence de cinq ponts hydrogène formés par les hydroxyles des carbones 3, 5, 7, 3' et 5' de la molécule. Le premier pont hydrogène est observé entre l'hydroxyle porté par le C3 et l'oxygène cyclique du désoxyribose lié à la base G11 (Myr-C3-OH...O-G11; $d = 2.04 \text{ \AA}$).

L'hydroxyle en C5 établit un deuxième pont hydrogène avec l'oxygène de la base T10 (Myr-C5-OH...O4-T10; $d = 2.18 \text{ \AA}$). Un troisième pont hydrogène est formé entre le groupement C7-OH et l'azote N₄ de la base cytosine C112 (Myr-C7-OH...N₄-C112; $d = 2.47 \text{ \AA}$). On retrouve un quatrième pont hydrogène entre l'hydroxyle en 3' et l'oxygène du résidu Thr718 (Myr-C3'-OH...O-Thr718; $d = 1.92 \text{ \AA}$). Le cinquième pont est formé entre le groupement C5'-OH et l'oxygène du sucre lié à la base G12 (Myr-C5'-OH...O-GC12; $d = 1.81 \text{ \AA}$). Enfin la myricétine est stabilisée par des interactions hydrophobiques formées par la base adénine (A113) et les résidus : Tyr 723, Gly 717 et Leu 721. (Figure 24)

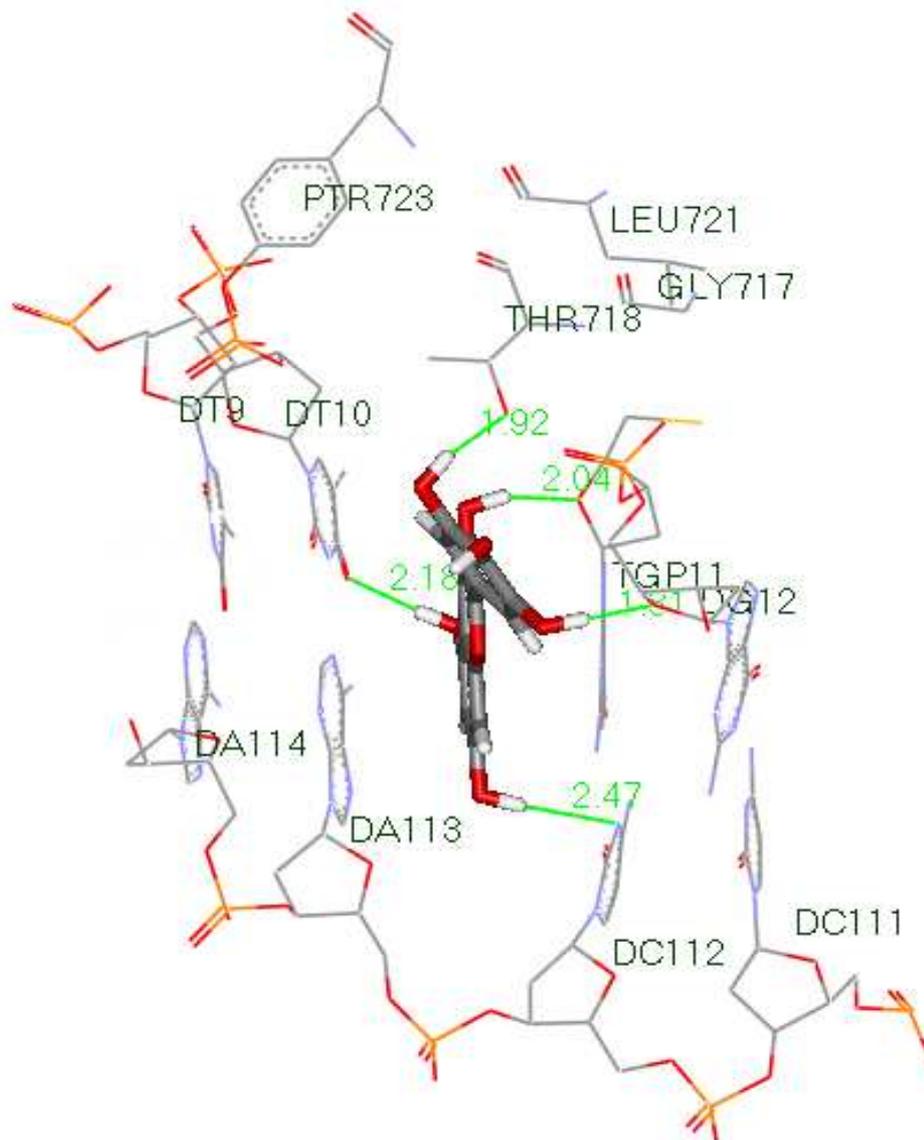


Figure 24 : Mode de liaison de la myricétine avec le site actif de l'ADN topoisomérase I

b/ Fisétine

Ce composé forme quatre ponts hydrogène dont le premier se trouve entre l'hydroxyle en position C3 et l'oxygène O4 de la base T10 (Fis-C3-OH...O4-T10; $d = 2.20 \text{ \AA}$), le deuxième entre le groupement C7-OH et l'oxygène du sucre lié à la base A113 (Fis-C7-OH...O-A113; $d = 2.10 \text{ \AA}$), le troisième entre l'hydroxyle en position C3' et l'oxygène O2 de la base T10 (Fis-C3'-OH...O2-T10; $d = 2.24 \text{ \AA}$) et enfin le quatrième entre l'hydroxyle porté par C4' et l'oxygène du résidu ASN722 (Fis-C4'-H...O-ASN722; $d = 1.84 \text{ \AA}$). La molécule est stabilisée également par les bases G11, C112 et les résidus : Leu 721, Tyr723 et Gly717 par des interactions hydrophobiques. (Figure25)

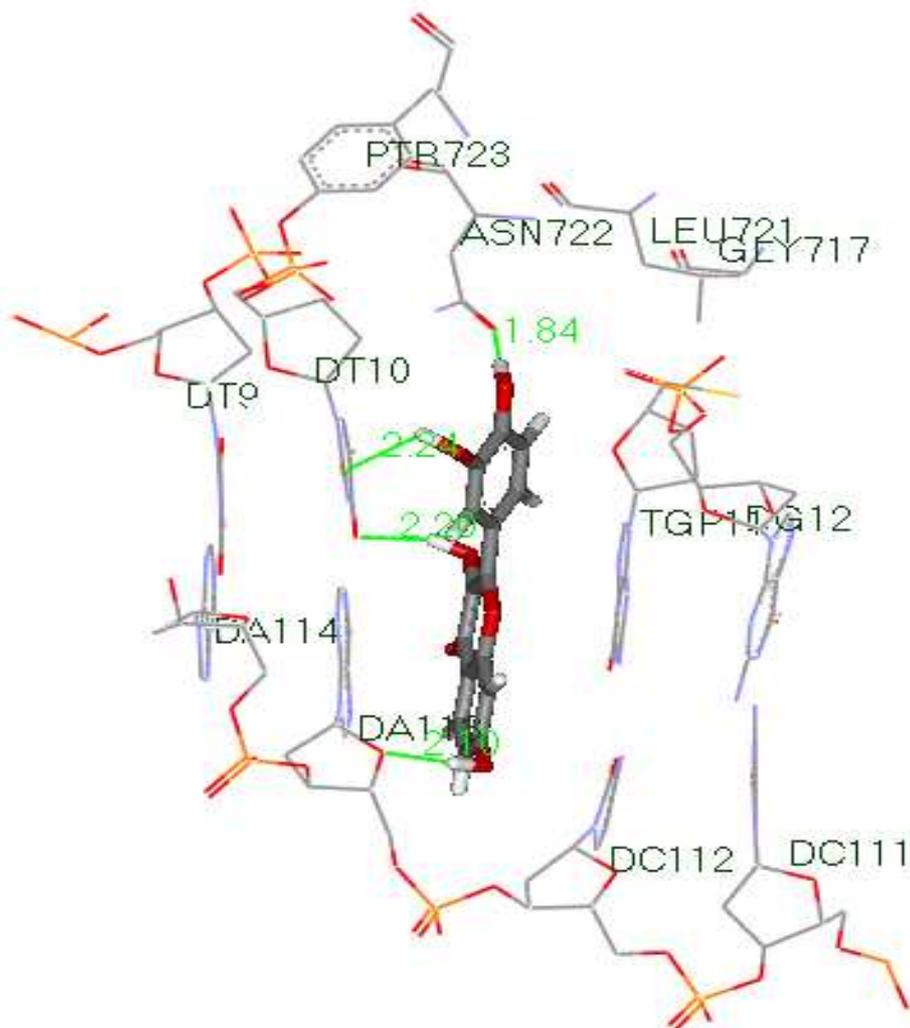


Figure 25 : Mode de liaison de la fisétine avec le site actif de l'ADN topoisomérase I

c/ Quercétine

L'analyse visuelle du complexe nous révèle trois liaisons hydrogènes, dont l'une entre le groupement C3-OH de la quercétine et l'oxygène du sucre lié à la base G11 (Que-C3-OH...O-G11; $d = 2.13 \text{ \AA}$), une autre liaison entre le groupement C3'-OH et l'oxygène du sucre lié au G12 (Que-C3'-OH...O-G12; $d = 2.03 \text{ \AA}$), et enfin une troisième liaison entre l'hydroxyle en position C4' et OP₁ lié au G12 (O-H...O - G12; $d = 1.99$). La molécule est stabilisée par les bases A113, C112, T10 et les résidus: Gly717, Tyr723 et Leu 721 par des interactions hydrophobiques. (figure26)

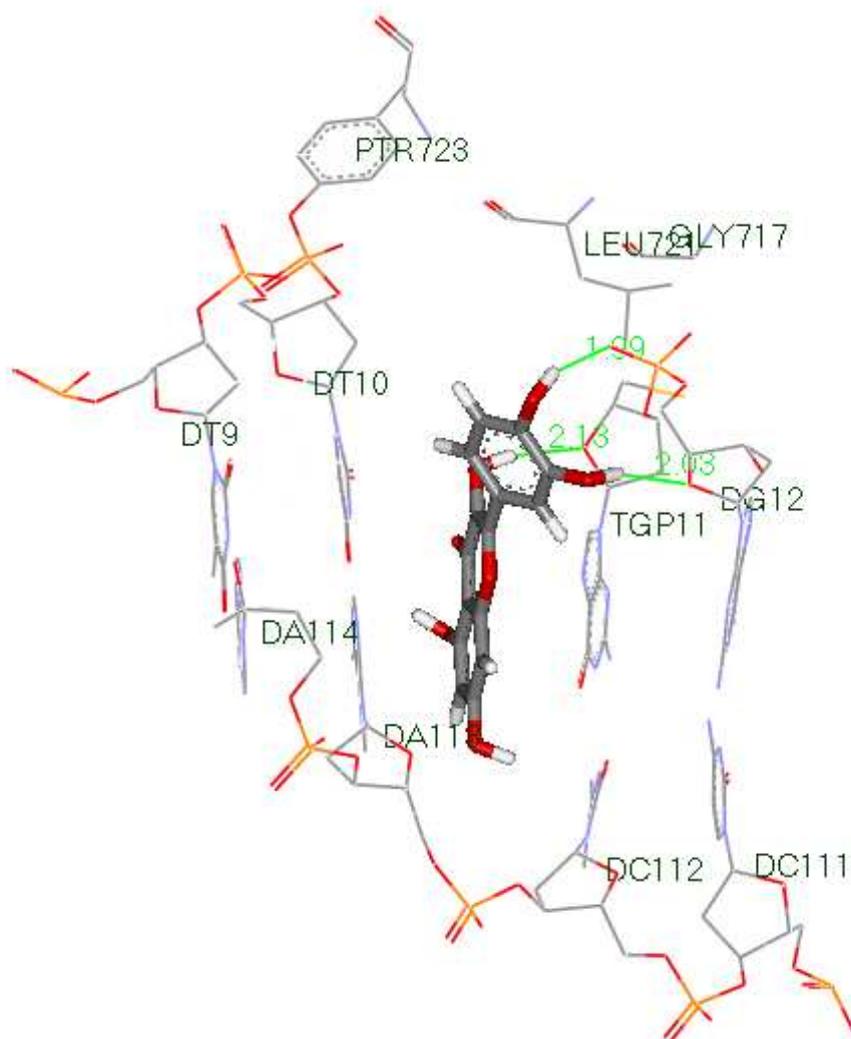


Figure 26 : Mode de liaison de la quercétine avec le site actif de l'ADN topoisomérase I

d/Apigénine

La visualisation des interactions entre l'apigénine et L'ADN topoisomérase I montre trois ponts hydrogène, dont le premier est formé entre l'hydroxyle en position C5 de la molécule et l'oxygène O6 de la base G11 (Api-C5-OH...O6-G11, $d = 2.18 \text{ \AA}$), le deuxième pont hydrogène est établi entre le groupement C7-OH et l'oxygène O4 de la base T10 (Api-C7-OH...O4-T10, $d = 1.98 \text{ \AA}$), et le troisième pont est formé entre le groupement C4'-OH et OP₁ lié au G12 (Api-C7--H...O-G12, $d = 1.89 \text{ \AA}$). Enfin la molécule est stabilisée par les noyaux des bases C112, A113 et les résidus: Tyr723, Gly717 et Leu 721 par des interactions hydrophobiques. (Figure 27)

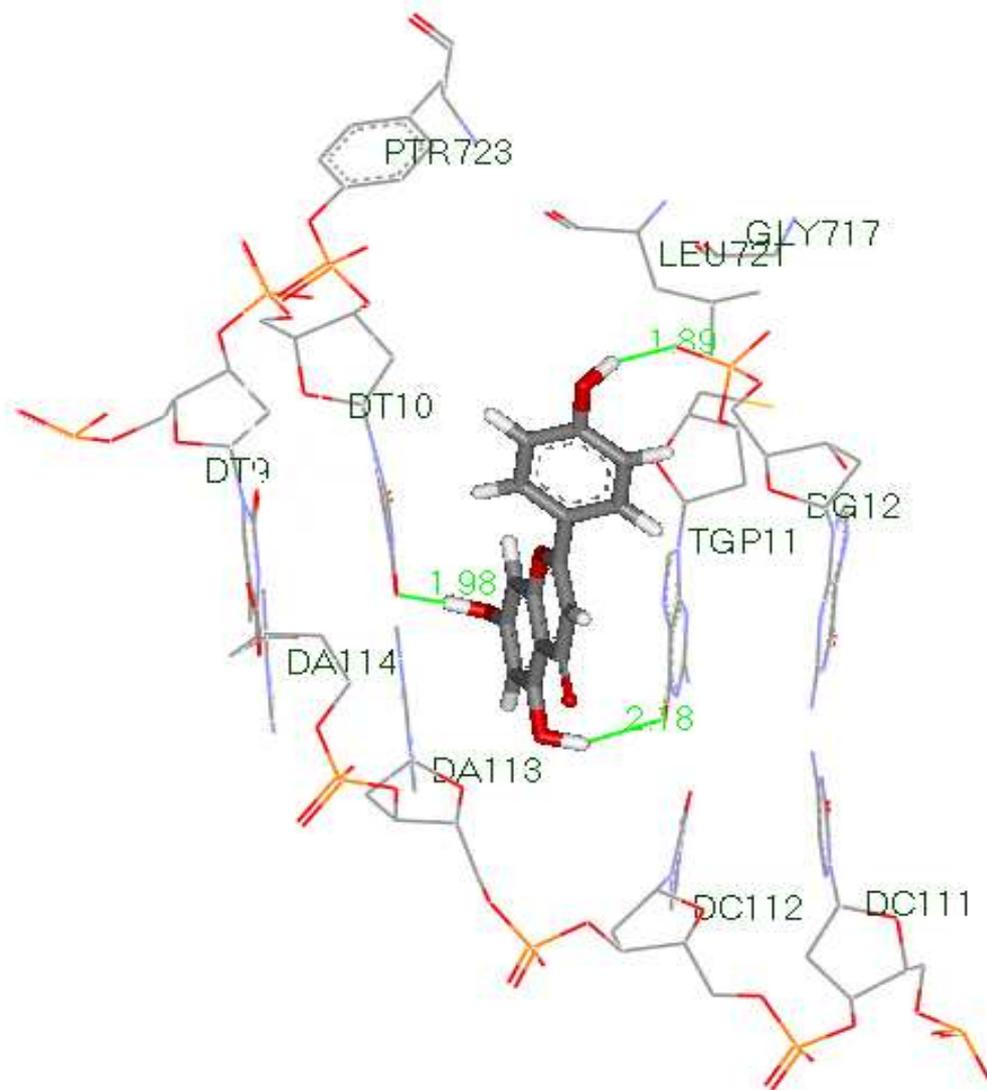


Figure 27 : Mode de liaison de l'apigénine avec le site actif de l'ADN topoisomérase I

Les résultats du docking montrent que les trois flavonols : myricétine, fisétine et quercétine avec des affinités de 7.01 M^{-1} , 6.09 M^{-1} et 5.53 M^{-1} présentent une meilleure activité inhibitrice que la flavone apigénine dont l'affinité est égale à 4.16 M^{-1} . Sachant que la structure de l'apigénine diffère des flavonols par l'absence du groupement OH en position C3, montrant ainsi le rôle majeur que joue cet hydroxyle dans l'activité inhibitrice des flavonoïdes, qui est en accord avec les travaux expérimentaux de Constantino A. et *al* (1995) [81].

En plus de l'hydroxyle en C3, le groupement OH en C3' présent chez les trois flavonols cités plus haut forme des liaisons hydrogènes avec les résidus du site actif. Ceci démontre l'importance de ce groupement qui intervient pour une meilleure activité inhibitrice vis-à-vis de la topoI. Le rôle de ce groupement est également observé par la faible affinité que présente l'apigénine en comparaison avec les trois flavonols étudiés, cette molécule ne contient pas d'hydroxyle en C3', ces résultats sont en parfait accord avec les résultats expérimentaux de

Michael R. et *al* (2004) [85].

En comparant la structure des trois flavonols, on peut constater que seule la myricétine possède sur son noyau latérale un OH en C5'. Ce dernier joue un rôle majeur dans l'activité inhibitrice de cette molécule qui jusque là, présente la meilleure affinité avec la topoI, conformément aux résultats de Lopez-Lazaro M. et *al* (2010) [21].

3. Substitution de la myricétine :

Dans le but de développer de nouveaux inhibiteurs plus efficaces de la Topo I, nous avons utilisé comme structure de départ, le meilleur inhibiteur parmi les flavonoïdes étudiés : la myricétine. (Voir fig28)

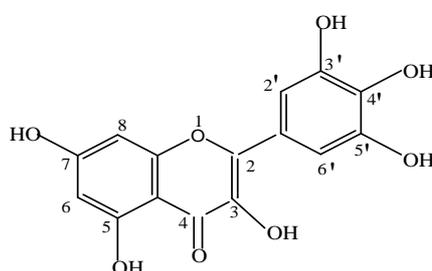


Figure 28 : Structure de la myricétine

Afin d'améliorer l'affinité de la myricétine, plusieurs groupements fonctionnels présentant la capacité d'engager des liaisons hydrogène avec le site actif de l'enzyme ont été introduits sur différentes positions de la molécule. Il s'agit des fonctions : alcool (-OH); thiol (-SH); acide (-COOH) et amine (-NH₂) [86]. (Voir tableau 5)

Tableau 5: Résultats de la substitution

Les groupements Position De la substitution	COOH	SH	NH ₂	CH ₂ -OH
	Affinité (M ⁻¹)			
C3	4.91 1	4.11 2	6.81 3	6.02 4
C5	8.11 5	6.01 6	6.88 7	5.35 8
C6	6.54 9	5.93 10	5.40 11	6.88 12
C7	7.80 13	5.79 14	7.55 15	7.62 16
C8	6.12 17	5.69 18	6.44 19	6.69 20
C2'	4.63 21	4.94 22	5.35 23	5.38 24
C3'	8.11 25	7.84 26	8.22 27	8.79 28
C4'	8.36 29	6.36 30	8.12 31	6.98 32

C5'	6.82	7.55	7.22	7.56
	33	34	35	36
C6'	5.47	4.11	5.64	4.08
	37	38	39	40

Le tableau montre que parmi les différentes substitutions réalisées, seuls 13 composés ont présenté des affinités supérieures à celle de la myricétine qui est de 7.01 M^{-1} . Il s'agit des composés portant les numéros : 5,13,15,16,25,26,27,28,29,31,34,35 et 36. Par ailleurs les résultats indiquent que les substitutions en C3', C4', C5', C5 et C7 ont nettement amélioré l'affinité de la myricétine montrant le rôle essentiel de ces carbones dans l'inhibition de la topo I, contrairement aux positions C2', C6', C6 et C8 où une baisse de l'affinité est observée.

Enfin, il faut souligner l'importance de l'hydroxyle en C3 car l'utilisation d'autres groupements fonctionnels dans cette position n'a pas pu améliorer l'affinité de la molécule étudiée.

4. Application de la règle de Lipinski sur les 13 composés

Le tableau ci-dessous présente les résultats de la règle de 5 de Lipinski sur les 13 composés présentant une affinité supérieure à celle de la myricétine.

Tableau 6: résultats de la règle de Lipinski sur les 13 composés

Composés	Affinité (M ⁻¹)	Descripteur				
		PM	DH	AH	LF	LogP
5	8.11	346.2	6	9	2	2.21
13	7.80	346.2	6	9	2	2.21
15	7.55	317.2	6	8	1	2.22
16	7.62	32.2	6	8	2	2.22
25	8.11	346.2	6	9	2	2.42
26	7.84	334.2	5	7	1	3.2
27	8.22	317.2	6	8	1	2.22
28	8.79	332.2	6	8	2	2.22
29	8.36	346.2	6	9	2	2.42
31	8.12	317.2	6	8	1	2.22
34	7.55	334.2	5	7	1	3.2
35	7.22	317.2	6	8	1	2.22
36	7.56	332.6	6	8	2	2.22

Les résultats du tableau montrent que les 13 composés s'inscrivent parfaitement dans la marge des critères imposés par la règle de Lipinski. Parmi ces composés, il ressort que le composé 28 présente la meilleure affinité (8.79 M⁻¹).

Une étude plus poussée sur le composé 28 a été réalisée en faisant appel à d'autres propriétés physico-chimiques complémentaires de la règle de Lipinski [75]. Il s'agit de :

- nombre d'halogènes (F, Cl, Br, I) ≤ 7
- chaînes alkyles ≤ -(CH₂)₆-CH₃
- nombre de cycles ≤ 6
- au moins un atome d'azote ou d'oxygène dans la molécule.

L'étude a donné les résultats suivants (tableau 7).

Tableau 7 : Autre propriétés physico-chimiques du composé 28

Descripteur	Nombres d'halogènes	Longueur de la chaîne alkyle	Nombre de cycle	Nombre d'azote ou d'oxygène
Composé 28	0	Dépourvue de Chaîne alkyle	3	8 atomes d'oxygènes

Ce complément d'information sur d'autres propriétés moléculaires du composé 28 nous confirme qu'il s'agit d'une molécule biologiquement très active dont l'effet inhibiteur sur la topo I doit être vérifié expérimentalement par des tests *in vitro* et *in vivo*.

5. Etude des interactions du composé 28 :

L'étude du composé 28 a été complétée par une visualisation des interactions établit entre ce composé et l'ADN topoisomérase I, cette molécule est obtenu par substitution de l'hydroxyle en C3' par le groupement CH₂OH. Cette substitution a entraîné la formation d'un pont hydrogène entre l'hydroxyle de ce groupement et le résidu Asn722 de la topo I avec une distance de 1.89 Å, (-C3'-CH₂-OH...O-Asn722). On voit également apparaitre un pont hydrogène entre le groupement C3-OH du composé 28 et l'azote N₇ de la base G11 (C3-OH...N₇-G11; d = 2.12 Å). Le groupement C7-OH établit une liaison hydrogène avec l'oxygène du sucre lié à l'adénine 113 (C7-OH...O-A113; d = 2.06 Å). Enfin le groupement C5'-OH forme une liaison hydrogène avec l'oxygène de la base T10 (C5'-OH...O-T10; d = 1.96 Å). La molécule est stabilisée par d'autres interactions de type hydrophobique faisant intervenir les résidus Leu 721, Tyr723, C112 et Gly717. (figure29)

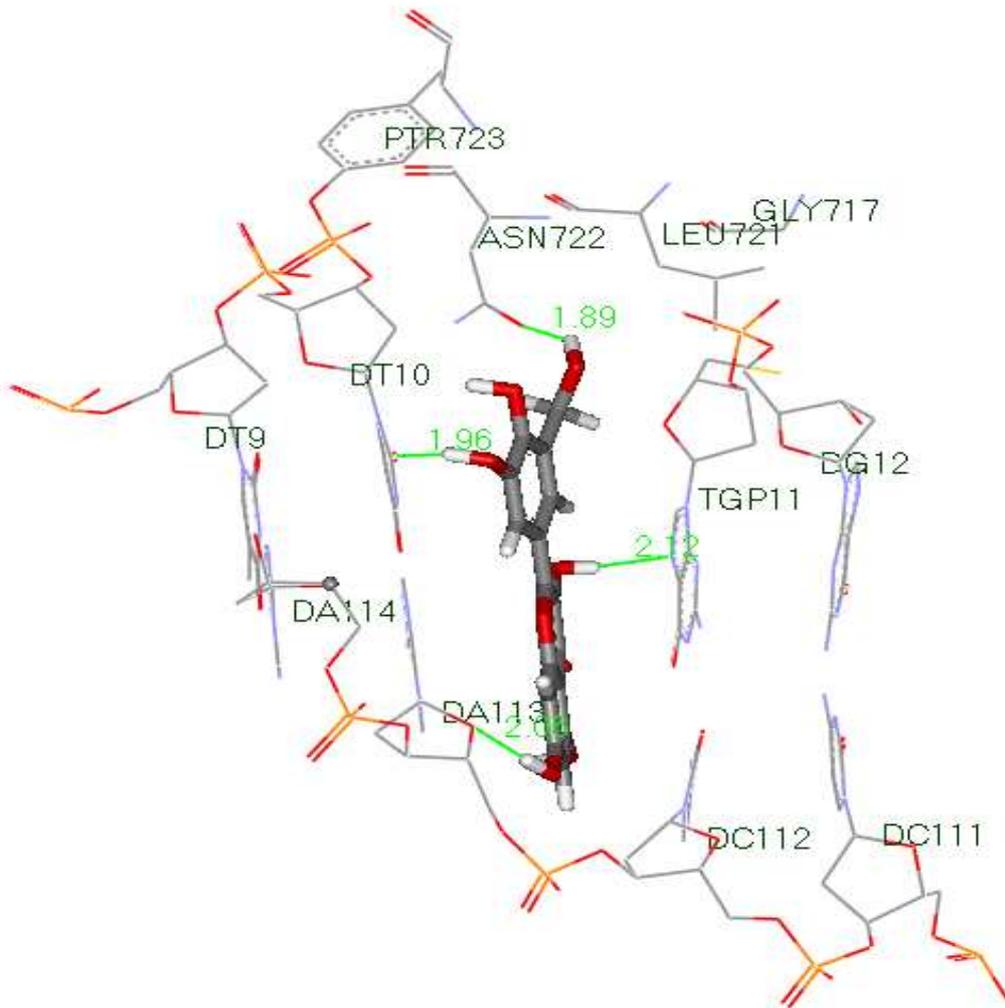


Figure 29 : Mode de liaison du composé 28 avec le site actif de l'ADN topoisomérase I

Conclusion et perspectives

Conclusion et perspectives

Ce mémoire de thèse avait pour but le développement *in silico* de nouvelles structures hypothétiques agissant comme inhibiteurs très puissants de la topoisomérase I dans le domaine de la chimiothérapie cancéreuse. Pour cela nous nous sommes appuyés sur des récents travaux expérimentaux montrant que les quatre flavonoïdes myricétine, fisétine, quercétine et apigénine sont des inhibiteurs de la topoisomérase I. Dans cette optique, à partir de modifications structurales variées, nous avons cherché à développer une nouvelle molécule anticancéreuse en conservant la structure de base des flavonoïdes. Afin d'atteindre cet objectif nous sommes passés par plusieurs étapes.

Dans un premier temps nous avons testé la fiabilité du programme Surflex utilisé dans ce travail. Pour cela nous avons fait appel au test RMSD (root-mean-square derivation), permettant de comparer la déviation entre la structure simulée par Surflex et celle obtenue par cristallographie. Sur les 134 complexes pris arbitrairement de la PDB, 85.8 % ont présenté un RMSD inférieur à 2 Å, témoignant ainsi de la haute performance du programme Surflex. L'application de ce test à six complexes de l'ADN Topoisomérase I humaine disponibles dans la PDB a donné à chaque fois un RMSD inférieur à 2 Å, l'analyse visuelle de ces résultats montre une superposition maximale de la conformation optimale du ligand calculée par Surflex et la géométrie du même ligand pris initialement de la PDB. Ce test nous a permis de conclure que Surflex est un programme de docking moléculaire très fiable qu'on peut utiliser en toute confiance pour modéliser les interactions des quatre flavonoïdes cités plus haut avec le site actif de la topoisomérase I.

Dans un deuxième temps, la modélisation par docking moléculaire nous a permis d'évaluer l'affinité des flavonoïdes étudiés et de retenir la myricétine avec une valeur de 7.01 M^{-1} comme le meilleur inhibiteur de la topo I.

Ce résultat est en parfait accord avec les travaux de Lopez-Lazaro M., et *al.* (2010), dont les tests *in vitro* sur l'effet inhibiteur de la myricétine donnent la plus faible valeur de l'IC₅₀ ($11.1 \pm 2.0 \mu\text{M}$).

Dans un troisième temps, nous avons pu élucider les mécanismes d'interaction entre les flavonoïdes et la topo I par visualisation des différents types de liaisons faibles mis en jeu. Les résultats montrent que les OH en C3, C5, C7, C3' et C5' du meilleur inhibiteur, qui est la

myricétine forment des liaisons hydrogène avec l'ADN et le site actif de la topo I jouant ainsi un rôle essentiel dans l'inhibition de cette dernière.

Dans un quatrième temps nous avons développé des nouveaux inhibiteurs du topo I présentant une affinité plus élevée que la myricétine. Pour cela nous nous sommes servis de la myricétine comme structure de départ sur laquelle nous avons réalisé différentes substitutions. Parmi les 40 composés obtenus seulement 13 ont présenté une affinité nettement supérieure à celle de la myricétine. La meilleure affinité (8.79 M^{-1}), est donnée par le composé 28 qui provient de la substitution du groupement hydroxyle en position C3' de la myricétine par le groupement alcoolique CH_2OH .

La dernière partie de ce travail nous a permis nous renseigner sur les éventuelles propriétés biologiques d'absorption, de digestion, de métabolisme et d'excrétion des 13 composés par simple application de la règle de 5 de Lipinski. Les résultats montrent que ces composés se présentent comme des inhibiteurs puissants de la topoisomérase I humaine.

Sur la base de ces résultats préliminaires, nous envisageons d'approfondir l'étude de l'effet des flavonoïdes sur l'inhibition de la topoisomérase I, en faisant appel à d'autres programmes d'étude de docking et de simulation *in silico* des propriétés ADME notamment. Notre étude ici n'étant qu'initiale, une étude expérimentale complémentaire *in vitro* et/ou *in vivo* devra permettre de vérifier les résultats théoriques obtenus *in silico*.

Bibliographie

Bibliographie

- [1] : Genoux E. Dérivés de flavonoïdes et de vérapamil comme ligands des transporteurs MRP1 et ABCG2 : De la conception à l'activité Anticancéreuse. Thèse de Doctorat: Chimie-Biologie. Grenoble : Université de Grenoble. France. 2011. 134p.
- [2] : Ghedira K. Flavonoids: structure, biological activities, prophylactic function and therapeutic uses. *Phytothérapie*. 2005 ; **3** : 162-169.
- [3] : Hadj Salem. J. Extraction, identification, caractérisation des activités biologiques de flavonoïdes de NITRARIA retusa et synthèse de dérivés acyles de ces molécules par voie enzymatique. Thèse de Doctorat : Procédés Biotechnologiques et Alimentaires. Nancy: Université de Nancy. France. 2009. 251p.
- [4] : Wenying R., Zhenhua Q., Hongwei W., Lei Z., Li Z. Flavonoids : Promising Anticancer Agents. *Medicinal Research Reviews*. 2003 ; **23** : 519-534.
- [5] : Boege F., Straub T., Kehr A., Boesenberg C., Christiansen K., Andersen A., et al. Sélectionnés flavones roman inhiber la liaison d'ADN ou de l'étape religature ADN topoisomérase I. *Journal of Biological Chemistry*. 1996; **271**: 2262-2270.
- [6] : Yamashita Y., Kawada SZ, Nakano H. L'induction de la topoisomérase II clivage mammifères ADN dépendante par les flavonoïdes nonintercalative, génistéine et orobol. *Biochemical Pharmacology*. 1990; **39**: 737-744.
- [7] : Wang J. C. rôles cellulaires des ADN topoisomérases: Une perspective moléculaire *Nature Reviews Molecular Cell Biology*. 2002; **3** : 430-440.
- [8]: Petron J., Dwyer J. Flavonoids: dietary occurrence and biochemical activity. *Nutrition Research*. 1998; **18**: 1995-2018.
- [9] : Manach C., Scalbert A., Morand C., Remy C., Jiménez L. Polyphenols: food sources and bioavailability . *American Journal of Clinical Nutrition*. 2004; **79**: 727-747.
- [10] : Scalbert A., Williamson G. Dietary intake and bioavailability of polyfenols . *Journal of Nutrition*. 2000; **130**: 2073-2085.

- [11]: Aherne S. A., O'Brien N. M.. Dietary flavonols: chemistry, food content, and métabolisme. *Nutrition*. 2002; **18**: 75-81.
- [12] : Hertog M. G. L., Hollman P. C. H., Katan M. B, Kromhout D. Estimation of daily intake of potentially anticarcinogenic flavonoids and their determinants in adults in the netherlands . *Nutrition and Cancer*. 1993; **20**: 21-29.
- [13]: Hertog M. G. L., Kromhout D., Aravis C., Blackburn H., Buzina R. B., Fidanza F., et al. Flavonoids , intake and long-term risk of coronary heart disease and cancer in the seven countries study. *Archives of Internal Medicine*. 1995; **155** : 381-386.
- [14]: Elicoh-Middleton J., Chithan K., Theoharis C., Effect of plant flavonoids on mammalian cells: implications for inflammation, heart diseases and cancer. *Pharmacology and Experimental therapeutics*. 2000; **4**: 673-751.
- [15]: Bensegueni. A. Etude théorique des métabolites secondaires des végétaux et des composés de synthèse sur le plan de l'activité biologique : simulation par docking (arrimage) moléculaire sur la lipoxigénase et la cyclooxygénase. Thèse de Doctorat : Biochimie appliquée. Constantine : Université Mentouri Constantine. Algérie. 2007. 91p.
- [16]: Manach C., Scalbert A., Morand C., Remy C., Jiménez L. Polyphenols: food sources and bioavailability . *The American Journal of Clinical Nutrition*. 2004; **79**: 727-747.
- [17]: Criqui M.H., Ringel, B.L. Does diet or alcohol explain the French paradox? *Lancet*. 1994; **344**:1719-1723.
- [18]: Renaud, S. ; de Lorgeril, M. Wine, alcohol, platelets, and the French paradox for coronary heart disease. *Lancet*. 1992; **339**: 1523-1526.
- [19]: Milane H. La quercitrine et ses dérivés: molécules : caractère prooxydant ou capteurs de radicaux libres; Etudes et applications thérapeutiques. Thèse de doctorat: Chimie. Strasbourg : Université de Louis Pasteur Strasbourg I. France. 2004. 230p.
- [20]: Brusselmans K., Vrolix R., Verhoeven G., Swinnen J. V. Induction of cancer cell apoptosis by flavonoids is associated with their ability to inhibit fatty acid synthase activity. *Journal of biological chemistry*. 2005; **280**: 5636-5645.
- [21] : Lopez-Lazaro M., Willmore E. A., Austin C. The dietary flavonoids myricetin and fisetin act as dual inhibitors of DNA topoisomerases I and II in cells. *Mutation Research*. 2010; **696**:

41-47.

[22]: Hanahan D., Weinberg R. A., The hallmarks of cancer *Cell*. 2000 ; **100** : 57-70.

[23]: Madonna S. Innovation moléculaire à visée thérapeutique : conception, synthèse et évaluation des propriétés anticancéreuses de nouveaux dérivés du (*N*-(5-méthyl)-quinoléin-8-ol)amine *N*-substitués. Thèse de Doctorat: Chimie Organique. Marseille : Université de la Méditerranée AIX-Marseille II. France. 2010. 201p.

[24]: Dufort. S. Vectorisation de biomolécules pour l'imagerie et la thérapie des cancers. Thèse de Doctorat : Biologie cellulaire. Grenoble : Université de Grenoble. France. 2010. 150p.

[25]: Pottier. A. Conception, synthèse et évaluation pharmacologique de pyrrolo [3,4-*b*] Quinoleines condensées, ligands potentiels de l'ADN. Thèse de Doctorat : Pharmacochimie. Lille : Université de Lille. France. 2003. 238p.

[26]: Krzaczkowski.L. *Recherche exploratoire de substances actives en pharmacologie antitumorale à partir des Bryophytes*. Thèse de Doctorat: Pharmacologie. Toulouse : Université de Toulouse III. France. 2008. 199p.

[27] : Chen A. Y. Liu L. F. "DNA topoisomerases: essential enzymes and lethal targets". *Annual Review of Pharmacology and Toxicology*. 1994; **34**: 191-218.

[28] : Wang J. C. "Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine." *Quarterly Reviews of Biophysics*. 1998; **31**: 107-44.

[29] : Wang J. C. "DNA topoisomerases." *The Annual Review of Biochemistry*. 1996; **65**: 635-92.

[30] : Wang J. C. "Cellular roles of DNA topoisomerases: a molecular perspective". *Nature Reviews Molecular Cell Biology*. 2002; **3**: 430-40.

[31] : Wang J. C. "Interaction between DNA and an Escherichia coli protein omega". *The Journal of Molecular Biology*.1971; **55**: 523-33.

[32] : Champoux J. J., Dulbecco R. "An activity from mammalian cells that untwists superhelical DNA--a possible swivel for DNA replication (polyoma-ethidium bromide-mouse-embryo cells-dye binding assay)". *Proceedings of the National Academy of Sciences USA*. 1972 ; **69**: 143-6.

[33] : Gellert M., Mizuuchi K., O'Deaet M. H., Nash H. A. "DNA gyrase: an enzyme that

introduces superhelical turns into DNA." *Proceedings of the National Academy of Sciences USA*. 1976 ; **3**: 3872-6.

[34] : Wang J. C. "DNA topoisomerases." *The Annual Review of Biochemistry*.1996; **65**: 635-92.

[35]: Nitiss J. L. "Investigating the biological functions of DNA topoisomerases in eukaryotic cells." *Biochimica et Biophysica Acta*. 1998; **1400**: 63-81.

[36]: Wang J. C., Caronet P. R., Kim R. A. "The role of DNA topoisomerases in recombination and genome stability: a double-edged sword?". *Cell*. 1990; **62**: 403-6.

[37]: Champoux J. J. "DNA topoisomerases: structure, function, and mechanism". *The Annual Review of Biochemistry*. 2001; **70**: 369-413.

[38] : D'arpa P., Machlin P., Ratrie H., Rothfield N. F., Cleveland D. W., Earnshaw W. C. cDNA cloning of human DNA topoisomerase I : catalytic activity of a 67.7-kDa carboxyl-terminal fragment. *Proceedings of the National Academy of Sciences USA*. 1988; **85** : 2543-2547.

[39] : Stewart L., Ireton G. C., Champoux J. J. The domain organization of human topoisomerase I. *Journal of Biological Chemistry*. 1996; **271**: 7602-7608.

[40] : Madden K. R., Champoux J. J. Overexpression of human topoisomerase I in baby hamster Kidney cells : hypersensitivity of clonal isolates to camptothecin. *Cancer research*. 1992; **52**: 525-532.

[41] : Stewart L., Ireton G. C., Champoux J. J. Reconstitution of human topoisomerase I by fragment complementation. *The Journal of Molecular Biology*. 1997; **269**: 355-372.

[42] : Wang J. C., DNA topoisomerases. *The Annual Review of Biochemistry*. 1996; **65**: 635-92.

[43]: Stewart L., Redinbo M. R., Qiu X. A model for the mechanism of human topoisomerase I. *Science*. 1998; **279**: 1534-41.

[44] : Caserta M., Amadei A., Di Mauro E., Camilloni G. In vitro preferential topoisomerization of bent DNA. *Nucleic Acids Research*. 1989; **17**: 8463-8474.

[45] : Krogh S., Mortensen U. H., Westergaard O., Bonven B. J. Eukaryotic topoisomerase I-DNA interaction is stabilized by helix curvature. *Nucleic Acids Research*. 1991; **19**: 1235-41.

[46]: Champoux J. J. DNA is linked to the rat liver DNA nicking-closing enzyme by a

phosphodiester bond to tyrosine. *Journal of Biological Chemistry*. 1981; **256**: 4805-4809.

[47] : Lynn R. M., Bjornsti M. A., Caron P. R., Wang J. C. Peptide sequencing and site-directed mutagenesis identify tyrosine-727 as the active site tyrosine of *Saccharomyces cerevisiae* DNA topoisomerase I. *Proceedings of the National Academy of Sciences USA*. 1989 ; **86**: 3559-3563

[48] : Christiansen K., Svejstrup A. B., Andersen A. H, Westergaard O. Eukaryotic topoisomerase I-mediated cleavage requires bipartite DNA interaction. Cleavage of DNA substrates containing strand interruptions implicates a role for topoisomerase I in illegitimate recombination. *Journal of Biological Chemistry*. 1993; **268**: 9690-9701.

[49]: Stivers J. T., Jagadeesh G. J., Nawrot B., Stec W. J, Shuman S. Stereochemical outcome and kinetic effects of Rp- and Sp-phosphorothioate substitutions at the cleavage site of vaccinia type I DNA topoisomerase. *Biochemistry*. 2000; **39**: 5561-5572.

[50]: Svejstrup J. Q., Christiansen K., Andersen A. H., Lund K., Westergaard O. Minimal DNA duplex requirements for topoisomerase I-mediated cleavage in vitro. *Journal of Biological Chemistry*. 1990; **265**: 12529-12535.

[51]: Capranico G., Binaschi M. DNA sequence selectivity of topoisomerases and topoisomerase poisons. *Biochimica et Biophysica Acta*. 1998; **1400**: 185- 194.

[52]: Kroeger P. E., Rowe T. C. Analysis of topoisomerase I and II cleavage sites on the *Drosophila* actin and Hsp70 heat shock genes. *Biochemistry*. 1992; **31**: 2492-2501.

[53]: Gilmour D. S., Pflugfelder G., Wang J. C., Lis J. T. Topoisomerase I interacts with transcribed regions in *Drosophila* cells. *Cell*. 1986; **44**: 401-407.

[54]: Liu L. F., Wang J. C. Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences USA*. 1987; **84**: 7024-7027.

[55]: Wang L. K., Johnson Hecht S. M. Inhibition of topoisomerases I function by nitidine and fagaronine. *Chemical Research in Toxicology*. 1993; **6**: 813-818.

[56] : Wargnier R. Dosage de l'ADN topoisomerase I dans les lignes cellulaires issues de tumeurs solides et approche nanotechnologique du dosage simultané de la protéine et de son gène. Thèse Doctorat: Biomolécules et Dynamique Cellulaire. Reims Champagne-Ardenne : Université de Reims Champagne-Ardenne. France. 2004. 146p.

[57] : Bjornsti M. A., Benedetti P., Viglianti G. A., Wang J. C. Expression of human DNA topoisomerase I in yeast cells lacking yeast DNA topoisomerase I : restoration of sensitivity of the cells to the antitumor drug camptothecin. *Cancer research*. 1989; **49**: 6318-6323.

[58]: Gupta M., Fujimori A., Pommier Y. Eukaryotic DNA topoisomerase I. *Biochim Biophys*

Acta 1995; **1262** : 1-14

[59]: Pommier Y, Jenkins J, Kohlhagen G, Leteurtre F. DNA recombinase activity of eukaryotic DNA topoisomerase I; effects of camptothecin and other inhibitors. *Mutat Res* 1995; **337** : 135-145.

[60]: Pommier Y, Pourquier P, Fan Y, Strumberg G D. Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme. *Biochimica et Biophysica Acta*. 1998; **1400**: 83-105.

[61]: Redinbo M. R., Champoux J.J., Hol W.G. Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA. *Biochemistry*. 2000; **39**: 6832-6840.

[62]: Inhee C., Choonmi K., Sun C. Binding Mode Analysis of Topoisomerase Inhibitors, 6-Arylamino-7-chloro-quinazoline-5,8-diones, within the Cleavable Complex of human Topoisomerase I and DNA. *Arch Pharm Res*. 2007; **30**: 1526-1535.

[63]: Arjun B., Pritam T., Radha K et al. 2,6-Dithienyl-4-furyl pyridines: Synthesis, topoisomerase I and II inhibition, cytotoxicity, structure–activity relationship, and docking study. *Bioorganic & Medicinal Chemistry Letters*. 2010; **20**: 42–47.

[64] : Karen J. H., Elizabeta P., Howard S. H. Recent developments in the clinical activity of topoisomerase-1 inhibitors. *Update on Cancer Therapeutics*.2006; **1**: 117–145.

[65]: Gautier M. Etude structurale du domaine d'interaction du récepteur de l'élastine. Approches biochimiques, biophysiques et bioinformatiques. Thèse de doctorat d'université : Biomolécules et Dynamique Cellulaire. Reims Champagne-Ardenne : Université de Reims Champagne-Ardenne. France. 2005. 240 p.

[66]: Thiriote E. Modélisation de la reconnaissance et de la catalyse enzymatiques : développements méthodologiques et détermination du mécanisme des Méthionine Sulfoxyde Réductases de classe A. thèse de doctorat d'université : Chimie informatique et théorique. Nancy : Université de Nancy. France. 2009. 291 p.

[67] : Tocino Diaz JC. Développement d'un programme d'amarrage protéine-protéine. Thèse Master: Bioinformatique. Bruxelles : Université Libre de Bruxelles. Belgique. 2000. 72p.

[68]: Arrault A. Stratégies de docking-scoring assistées par analyse de données. Application au criblage virtuel des cibles thérapeutiques COX-2 et PPAR gamma. Thèse de doctorat d'université : Modélisation Moléculaire et Chemoinformatique. Orleans : Université d'Orleans. France. 2007. 176 p.

[69]: Morris G. M., Goodsell D. S., Halliday R. S. et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem*.1998; **19** : 1639-1662.

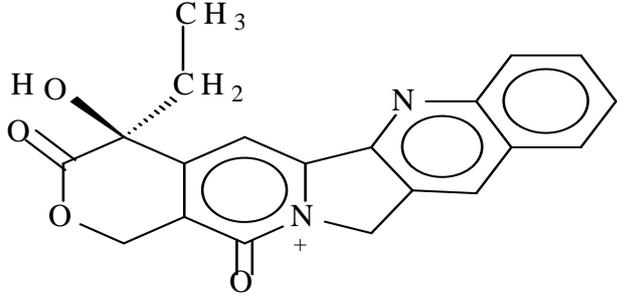
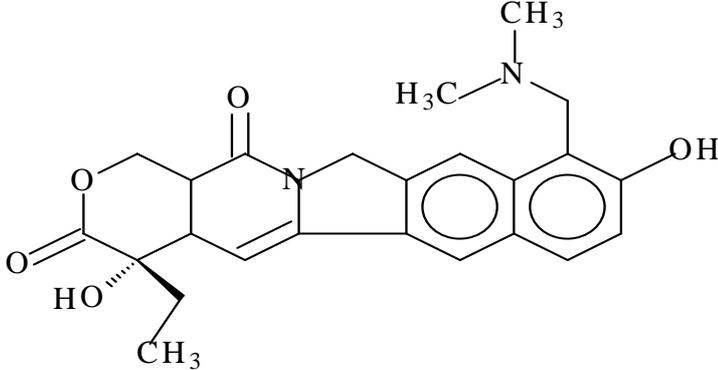
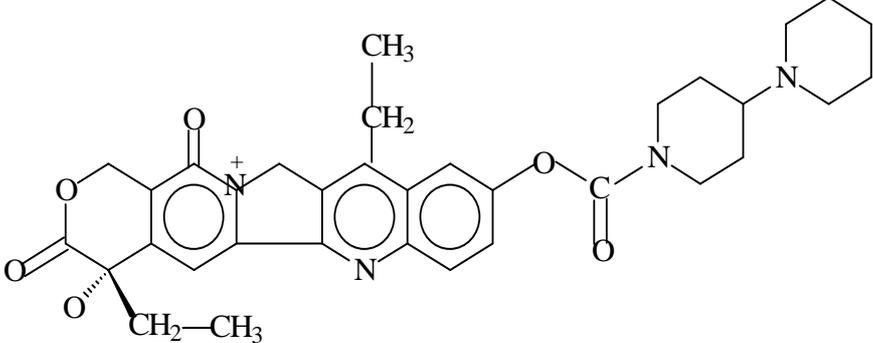
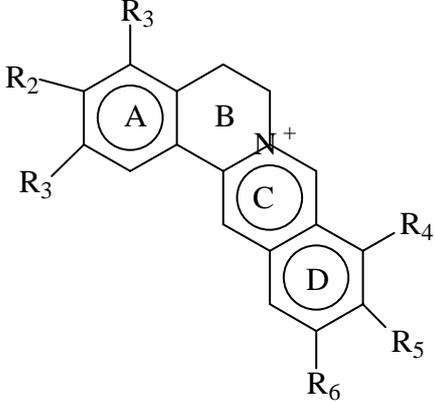
- [70]: Rarey M., Kramer B., Lengauer T., Klebe G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol.* 1996; **261** : 470-489.
- [71]: Jones G., Willett P., Glen R. C., Leach A. R., Taylor R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 1997; **267** : 727-48.
- [72]: Ewing T.J., Makino S., Skillman A. G., Kuntzi. D. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases., *J. Comput. Aided. Mol. Des.* 2001, **15** : 411-428.
- [73]: Jay A., Jain N. Surflex-Dock 2.1: Robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J Comput Aided Mol Des.* 2007, **21**:281–306.
- [74]: Grosdidier A. Conception d'un logiciel de docking et applications dans la recherche de nouvelles molécules actives. Thèse de doctorat en pharmacie. Grenoble : Université Joseph Fourier. France. 2007. 90p.
- [75]: Chikhi A. W. Calculs et modélisations des interactions peptide deformylase –substances antibactériennes à l'aide de techniques de "docking"(arrimage) moléculaire. Thèse Doctorat d'université : Biochimie appliquée. Constantine : Université Mentouri Constantine. Algérie. 2007. 126p.
- [76]: Staker B.L., Feese M.D., Cushman M., Pommier Y., Zembower D., Stewart L., Burgin A.B. Structures of three classes of anticancer agents bound to the human topoisomerase I-DNA covalent complex. *J Med Chem.* 2005; **48** : 2336- 2345.
- [77]: Dewar M. J. S., Zoebisch E. G., Healy E. F., Stewart J. J. P. AM1: A New General Purpose Quantum Mechanical Model. *Journal of the American Chemical Society.* 1985; **107**: 3902-3909.
- [78]: Chikhi A., Bensegueni A. Comparative Study of the Efficiency of Three Protein-Ligand Docking Program. *Journal of Proteomics and Bioinformatics.*2008; **1**: 161-165.
- [79]: Vieth M., Hirst J. D., Kolinski A., Brooks C. L. Assessing energy functions for flexible docking. *Journal of Computational Chemistry.* 1998; **19**: 1612-22.
- [80]: Lipinski C.A., Lombardo F., Domini B.W., Feney P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997; **23** : 3-25.

- [81]: Constantino A., Mehta R., Runyan C., Rao K., Vaughan A., Moon R. Flavonoids as DNA topoisomerase antagonists and poisons : structure-activity relationships. *J. Nat. Prod.* 1995; **58** : 217-225.
- [82]: Akbas SH, Timur M, Ozben T. The effect of quercetin on topotecan cytotoxicity in MCF-7 and MDA-MB 231 human breast cancer cells. *J. Surg. Res.* 2005, 125 (1): 49-55.
- [83]: Paliwal S, Sundaram J, Mitragotri S. Induction of cancer-specific cytotoxicity towards human prostate and skin cells using quercetin and ultrasound. *Brit. J. Cancer* 2005, 92: 499-502.
- [84]: Boege F.; Straub T., Kehr A., Boesenberg C., Christiansen K., Andersen A., Jakob F., Kohrle J. Selected novel flavones inhibit the DNA binding or the DNA relegation step of eukaryotic topoisomerase. *J.Biol. Chem.* 1996; **271**: 2262-2270.
- [85]: Michael R., Webb and Susan E., Ebeler. Comparative analysis of topoisomerase IB inhibition and DNA intercalation by flavonoids and similar compounds : structural determinates of activity. *Biochem. J.* 2004; **384** : 527-541.
- [86]: McMurry J. Organic chemistry. USA. Thomson. 1996. 1243p.

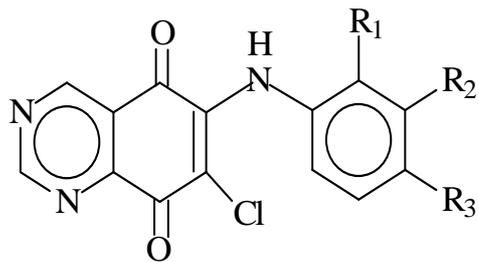
Annexes

Annexe 1

Structures de quelques poisons de l'AND topoisomérase I

Poisons	Structures
Camptothécine	 <p>The structure of Camptothecin is a complex pentacyclic alkaloid. It features a tropane-like bicyclic core (8-methyl-8-azabicyclo[3.2.1]octane) fused to a pyridine ring. This pyridine ring is further fused to a cyclohexene ring, which is substituted with a methyl group and a hydroxyl group. The cyclohexene ring is also fused to a piperidine ring, which is substituted with a methyl group and a hydroxyl group. The piperidine ring is fused to a tropane bicyclic system.</p>
Topotecan	 <p>The structure of Topotecan is a pentacyclic alkaloid. It features a tropane-like bicyclic core (8-methyl-8-azabicyclo[3.2.1]octane) fused to a pyridine ring. This pyridine ring is further fused to a cyclohexene ring, which is substituted with a methyl group and a hydroxyl group. The cyclohexene ring is also fused to a piperidine ring, which is substituted with a methyl group and a hydroxyl group. The piperidine ring is fused to a tropane bicyclic system.</p>
Irinotecan	 <p>The structure of Irinotecan is a pentacyclic alkaloid. It features a tropane-like bicyclic core (8-methyl-8-azabicyclo[3.2.1]octane) fused to a pyridine ring. This pyridine ring is further fused to a cyclohexene ring, which is substituted with a methyl group and a hydroxyl group. The cyclohexene ring is also fused to a piperidine ring, which is substituted with a methyl group and a hydroxyl group. The piperidine ring is fused to a tropane bicyclic system. The tropane bicyclic system is further substituted with a methyl group and a hydroxyl group.</p>
Protoberbérines	 <p>The structure of Protoberberines is a pentacyclic alkaloid. It features a tropane-like bicyclic core (8-methyl-8-azabicyclo[3.2.1]octane) fused to a pyridine ring. This pyridine ring is further fused to a cyclohexene ring, which is substituted with a methyl group and a hydroxyl group. The cyclohexene ring is also fused to a piperidine ring, which is substituted with a methyl group and a hydroxyl group. The piperidine ring is fused to a tropane bicyclic system. The tropane bicyclic system is further substituted with a methyl group and a hydroxyl group.</p>

dérivés du 6-Arylamino-7-chloro-quinazoline-5,8-diones



Annexe 2

Fichier PDB 1T8I

```
HEADER      ISOMERASE/DNA                                12-MAY-04
1T8I
TITLE       HUMAN DNA TOPOISOMERASE I (70 KDA) IN COMPLEX WITH THE
TITLE       2 POISON CAMPTOTHECIN AND COVALENT COMPLEX WITH A 22
BASE
TITLE       3 PAIR DNA DUPLEX
DBREF      1T8I A   174   765 UNP      P11387   TOP1_HUMAN   174
765
DBREF      1T8I B     1    10  PDB      1T8I     1T8I         1
10
DBREF      1T8I C    11    22  PDB      1T8I     1T8I        11
22
DBREF      1T8I D   101   122  PDB      1T8I     1T8I       101
122
SEQADV     1T8I PTR A   723 UNP      P11387     TYR    723 MODIFIED RESIDUE
SEQRES     1 B    10   DA  DA  DA  DA  DA  DG  DA  DC  DT  DT
SEQRES     1 C    12  TGP  DG  DA  DA  DA  DA  DA  DA  DT  DT  DT  DT
DT
SEQRES     1 D    22   DA  DA  DA  DA  DA  DT  DT  DT  DT  DT  DC
DC  DA
SEQRES     2 D    22   DA  DG  DT  DC  DT  DT  DT  DT  DT  DT
SEQRES     1 A   592  LYS  LYS  PRO  LYS  ASN  LYS  ASP  LYS  ASP  LYS  LYS
VAL  PRO
SEQRES     2 A   592  GLU  PRO  ASP  ASN  LYS  LYS  LYS  LYS  PRO  LYS  LYS
GLU  GLU
SEQRES     3 A   592  GLU  GLN  LYS  TRP  LYS  TRP  TRP  GLU  GLU  GLU  ARG
TYR  PRO
SEQRES     4 A   592  GLU  GLY  ILE  LYS  TRP  LYS  PHE  LEU  GLU  HIS  LYS
GLY  PRO
SEQRES     5 A   592  VAL  PHE  ALA  PRO  PRO  TYR  GLU  PRO  LEU  PRO  GLU
ASN  VAL
SEQRES     6 A   592  LYS  PHE  TYR  TYR  ASP  GLY  LYS  VAL  MET  LYS  LEU
SER  PRO
SEQRES     7 A   592  LYS  ALA  GLU  GLU  VAL  ALA  THR  PHE  PHE  ALA  LYS
MET  LEU
SEQRES     8 A   592  ASP  HIS  GLU  TYR  THR  THR  LYS  GLU  ILE  PHE  ARG
LYS  ASN
SEQRES     9 A   592  PHE  PHE  LYS  ASP  TRP  ARG  LYS  GLU  MET  THR  ASN
GLU  GLU
SEQRES    10 A   592  LYS  ASN  ILE  ILE  THR  ASN  LEU  SER  LYS  CYS  ASP
PHE  THR
SEQRES    11 A   592  GLN  MET  SER  GLN  TYR  PHE  LYS  ALA  GLN  THR  GLU
ALA  ARG
SEQRES    12 A   592  LYS  GLN  MET  SER  LYS  GLU  GLU  LYS  LEU  LYS  ILE
LYS  GLU
SEQRES    13 A   592  GLU  ASN  GLU  LYS  LEU  LEU  LYS  GLU  TYR  GLY  PHE
CYS  ILE
```


SEQRES 41 A 592 LEU GLU GLU GLN LEU MET LYS LEU GLU VAL GLN
 ALA THR
 SEQRES 42 A 592 ASP ARG GLU GLU ASN LYS GLN ILE ALA LEU GLY
 THR SER
 SEQRES 43 A 592 LYS LEU ASN PTR LEU ASP PRO ARG ILE THR VAL
 ALA TRP
 SEQRES 44 A 592 CYS LYS LYS TRP GLY VAL PRO ILE GLU LYS ILE
 TYR ASN
 SEQRES 45 A 592 LYS THR GLN ARG GLU LYS PHE ALA TRP ALA ILE
 ASP MET
 SEQRES 46 A 592 ALA ASP GLU ASP TYR GLU PHE
 MODRES 1T8I TGP C 11 DG
 MODRES 1T8I PTR A 723 TYR O-PHOSPHOTYROSINE
 HET TGP C 11 19
 HET PTR A 723 16
 HET EHD D 990 26
 HETNAM TGP 5'-THIO-2'-DEOXY-GUANOSINE PHOSPHONIC ACID
 HETNAM PTR O-PHOSPHOTYROSINE
 HETNAM EHD 4-ETHYL-4-HYDROXY-1,12-DIHYDRO-4H-2-OXA-6,12A-
 DIAZA-
 HETNAM 2 EHD DIBENZO[B,H]FLUORENE-3,13-DIONE
 HETSYN PTR PHOSPHONOTYROSINE
 HETSYN EHD CAMPTOTHECIN
 FORMUL 2 TGP C10 H14 N5 O6 P S
 FORMUL 4 PTR C9 H12 N O6 P
 FORMUL 5 EHD C20 H16 N2 O4

Annexe 3

Liste des 134 complexes et leur RMSD utilisés pour tester le programme Surfex

NUMERO	CODE	RMSD
1	1C1C	0.376
2	1COU	0.446
3	1FKO	0.051
4	1IKV	0.268
5	1TKT	0.323
6	2B5J	0.298
7	2RKI	0.308
8	2VG5	0.368
9	2VG6	0.450
10	3DLG	0.342
11	3FFI	0.252
12	3ISN	0.053
13	3M8P	0.493
14	1KCI	0.299
15	2WIH	0.409
16	2FHR	0.409
17	3FMM	0.483
18	3KWF	0.475
19	3D4L	0.372
20	1BQM	0.532
21	1BQN	0.868
22	1DTT	0.662
23	1EET	0.557
24	1EP4	0.648
25	1HNI	0.858
26	1HNV	0.565
27	1S1W	0.593
28	1S6P	0.517
29	1TV6	0.828
30	2BAN	0.756
31	2BE2	0.713
32	2RF2	0.633
33	2VG7	0.596
34	2WON	0.772
35	2ZE2	0.584
36	3DM2	0.868
37	3DOK	0.919
38	3DOL	0.923
39	3IOR	0.658
40	3LAK	0.526
41	3NBP	0.562
42	3EBI	0.782
43	3EBH	0.807
44	3HVC	0.708
45	2ZOQ	0.851
46	3GC8	0.859
47	2HPT	0.936
48	2EK9	0.729
49	2DQM	0.767
50	2B3L	0.882
51	1MU0	0.955
52	1RL4	0.999

53	3FMR	0.736
54	3E3U	0.882
55	3GCU	0.580
56	3F8S	0.798
57	3E43	0.882
58	2RIP	0.722
59	1YW9	0.938
60	1COT	1.379
61	1DTQ	1.432
62	1TKX	1.688
63	1TKZ	1.048
64	1VRT	1.458
65	2WOM	1.168
66	3C6T	1.113
67	3IOS	1.024
68	1AJ6	1.046
69	1ZXM	1.296
70	3FMQ	1.264
71	1GCQ	1.286
72	3GC7	1.016
73	2GU6	1.451
74	3D44	1.426
75	2B3H	1.119
76	1XRY	1.452
77	2EWS	1.402
78	1G2A	1.385
79	1BSJ	1.449
80	3B7U	1.435
81	3B7R	1.269
82	3CHR	1.550
83	2P98	1.492
84	1C1B	1.535
85	1TL1	1.937
86	1VRU	1.656
87	3BGR	1.904
88	3C6U	1.839
89	3DI6	1.739
90	3DLE	1.794
91	3DRP	1.628
92	3E01	1.469
93	3HVT	1.709
94	3IRX	1.989
95	3ITH	1.219
96	3LAL	1.539
97	1EG6	1.610
98	3B3V	1.952
99	2PRQ	1.763
100	3FDG	1.774
101	1X2B	1.582
102	1TKF	1.646
103	1LOK	1.623
104	1KQ0	1.966
105	1FT7	1.536
106	2AI7	1.660
107	3GCS	1.822
108	2Q93	1.856
109	3KED	1.636
110	2EVC	1.623
111	1YW8	1.889
112	1R5H	1.565
113	1R5G	1.790

114	1BOA	1.935
115	1FK9	2.123
116	2ZD1	2.153
117	2DL8	2.376
118	2EK8	2.088
119	1XRL	2.502
120	1TKH	2.332
121	1CP6	2.091
122	2W3U	2.095
123	1S17	2.173
124	2SV5	2.815
125	2GVS	3.091
126	2GG5	3.161
127	1YJ3	3.617
128	1XBU	3.501
129	2GG7	2.966
130	1MT3	2.634
131	1KQ9	2.538
132	1C27	3.220
133	3D27	3.154
134	1BSZ	2.171

Annexe 4

Résultats de l'affinité des quatre flavonoïdes et le composé 28 obtenu par le programme Surflex

Myricétine

```
final-0: 7.01 (crash -1.25 (self -0.53))  
final-1: 6.77 (crash -1.11 (self -0.43))  
final-2: 6.43 (crash -1.26 (self -0.50))  
final-3: 6.29 (crash -1.21 (self -0.63))  
final-4: 6.28 (crash -0.96 (self -0.76))  
final-5: 6.20 (crash -1.20 (self -0.63))  
final-6: 5.99 (crash -1.65 (self -0.81))  
final-7: 5.86 (crash -1.13 (self -0.31))  
final-8: 5.86 (crash -1.13 (self -0.31))  
final-9: 5.86 (crash -1.13 (self -0.31))  
  
C:\Users\TENTOU\Desktop\fin\etude pratique\myricétine>_
```

Fiséine

```
final-0: 6.09 (crash -0.97 (self -0.43))  
final-1: 5.83 (crash -0.96 (self -0.45))  
final-2: 5.44 (crash -0.65 (self -0.06))  
final-3: 5.44 (crash -0.65 (self -0.06))  
final-4: 5.09 (crash -1.47 (self -0.58))  
final-5: 5.08 (crash -1.09 (self -0.49))  
final-6: 5.06 (crash -1.05 (self -0.60))  
final-7: 4.62 (crash -1.85 (self -0.37))  
final-8: 3.85 (crash -0.88 (self -0.00))  
final-9: 3.60 (crash -0.54 (self -0.18))  
  
C:\Users\TENTOU\Desktop\fin\etude pratique\Fiséine>_
```

Quercétine

```
final-0: 5.53 (crash -0.41 (self -0.05))  
final-1: 5.53 (crash -0.41 (self -0.05))  
final-2: 4.99 (crash -0.67 (self -0.32))  
final-3: 4.97 (crash -0.89 (self -0.11))  
final-4: 4.75 (crash -0.59 (self -0.20))  
final-5: 4.73 (crash -0.52 (self -0.36))  
final-6: 4.68 (crash -0.87 (self -0.15))  
final-7: 4.49 (crash -0.54 (self -0.29))  
final-8: 4.12 (crash -0.88 (self -0.27))  
final-9: 4.11 (crash -0.79 (self -0.59))  
  
C:\Users\TENTOU\Desktop\fin\etude pratique\Quercétine>
```

Apigénine

```
(4) [ret_conf 1.89] [ret_conf -3.58] final-0: 4.16 (crash -1.86 (self -0.44))
final-1: 3.15 (crash -3.42 (self -0.45))
final-2: 2.78 (crash -1.45 (self -0.28))
final-3: 2.67 (crash -2.18 (self -0.91))
final-4: 2.47 (crash -4.28 (self -0.81))
final-5: 2.46 (crash -2.27 (self -0.91))
final-6: 2.21 (crash -1.73 (self -0.81))
final-7: 2.08 (crash -3.39 (self -0.48))
final-8: 1.71 (crash -2.82 (self -0.64))
final-9: 1.66 (crash -2.92 (self -0.78))

C:\Users\TENTOU\Desktop\fin\étude pratique\apigénine)
```

Composé 28

```
final-0: 8.79 (crash -1.55 (self -0.83))
final-1: 8.65 (crash -1.65 (self -0.81))
final-2: 8.51 (crash -1.52 (self -0.69))
final-3: 8.13 (crash -1.51 (self -0.83))
final-4: 8.06 (crash -1.41 (self -0.84))
final-5: 7.99 (crash -1.36 (self -0.63))
final-6: 7.49 (crash -0.85 (self -0.57))
final-7: 7.15 (crash -1.83 (self -0.87))
final-8: 7.12 (crash -0.92 (self -0.76))
final-9: 6.93 (crash -1.56 (self -0.87))

C:\Users\TENTOU\Desktop\fin\étude pratique\Composé 28)
```

Annexe 5

Résultats de l'application des propriétés physico-chimiques sur le composé 28

Your request parameters:

```
data format : smiles
  Data size : 1 compounds.
  Toxic atoms :
Molecular weight : min 200.0 max 600.0
  Hydrogen donors : min 0.0 max 6.0
Hydrogen acceptors : min 0.0 max 12.0
  Flexible bonds : min 0.0 max 15.0
  Rigid bonds : min 0.0 max 50.0
  Ring number : min 0.0 max 7.0
  Ring size : min 0.0 max 12.0
  Atom number : min #carbons: 5.0 min #non carbons 2.0
Ratio carbon/herero : min 0.1 max 1.0
  Charge number : min 0.0 max 3.0
  Total charge : min -2.0 max 2.0
    logP : min -2.0 max 6.0
  Polar Surface Area : min 0.0 max 150.0
  Results URL : http://bioserv.rpbs.jussieu.fr/tmp/ADMETox/469124895
Will proceed ...
```

List of compounds passing the filter:

```
C1(C(=C(c2cc(c(c(c2)CO)O)O)Oc3c1c(cc(c3)O)O)O)=O
C:\Users\TENIOU\Desktop\34\34.pdb
```

Detailed output for each compound:

```
MW : Moléculaire weight
Drs : Donors
Ars : Acceptors
FB : flexible bonds
RB : Rigid Bonds
#R : Ring Number
RL : Ring Length
C : carbons
nC : non carbons
C/nC : ratio non carbons/carbons
#Chrg : number of charges
Chrg: Total Charge
LogP : logP (octanol / water)
PSA: Polar surface area
```

```
C1(C(=C(c2cc(c(c(c2)CO)O)O)Oc3c1c(cc(c3)O)O)O)=O
C:\Users\TENIOU\Desktop\34\34.pdb : MW : 332.2 Drs : 6 Ars : 8 FB : 2 RB : 18
#R : 3 RL : 6 C : 16 nC : 8 C/nC : 0.500000 #Chrg : 0 Chrg : 0 LogP :
2.220000 PSA : 147.680000
```

Les résumés

Thème : conception *in silico* de nouveaux inhibiteurs de la topoisomérase I de cellules cancéreuses

Résumé

L'ADN Topoisomérase I humaine (topo I) est la cible de plusieurs drogues utilisées en chimiothérapie cancéreuse. Ces drogues forment des complexes ternaires "topo-ADN-drogue" entraînant la mort des cellules cancéreuses. Plusieurs flavonoïdes ont montré une activité anticancéreuse en inhibant la topo I.

En faisant appel aux méthodes de modélisation moléculaire plus particulièrement le docking par le programme Surflex, nous avons évalué l'activité inhibitrice de quatre flavonoïdes : décrits récemment par la littérature comme inhibiteurs potentiels de la topo I. Il s'agit de : la myricétine, la fisétine, la quercétine et l'apigénine. La Myricétine présentant l'IC₅₀ le plus bas ($11.1 \pm 2.0 \mu\text{M}$) a donné par simulation avec Surflex la meilleure affinité d'interaction (7.01 M^{-1}). Afin de développer des nouvelles molécules plus puissantes que la myricétine, nous avons réalisé sur cette dernière plusieurs substitutions. Le changement de l'hydroxyle en C3' de la myricétine par un groupement alcoolique CH₂OH améliore nettement l'affinité qui passe de 7.01 M^{-1} à 8.79 M^{-1} .

L'application de la règle de 5 de Lipinski nous renseigne de manière positive sur les propriétés ADME de cette nouvelle molécule qui se présente comme un inhibiteur potentiellement plus actif que la myricétine.

Mots clés: L'ADN Topoisomérase I; Flavonoïdes; Modélisation moléculaire; Docking moléculaire; Surflex; Affinité; Myricétine.

Title: design *in silico* of new inhibitors of the DNA topoisomerase I of cancer cells

Summary

Human DNA Topoisomérase I (topo I) is the target of several drugs commonly used in cancer chemotherapy. These drugs induce topo-DNA complexes that eventually trigger cell death. Several flavonoides have been shown as inhibitors of topo I, having an anticancer activity.

Using the methods of molecular modeling in particular the docking by the Surflex program, we theoretically evaluated the affinity of four flavonoids: myricetin, fisetin, quercetin and apigenin, presented in recent studies as inhibitors of topo I. Myricetin with the lowest IC_{50} ($11.1 \pm 2.0 \mu\text{M}$) gave the best affinity (7.01 M^{-1}). For the development *in silico* of novel molecules most powerful myricetin, we have made several substitutions. Replacing the hydroxyl group carried by the carbone C3' of the myricetin by an alcoholic group CH_2OH significantly improves the affinity increases from 7.01 M^{-1} to 8.79 M^{-1} .

The application of the rule of Lipinski informs in a positive way about the ADME properties of this new molecule that is as an inhibitor potentially more active than myricetin.

Key words: DNA topoisomérase I; Flavonoids; Molecular modeling; Molecular Docking; Surflex; Affinity; Myricetin.

العنوان: اقتراح مثبطات جديدة لـ ADN topoisomérase I للخلايا السرطانية باستعمال برامج النمذجة الجزيئية

الملخص

ADN topoisomérase I هو إنزيم مستهدف من طرف عدة أدوية تستعمل عادة في العلاج الكيميائي للسرطان. تشكل هذه الأدوية مركبات ثلاثية " إنزيم-ADN-دواء " متسببة في موت الخلايا السرطانية. وقد ثبت من خلال دراسات علمية أن العديد من الفلافونويدات لها القدرة على تثبيط الإنزيم topoI و قتل الخلايا السرطانية .

باستعمال طرق النمذجة الجزيئية بالأخص Docking بواسطة البرنامج Surflex قمنا بالتقييم النظري لجاذبية أربعة فلافونويدات " myricétine ، fisétine ، quercétine ، apigénine "، علما أن هذه الجزيئات هي مثبطات للإنزيم topoI بحسب الدراسات الحديثة. المركب myricétine له IC_{50} الأقل بقيمة (11.1 ± 2.0) ميكرومول) والتي أعطت أفضل جاذبية للإنزيم ($7.01 M^{-1}$). من اجل البحث نظريا حسب النمذجة الجزيئية عن جزيئات أكثر فعالية من myricétine قمنا بإدخال مستبدلات على مختلفة على هذه الأخيرة. استبدال مجموعة الهيدروكسيل المحمولة على الكربون 3' بمجموعة كحولية CH_2OH أدى إلى ارتفاع الجاذبية من $7.01 M^{-1}$ إلى $8.79 M^{-1}$.

تطبيق قاعدة Lipinski أعطت نتائج ايجابية خاصة بالخصائص ADME لهذه الجزيئة الجديدة التي تعتبر مثبط أكثر فعالية من myricétine.

الكلمات المفتاحية :

ADN topoisomérase I، الفلافونويدات، النمذجة الجزيئية، Docking، Surflex، الجاذبية، Myricétine.