

REPUBLIQUE ALGERIENNE DEMOCRATIQUE et POPULAIRE
MINISTERE de l'ENSEIGNEMENT SUPERIEUR et de la RECHERCHE
SCIENTIFIQUE
UNIVERSITE MOULOUD MAMMERI de TIZI-OUZOU

FACULTE des SCIENCES

DEPARTEMENT de CHIMIE



MEMOIRE

Présenté par

Melle. Malki Djaouida

Pour l'obtention du diplôme de

MASTER

Spécialité : Chimie

Option : Chimie Pharmaceutique

Thème

Etude de mode de liaison de guanosine diphosphate et guanosine triphosphate avec l'Arf1, protéine impliquée dans le cancer du sein, en présence et en absence de Mg^{2+}

Devant le jury composé de:

M. SAAL Amar
M. GUECHTOULI Nabila
M. BENKHEMOU Malika
M. Bouaziz-Terrachet Souhila

M.C.B à l'UMMTO
M.C.B à l'UMMTO
M.A.A à l'UMMTO
M.C.B à l'UMBB

Président
Examinatrice
Co-Promotrice
Promotrice

Promotion 2015/2016

Dédicaces

Je dédie ce travail après une année de compétence et de sérieux à :

-ceux qui m'ont offert la vie :

Mon père (que Dieu le protège)

Ma mère qui (Que Dieu la protège)

- Mon cher mari avec qui je partagerais le reste de ma vie Ferhat et sa famille.

- Mes chers frères :

Mehenna, Arezki, Sofiane et sa femme Farida et leur petit ange Yastene.

- Mes sœurs :

Saida, Hamama (avec ses adorables filles Amal, Lila, Sabiha), Rosa (avec ses deux trésors Yacine et Adel), Akila et son mari Hakim.

-Sans oublier tous ceux qui ont une grande valeur dans mon cœur, mes amis.

Remerciement

Avant tout, nous remercions le dieu qui nous a donné la force et le courage pour réaliser et achever ce travail dans des bonnes conditions.

Je tiens à remercies chaleureusement et profondément madame **Souhila BOUAZIZ-TERRACHET**, Maître de conférences à l'UMBB pour m'avoir assuré l'encadrement de mon mémoire de Master. Merci énormément pour votre soutien, votre confiance, vous orientation judicieuse et aussi votre disponibilité pour accomplir mon travail de recherche, je vous suis très reconnaissante. Merci encore madame **BOUAZIZ**.

Je tiens aussi mes remerciements les plus sincères à ma Co-promotrice Madame **MALIKA BENKHEMOU**, chargée de cours à l'UMMTO pour sa générosité, sa gentillesse, sa patience et ses précieux conseils pour l'établissement de ce travail.

J'adresse mes vifs remerciements à monsieur **SAAL AMAR**, Maître de conférences à l'université de Mouloud Mammeri, qui m'a fait l'honneur de présider le jury de ce mémoire.

Toute ma gratitude va à Madame **NABILA GUECHTOULI**, Maitre de conférences à l'UMMTO, pour avoir accepté de participer à ce jury et d'avoir accepté de lire et de corriger ce mémoire.

Je ne saurais oublier dans ces remerciements la Direction de la Faculté de Chimie, Monsieur le Doyen, le personnel technique et administratif ainsi que tous ceux qui, de près ou de loin ont contribué à la réalisation de ce travail.

TABLE DES MATIERES

Dédicaces

Remerciements

Liste des abréviations

Liste des figures

Liste des tableaux

Introduction Générale1

Références bibliographiques.....3

Chapitre I : Rappel bibliographiques sur la protéine ADP-Ribosylation factor1 (Arf1)

I.	Introduction.....	4
II.	Définition des protéines G.....	4
II.1-	Les protéines G hétérotrimériques.....	4
II.2-	Les protéines G monomériques.....	5
II.2.1-	Les protéines G monomériques de la famille EF-TU.....	5
II.2.2-	La super famille des protéines Ras.....	5
a-	Protéine RAS.....	6
b-	La famille des RHO.....	6
c-	Les membres de la famille RAN.....	7
d-	La famille RAB.....	7
e-	La famille des ARF.....	7
•	La sous famille Arl.....	7
•	La protéine SARs.....	7
•	La protéine ARD1.....	8
•	La protéine ARP.....	8
•	La sous-famille Arf.....	8
III.	Les protéines Arf1.....	9
III.1-	Définition.....	9
III.2-	Structure de l'Arf1.....	9
IV.	Mécanisme d'activation et d'inactivation d'ARF.....	11
IV.1-	Le guanosine triphosphate (GTP).....	12
IV.2-	Le guanosine diphosphate (GDP).....	13
IV.3-	Les ARF-GEFs.....	13
IV.4-	Les ARF-GAPs.....	13
IV.3-	Le cycle GDP/GTP de l'Arf1.....	14
V.	Rôle de Mg ²⁺ dans le mécanisme d'Arf1.....	14
	Références bibliographiques.....	16

Chapitre II : Les principales approches de la modélisation moléculaire

I.	Introduction	18
II.	Les méthodes de la chimie théorique	18
II.1-	Les méthodes de la chimie quantiques	18
II.1.1-	Equation de Schrödinger	18
II.1.2-	Approximation de Born-Oppenheimer	19
II.1.3-	Approximation orbitale	20
II.1.3-	Approximation LCAO	21
II.1.4-	Les différentes méthodes de la chimie quantiques	21
a-	Les méthodes ab-initio	21
b-	Les méthodes de la théorie fonctionnelle de la densité	22
II.2-	Les méthodes de la mécanique moléculaire	22
II.2.1-	Définition de la mécanique moléculaire	22
II.2.2-	Champ de force.....	23
II.2.3-	Les différentes méthodes de la mécanique moléculaire	24
a-	La méthode de steepest descent.....	24
b-	La méthode du gradient conjugué	24
c-	Méthode de Newton-Raphson	24
d-	Recuit Simulé	25
e-	Monté Carlo.....	25
III-	Les méthodes de drug design.....	25
III.1-	Le QSAR.....	26
III.2-	Le docking	26
IV-	Le logiciel Autodock	26
IV.1-	Présentation	26
IV.2-	AutoDockTools (ADT).....	27
IV.3-	Principe d'utilisation d'AutoDock.....	27
IV.3.1-	Préparation du récepteur	27
IV.3.2-	Description de la macromolécule sous la forme de grilles	28
IV.3.3-	Préparation des ligands	29
IV.3.4-	Recherche des conformations de docking.....	29
IV.3.5-	Paramètres d'un algorithme génétique.....	29
IV.3.6-	Analyse et évaluation des résultats.....	30
	Références bibliographiques.....	32

Chapitre III : Etude par docking moléculaire du rôle de l'ion Mg²⁺ dans la stabilité des complexes GDP/Arf1 et GTP/Arf1

I.	Introduction	34
II.	Méthodologie.....	34
III.	II.1- Structures de départ	34
II.1.1-	Préparation des structures de la protéine.....	34
II.1.2-	Préparation des ligands	35
II.2.	Méthodologie utilisée	35
II.2. 1-	Préparation des structures du récepteur	35

II.2.2- Localisation du site actif.....	36
II.2. 3- Démarche suivie dans les calculs de docking.....	36
a- Paramètres de calcul du grid	36
b- Paramètres de calculs de docking.....	37
IV. Résultats et discussion	37
III.1- Validation des paramètres de calcul utilisés	37
III.2- Etude énergétique des conformations obtenue	39
III.3- Etude des liaisons Hydrogènes	42
III.4-Etude des distances interatomiques entre l'ion Mg^{2+} et les ligands GTP et GDP	49
V. Conclusion	51
Références bibliographiques.....	53
Conclusion Générale.....	54

ABREVIATIONS

ADT: AutoDockTools.

AG: algorithme génétique.

AGL : algorithme génétique lamarckien.

AM1: Austin Model 1.

AMBER: Assisted Model Building and Energy Refinement.

ARD1 : Arf-Ribosylation Domain1.

Arf1 : ADP ribosylation facteur 1.

Arl: Arf like.

ARP: ARF-Related Protein.

BO: Born-Oppenheimer.

CFF: Consistent Force Field.

CHARMm: Chemistry at HARvard Macromolecular Mechanics.

CVFF: Consistent Valence Force Field.

DFT: théorie de la fonctionnelle de la densité.

Dpf: Dock Parameters File.

EF-TU: Elongation factor thermo unstable.

eV : électronvolt.

GAP: GTPase activating protein.

GDP: guanosine diphosphate.

GEF: guanosine nucleotide exchange factor.

gpf: Grid Parameters File.

GROMOS: GROnigen MOlecular Simulation.

GTP: guanosine triphosphate.

kDA: kilo Daltons.

LCAO: Linear Combination of Atomic Orbitals.

MC: Monté Carlo.

MM: mécanique moléculaire.

MMFF: Merck Molecular Force Field.

MNDO: Modified Neglect of Diatomic Differential Overlap.

OA: orbitales atomiques.

OPLS: Optimised potentials for Liquid Simulation.

Pdb: Protein Data Bank.

PM3: Parametric Method 3.

QSAR: Quantitative Structure Activity Relationship.

Rab: Ras-like nuclear protein in brin

Ran: Ras-like nuclear protein.

Ras: Rat Sarcoma.

Rho: Ras homologue.

RSMC: Monte-Carlo avec Recuit Simulé.

Sar1: Secretion-Associated and Ras-superfamily-related gene.

SD : steepest descent.

SE: Méthodes Semi-Empiriques.

SPASIBA: Spectroscopic Potentiel Algorithm for Simulating Biomolecular conformational Adaptability.

VDW: Van Der Waals.

LISTE DES FIGURES

Figure 1: la superfamille Ras des protéines G	6
Figure 2 : Caractéristiques structurales d'arf1. (A) Structure 3D de l'arf1-GDP, (B) Séquence en acides aminés de l'arf1	10
Figure 3 : mécanisme d'activation de l'ARF	12
Figure 4 : Structure de GTP.....	12
Figure 5 : Structure de GDP	13
Figure 6 : Mécanisme d'activation d'arf1 par sa GEF en présence de la membrane	14
Figure 7 : a) un exemple de liaison de la guanine (G) avec une base nucléique (la cytosine (C)) ; b) la structure de la guanine (G).....	15
Figure 8 : Interface graphique d'AutoDock	27
Figure 9: Schéma d'une grille de potentiel englobant le site actif d'une macromolécule	28
Figure 10: les différents outils d'analyse de résultats de docking disponibles dans ADT. Les histogrammes à gauche représentent les clusters	31
Figure 11 : (A) Structure de GTP et (B) Structure de GDP.....	35
Figure 12 : superposition des deux conformations de A) GDP et B) GTP dans le site actif de l'Arf1 : les conformations issues du docking sont colorées en vert et les conformations cristallographiques en rouge.....	38
Figure 13 : Classification des conformations obtenues par les calculs de docking pour le ligand GTP A) en absence et B) en présence de Mg^{2+} selon l'énergie croissante.....	40
Figure 14 : Classification des conformations obtenues par les calculs de docking pour le ligand GDP A) en présence et B) en absence de Mg^{2+} selon l'énergie de liaison croissante.....	41
Figure 15 : Classification des conformations obtenues par les calculs de docking pour le ligand GTP A) en présence et B) en absence de Mg selon l'énergie et le nombre de liaisons hydrogène	43
Figure 16 : visualisation des conformations pour le système GTP/Arf1 obtenue (A) en présence de Mg^{2+} et (B) en absence de Mg^{2+} . La protéine, colorée en gris, est représentée en mode cartoon, le ligand GTP en tiges et les résidus du site actif en ligne	44
Figure 17 : Classification des conformations obtenues par les calculs de docking pour le ligand GDP A) en présence B) en absence de Mg selon l'énergie et le nombre de liaisons hydrogène	46
Figure 18: visualisation des conformations pour le complexe GDP/Arf1 obtenue (A) en présence de Mg^{2+} et (B) en absence de Mg^{2+} . La protéine, colorée en gris, est représentée en mode cartoon, le ligand GTP en tiges et les résidus du site actif en ligne	48
Figure 19: les énergies de liaison en fonction des distance entre A) Mg^{2+} et GTP B) Mg^{2+} et GDP.....	51

LISTE DES TABLEAUX

Tableau 1 : paramètres des grilles utilisés dans notre travail36

Tableau 2 : les énergies de liaison, les acides aminés constituant le site actif et les liaisons hydrogène entre le ligand GTP et les résidus du site actif de l'Arf1 en présence et en absence de Mg^{2+} , les résidus conservés sont colorés en vert45

Tableau 3: les énergies de liaison, les acides aminés constituant le site actif et les liaisons hydrogène entre le ligand GDP et les résidus du site actif de l'Arf1 en présence et en absence de Mg^{2+} , les résidus conservés sont colorés en vert49

Résumé

L'ADP ribosylation factor1 (Arf1) est une petite protéine G localisée au niveau de l'appareil de Golgi où elle exerce plusieurs fonctions homéostatiques comme le transport vésiculaire. Contrairement aux autres membres de cette famille de petites protéines G, Arf1 est surexprimée dans les cellules invasives du cancer du sein et elle est impliquée dans la migration et la prolifération des cellules cancéreuses. Ainsi, Arf1 constitue une cible potentielle dans le traitement anti-cancer du sein.

L'objectif principal de notre travail est d'étudier le mode de liaison des deux ligands, GTP et GDP, avec les résidus du site de liaison de l'Arf1 en présence et en absence de l'ion Mg^{2+} . Cette étude est réalisée à l'aide d'une technique de la modélisation moléculaire basée sur la cible biologique, en l'occurrence le docking moléculaire.

Les résultats de la présente étude, qui sont en parfait accord avec les résultats de la littérature, montrent que l'ion Mg^{2+} favorise l'interaction des ligands GTP et GDP aux résidus tapissant le site actif de la protéine en particulier ASP93, ASP96 et THR31. Nous pouvons dire que cette étude constitue un outil potentiel dans la compréhension du mécanisme d'activation et d'inactivation de la protéine Arf1.

Mots clés : Arf1, GDP, GTP, docking moléculaire, anti-cancer .

Introduction générale

L'outil informatique a été mis en œuvre dans le domaine chimique dans le but de modéliser les structures les plus complexes de manière de plus en plus rigoureuse et réaliste. Tout chimiste peut aujourd'hui réaliser une étude plus ou moins exhaustive et précise des systèmes moléculaires grâce à la modélisation moléculaire

La modélisation moléculaire est une nouvelle technique de compréhension des phénomènes chimiques comme les mouvements dynamiques des molécules et la prédiction qualitative des propriétés des molécules. Elle regroupe aussi les techniques de visualisation, de manipulation, d'analyse et de calcul de la structure spatiale [1-2].

La modélisation moléculaire permet de prédire l'interaction entre le ligand et sa cible. Pour cela, des outils spécifiques ont été développés tel que le docking moléculaire ayant pour but d'obtenir un ensemble de conformations du ligand de plus basse énergies issu de son interaction avec la cible [3].

Par ce travail, nous souhaitons étudier le mode de liaison du GDP et GTP avec l'Arf1. GTP et GDP sont deux coenzymes de transfert de groupement phosphate, ce sont des donneurs de ce groupement. L'ADP ribosylation factor1 (Arf1) est une petite protéine G localisée au niveau de l'appareil de Golgi où elle exerce plusieurs fonctions homéostatiques comme le transport vésiculaire. Contrairement aux autres membres de cette famille de petites protéines G, Arf1 est surexprimée dans les cellules invasives du cancer du sein et elle est impliquée dans la migration et la prolifération des cellules cancéreuses. Ainsi, Arf1 constitue une cible potentielle dans le traitement anti-cancer du sein.

Notre travail a pour objectif d'étudier le mode de liaison de ces deux ligands, GTP et GDP, avec les résidus du site de liaison de l'Arf1 en présence et en absence de l'ion Mg^{2+} . Cette étude est réalisée à l'aide d'une technique de la modélisation moléculaire basée sur la cible biologique, en l'occurrence le docking moléculaire

Ce travail s'articule autour de trois chapitres avec une introduction et une conclusion générale.

Le premier chapitre résume les résultats de notre recherche bibliographique sur les petites protéines G notamment la protéine Arf1, sa structure cristallographique, son mécanisme d'activation ainsi que le rôle de Mg^{2+} dans ce mécanisme.

Dans le deuxième chapitre nous avons parlé des différentes méthodes de la modélisation moléculaire utilisées dans ce travail. Nous avons commencé par une définition des méthodes de la chimie théorique à savoir les méthodes de la mécanique quantique et celles de la mécanique moléculaire. Par la suite, nous avons parlé des méthodes de drug design en particulier le docking moléculaire que nous avons utilisé dans nos travaux de recherche présentés dans ce mémoire. La dernière partie de ce chapitre a été réservée à la présentation du logiciel que nous avons utilisé pour la réalisation de présent travail, en l'occurrence de logiciel Autodock [4] et son interface graphique AutodockTools [5].

Le troisième chapitre est réservé à l'étude par docking moléculaire du rôle de l'ion Mg^{2+} dans la stabilité des complexes GDP /Arf1 et GTP/Arf1 pour voir l'effet de cet ion sur la liaison entre ces deux ligands et la protéine Arf1.

On complète ce travail par une conclusion générale qui est une récapitulation des différentes approches empruntées dans ce présent travail de recherche avec un résumé de l'essentiel des résultats obtenus et quelques perspectives qui peuvent être une extension de cette étude.

Référence bibliographique:

- [1] Leach AR, Molecular Modeling: Principals and Applications, Edimbourg: Addison Wesley Longman Ltd, 1996, 1-24.
- [2] Graham LP, Chimie pharmaceutique, *de boeck*, 2003, 319-334.
- [3] Amandine SC, Ciblage thérapeutique du Complexe Récepteur de l'Elastine dans le cadre de l'invasion tumorale du mélanome : Aspects mécanistiques et impact de l'âge. Université de reims champagne –ardenne, 2015.
- [4] Morris GM, Goodsel DS, Halliday RS et *al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, 1998, 19: 1639-1662
- [5] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, and Olson AJ, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *Journal of Computational Chemistry*, 2009, 30(16): 2785–2791.

CHAPITRE I

RAPPELS BIBLIOGRAPHIQUES SUR LA PROTEINE ADP- RIBOSYLATION FACTOR 1 (ARF1)

I. Introduction :

Le cancer, appelé également tumeur maligne, est une maladie caractérisée par une prolifération cellulaire anormale dans un tissu normal de l'organisme [1].

Le cancer du sein est une tumeur maligne qui se développe au niveau de la glande mammaire. Ce type de cancer est le plus fréquent chez la femme et le nombre de cas est en constante augmentation. En Algérie, cette maladie représente plus de 50% des cancers gynécologiques chez les femmes [2].

Des études récentes ont montré que la petite protéine ADP ribosylation facteur 1 (arf1) est surexprimée dans les cellules hautement invasives du cancer de sein. Nous résumons dans ce chapitre l'état actuel des connaissances concernant ces protéines ainsi que leur cycle d'activation.

II. Définition des protéines G :

Les petites protéines G appartiennent à la famille des GTPase. Elles existent avec une forte proportion chez les eucaryotes. Ce sont des protéines ayant une masse moléculaire variant entre 20 et 30 kDa [3].

Les protéines G contrôlent divers aspects de la vie cellulaire tels que la transduction du signal, la réorganisation du cytosquelette et les mécanismes de transport vésiculaire [3]. Elles sont classées en deux superfamilles [4] :

- Les protéines G hétérotrimériques.
- Les protéines G monomériques.

II.1- Les protéines G hétérotrimériques :

Les protéines G hétérotrimériques se trouvent généralement liées à la membrane, elles sont constituées par l'assemblage de sous-unités alpha ($G\alpha$), beta ($G\beta$) et gamma ($G\gamma$). Chacune de ces composantes présente une diversité structurale : la sous-unité $G\alpha$ est la composante catalytique, tandis que les sous-unités $G\beta$ et $G\gamma$ sont des sous-unités régulatrices [4].

II.2- Les protéines G monomériques :

Les protéines G monomérique, appelées également les petites protéines G, sont composées d'une seule sous-unité plus petite (environ 20 kDa) non liées à la membrane. Cette famille de protéine se divise en deux branches [5] :

- ✓ Les protéines Monomériques de la famille *d'Elongation factor thermo unstable* (EF-TU)
- ✓ Les protéines G de la superfamille *Rat sarcoma* (Ras)

II.2.1- Les protéines G monomériques de la famille EF-TU :

Les protéines G monomériques de la famille *d'Elongation factor thermo unstable* (EF-TU), sont impliquées dans la synthèse protéique. Cette famille de protéines a été la première protéine G monomérique étudiée en détail [4].

II.2.2- La super famille des protéines Ras :

Depuis plusieurs années, de nombreuses petites protéines G de la superfamille *Rat sarcoma* (Ras) ont été identifiées et caractérisées. Ces protéines, de faibles poids (20kDa), sont impliquées dans plusieurs événements cellulaires tels que la migration, la prolifération, la différenciation ou le réarrangement du cytosquelette d'actine. Elles sont distribuées en cinq groupes : Ras, Rho, Rab, Ran et ARFs [6].

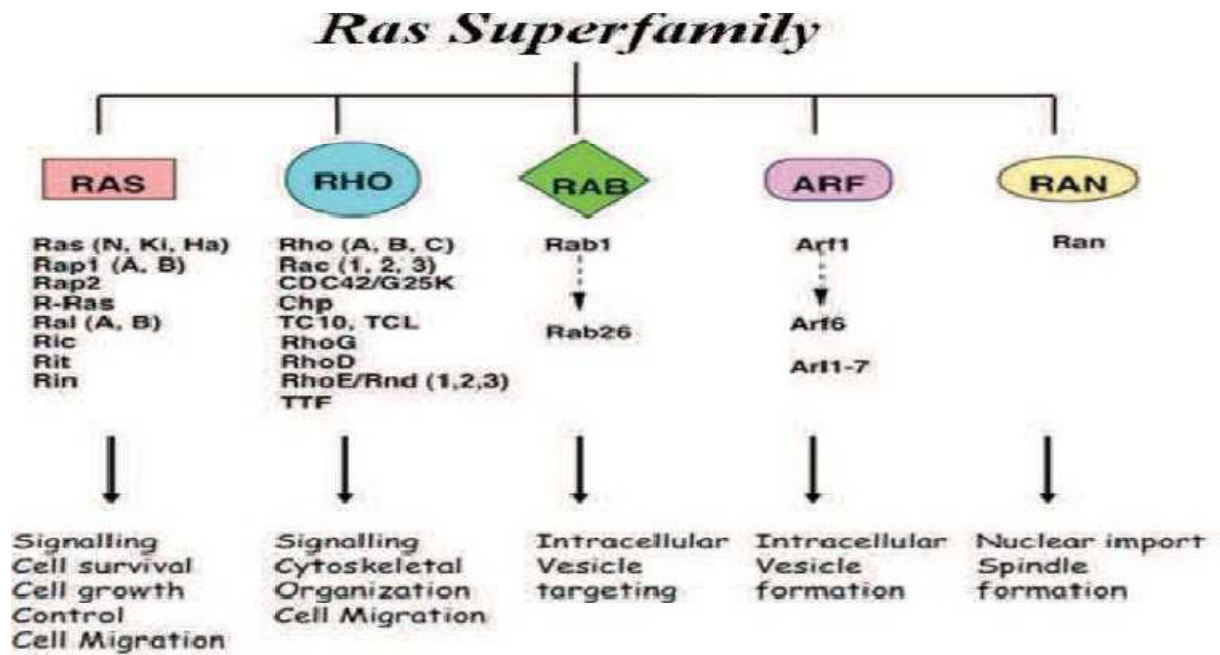


Figure 1: La superfamille Ras des protéines G

a- Protéine RAS :

La famille de protéines Ras (*Rat sarcoma*), est la plus étudiée de cette superfamille. Initialement, les protéines Ras ont été identifiées dans les années 1980. Des homologues cellulaires ont été par la suite identifiés. Ces protéines jouent un rôle majeur dans la transduction des signaux régulant la prolifération et la division cellulaire [4].

b- La famille des RHO:

Les petites protéines G monomériques de la famille des *Ras homologue* (Rho) regroupent au moins 22 protéines [7]. Chez l'homme, les membres de cette famille sont divisés en 6 classes: Rho, Rac, Cdc42, RhoBTB, Rnd et RhoT [8].

Les Rho GTPases jouent un rôle dans de nombreux processus cellulaires tels que la réorganisation du cytosquelette, la progression du cycle cellulaire, l'expression génique, la polarité cellulaire, la migration et la transformation cellulaire [9].

c- Les membres de la famille RAN :

Les protéines *Ras-like nuclear protein* (Ran) sont les protéines GTPases les plus abondantes dans les cellules. Leur fonction régule le transport nucléo cytoplasmique dans les cycles cellulaires des acides ribonucléiques (ARN) et des protéines [10].

d- La famille RAB :

Plus de 60 protéines *Ras-like protein in brin* (Rab) ont été identifiées dans les cellules de mammifères. Elles sont présentes sous de nombreuses formes et sont localisées dans des compartiments spécifiques selon leur fonction [11]. Les protéines Rab sont considérées comme des organisateurs du trafic intracellulaire assurant le criblage et le transport des vésicules [12].

e- La famille des ARF:

Les protéines de la famille *ADP ribosylation factor* (ARF) sont des protéines de 20 kDa. Elles comportent 30 membres chez l'homme [3]. La famille des protéines ARF est constituée par les sous-familles Arf, Arl, la protéine ARD1, la protéine Sar1 et la protéine ARP [5].

✓ La sous famille Arl:

Les protéines *Arf-like* (Arl) représentent une sous-famille de sept isoformes. Les protéines Arl possèdent des motifs caractéristiques fortement similaires en séquence. Ces protéines possèdent entre 63 et 67 % de similarité de séquence avec les protéines Arf [5].

✓ La protéine SARs :

La protéine *Secretion-Associated and Ras-superfamily-related gene* (SARs) a été découverte en 1989 dans la levure. Ces protéines présentent 54 % de similarité de

séquence avec les protéines Arfs. Elles possèdent la plus longue extension N-Terminale de la famille ARF (23 résidus) [5].

✓ **La protéine ARD1:**

La protéine Arf-Ribosylation Domain 1 (ARD1) découverte en 1986 est une protéine ubiquitaire de 64 kDa. Elle est constituée par 2 domaines : un domaine C-Terminal de 18 kDa, homologue aux protéines ARF (68 % de similarité de séquence avec Arf1) et un domaine N-Terminale de 46 kDa homologue aux ArfGAP [5].

✓ **La protéine ARP :**

La protéine ARF-Related Protein (ARP) est une protéine ubiquitaire, qui a été découverte en 1995. Elle représente la protéine la plus divergente de la famille ARF (56 % de similarité de séquence avec Arf1). L'ARP possède une extension N-Terminale de 17 résidus. [5]

✓ **La sous-famille Arf :**

Les protéines de la sous-famille Arf ont été les premières protéines de la famille ARF à être identifiées. Elles sont les plus étudiées et servent de modèle pour la famille des protéines ARF [5].

Les protéines Arfs sont présentes dans tous les organismes eucaryotes et sont très conservés. Leur identité de séquence est supérieure à 60%. Ils sont composés de six membres chez les mammifères. Elles jouent un rôle de régulation dans la formation de vésicules et le trafic cellulaire [13]. Ainsi que dans la dynamique du cytosquelette [14].

Les Arfs sont localisés au niveau des membranes dans toute la cellule, incluant la membrane plasmique et les membranes du système sécrétoire, endosomales et lysosomales [3]

Les protéines Arf sont constituées de six isoformes. Elles peuvent être classées, selon leur identité de séquence en acides aminés, en trois classes [3] :

La classe I : comprenant les Arf1, Arf2 et Arf3. Ces membres possèdent une identité de séquence de plus de 96%. Elles sont impliquées dans le contrôle de la formation de vésicules lors du transport intracellulaire [15].

La classe II : constituée des protéines Arf4 et Arf5. Elles ont une identité de séquence de 90% entre elles et de 80% avec les membres de la classe I [3]. La protéine arf4 est localisée au niveau de l'appareil de Golgi ou elle organise le trafic entre ces compartiments [16].

La classe III : cette classe est constituée uniquement de l'Arf6. Cette dernière possède une identité de séquence avec les autres classes qui varie de 64% à 69% [3]. Elle est localisée dans la membrane plasmique et dans les membranes endosomales [17]. Elle est impliquée dans la régulation du trafic membranaire et dans l'organisation du cytosquelette d'actine [18].

III. Les protéines Arf1 :

III.1- Définition :

La protéine Arf1 est l'isoforme la plus abondante de la famille Arf. Elle représente 90% des protéines Arfs présentes dans une cellule [5]. De ce fait et parce que, les autres protéines Arf ont été identifiées plus tardivement, Arf1 a été beaucoup plus caractérisée que ses isoformes. Elle est essentiellement localisée au niveau de l'appareil de Golgi [19].

III.2- Structure de l'Arf1 :

Les protéines Arf1 sont constituées de 181 acides aminés. Leur structure comporte une hélice alpha N-Terminale myristoylée qui leur permet de se fixer à la bicouche lipidique de la membrane plasmique. Les structures de ces protéines possèdent aussi deux régions principales appelées switch I et switch II reliées par une boucle *β -hairpin* appelée région interswitch.

La région Switch I est située entre les acides aminés 45 et 54. La région Switch II se situe entre les acides aminés 70 et 80 et l'interswitch se situe entre les acides aminés de 55 à 69 [20].

Dans la figure suivante, nous représentons la structure 3D et la séquence en acides aminés de la protéine arf1.

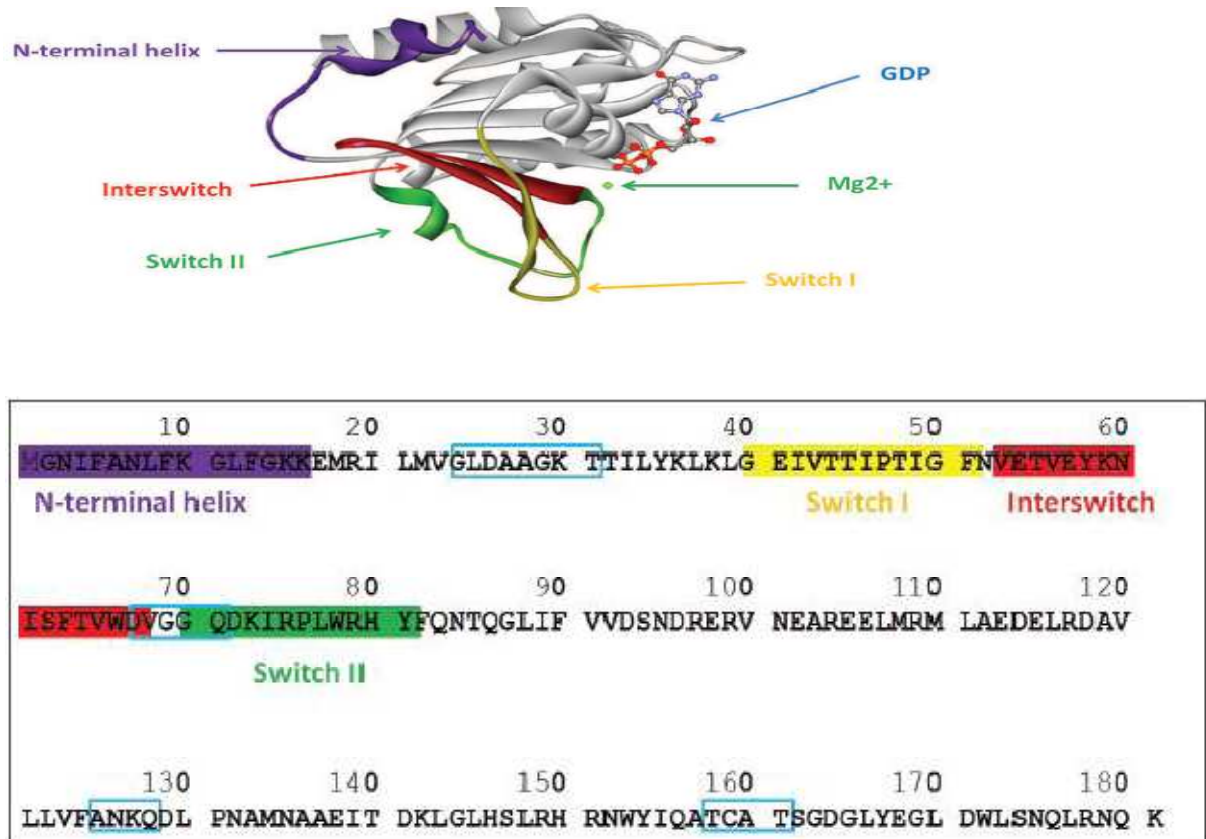


Figure 2 : Caractéristiques structurales d'arf1. (A) Structure 3D de l'arf1-GDP, (B) Séquence en acides aminés de l'arf1

La figure précédente montre que la structure de l'Arf1 est constituée de :

✓ **Une hélice α N-Terminale**

Représentée en violet et contient les acides aminés : **M** (Met : Méthionine), **G** (Gly : Glycine), **N** (Asn : Asparagine), **I** (Ile : Isoleucine), **F** (Phe : Phénylalanine), **A** (Ala : Alanine), **N** (Asn : Asparagine), **L** (Leu : Leucine), **F** (Phe : Phénylalanine), **K** (Lys :

Lysine), **G** (Gly : Glycine), **L** (Leu : Leucine), **F** (Phe : Phénylalanine), **G** (Gly : Glycine), **K** (Lys : Lysine), **K** (Lys : Lysine).

✓ **La partie Switch I**

Représenté en jaune, et regroupe les résidus : **G** (Gly : Glycine), **E** (Glu : Acide Glutamique), **I** (Ile : Isoleucine), **V** (Val : Valine), **T** (Thr : Thréonine), **T** (Thr : Thréonine), **I** (Ile : Isoleucine), **P** (Pro : Proline), **T** (Thr : Thréonine), **I** (Ile : Isoleucine), **G** (Gly : Glycine), **F** (Phe : Phénylalanine).

✓ **La partie interswitch**

Représenté en rouge et regroupe les acides aminés : **V** (Val : Valine), **E** (Glu : Acide Glutamique), **T** (Thr : Thréonine), **V** (Val : Valine), **E** (Glu : Acide Glutamique), **Y** (Tyr : Tyrosine), **K** (Lys : Lysine), **N** (Asn : Asparagine), **I** (Ile : Isoleucine), **S** (Ser : Serine), **F** (Phe : Phénylalanine), **T** (Thr : Thréonine), **V** (Val : Valine), **W** (Trp : Tryptophane), **D** (Asp : Acide Aspartique), **V** (Val : Valine).

✓ **La partie Switch II**

Représentée en vert, et contient les acides aminés : **G** (Gly : Glycine), **D** (Asp : Acide Aspartique), **K** (Lys : Lysine), **I** (Ile : Isoleucine), **R** (Arg : Arginine), **P** (Pro : Proline), **L** (Leu : Leucine), **W** (Trp : Tryptophane), **R** (Arg : Arginine), **H** (His : Histidine), **Y** (Tyr : Tyrosine), **F** (Phe : Phénylalanine).

✓ **Les résidus constituant le site actif de la protéine**

G (Gly : Glycine), **L** (Leu : Leucine), **D** (Asp : Acide Aspartique), **A** (Ala : Alanine), **A** (Ala : Alanine), **G** (Gly : Glycine), **K** (Lys : Lysine), **T** (Thr : Thréonine), **D** (Asp : Acide Aspartique), **V** (Val : Valine), **G** (Gly : Glycine), **G** (Gly : Glycine), **Q** (Gln : Glutamine), **A** (Ala : Alanine), **N** (Asn : Asparagine), **K** (Lys : Lysine), **Q** (Gln : Glutamine), **T** (Thr : Thréonine), **C** (Cys : Cystéine), **A** (Ala : Alanine), **T** (Thr : Thréonine), sont encadrés en bleu [21]

✓ **Les résidus constituant le site de liaison de Mg²⁺ (4A°)**

GDP1, GLU54, THR31

✓ **Les résidus constituant le site de liaison de GDP (3A°)**

ALA160, ASP129, THR32, THR31, GLY29, ALA28, ALA27, MG182

✓ **Les résidus constituant le site de liaison de GTP (3A°)**

ASP129, THR48, THR45, THR32, THR31, LYS30, GLY70, MG1002

IV. Mécanisme d'activation et d'inactivation d'ARF :

Les petites protéines G peuvent se trouver sous deux conformations actives et inactives. Elles sont dites actives si elles sont liées au guanosine triphosphate (GTP) ou inactives lorsqu'elles sont liées au guanosine diphosphate (GDP) [3]. Les structures chimiques de ces deux composés (GDP et GTP) sont représentées sur les **figures 4 et 5**. Elles passent d'une conformation à une autre suivant un cycle GDP/GTP comme c'est illustré dans la **figure 3**.

L'activation des Arf-GTPases résulte d'une interaction avec un facteur d'échange nucléotidique, (Guanine Exchange Factor ou GEF). Ce dernier stimule la dissociation du GDP pour le remplacer par du GTP. C'est sous cette forme liée au GTP, dite active, que les Arfs interagissent avec leurs effecteurs et les activeront.

Le passage de la forme active (protéine G liée au GTP) à la forme inactive (protéine G liée au GDP) permet le déclenchement ou l'inhibition des réactions biochimiques dans la cellule. Cette réaction est catalysée par une protéine nommée *GTPase activating protein* (GAP) assurant ainsi un retour vers la forme inactive liée au GDP [22].

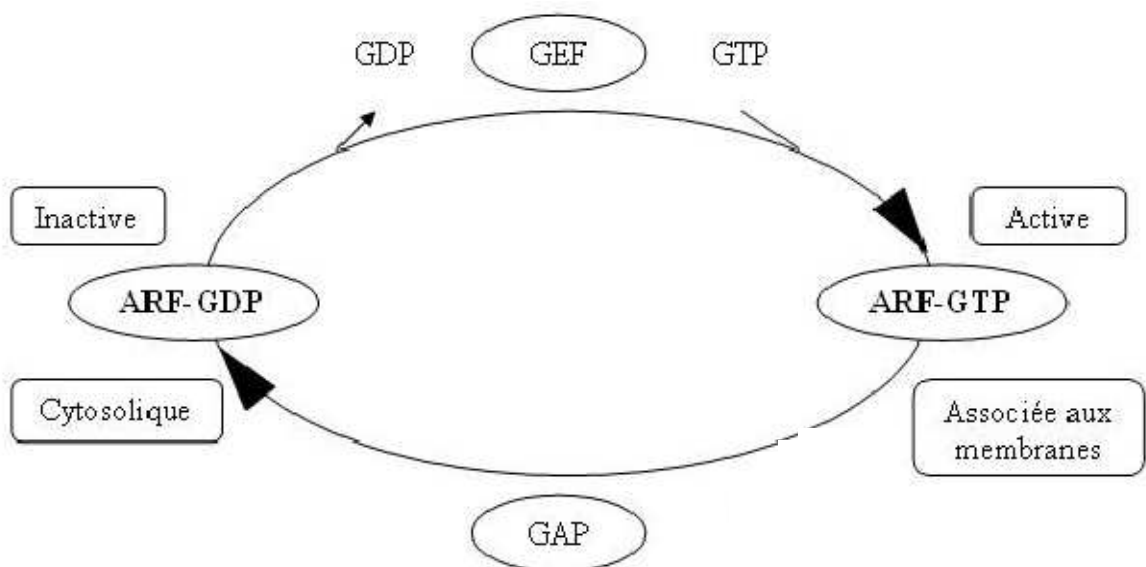


Figure 3 : Mécanisme d'activation de l'ARF

IV.1- Le guanosine triphosphate (GTP) :

C'est une coenzyme de transfert de groupement phosphate, c'est un donneur de phosphate.

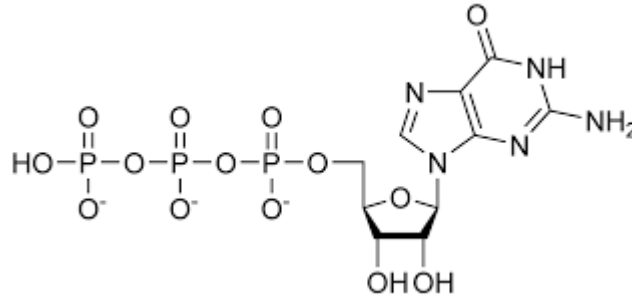


Figure 4 : Structure de GTP

IV.2- Le guanosine diphosphate (GDP) :

C'est une coenzyme de transfert de groupement de phosphate. Elle résulte de l'hydrolyse de la GTP, le groupement phosphate libéré peut être transférée sur une protéine.

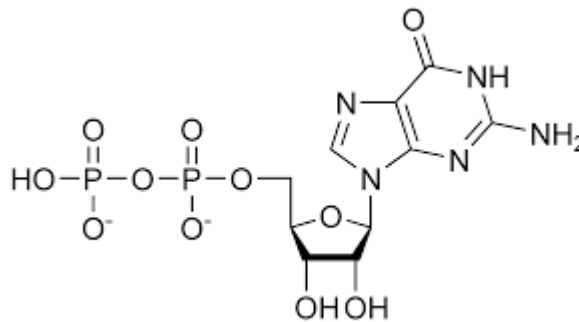


Figure 5 : Structure de GDP

IV.3- Les ARF-GEFs :

Toutes les structures, correspondant aux protéines Arf liées à la guanine nucleotide exchange factor (GEFs), identifiées possèdent un domaine Sec7, un module d'environ 200 acides aminés qui est suffisant pour catalyser l'échange du GDP par GTP [23]. Ces

protéines régulatrices ont pour fonction biochimique de stimuler la dissociation du nucléotide GDP (forme inactive) favorisant l'association du nucléotide GTP à l'Arf et la formation de la conformation active.

L'activation de la protéine Arf par ses GEFs peut être inhibée par stabilisation de complexes abortifs Arf-GDP-GEF en présence de la toxine bréfeldine A [22].

IV.4- Les ARF-GAPs :

Au moins 20 gènes codant pour les protéines Arf liées à la protéine activatrice de GTPase (ARF-GAPs) ont été identifiés chez l'homme. Toutes les ARF-GAPs contiennent un domaine conservé d'environ 70 acides aminés contenant un motif en doigts de zinc qui est responsable de l'activité catalytique [23].

IV.3- Le cycle GDP/GTP de l'Arf1:

Le cycle GDP/GTP de Arf1 est particulier par rapport aux autres petites protéines G. Cependant, l'étape d'activation nécessite qu'Arf1 s'associe à la membrane, alors que l'hydrolyse du GTP (étape d'inactivation) conduit à la dissociation d'Arf1 de la membrane [5]

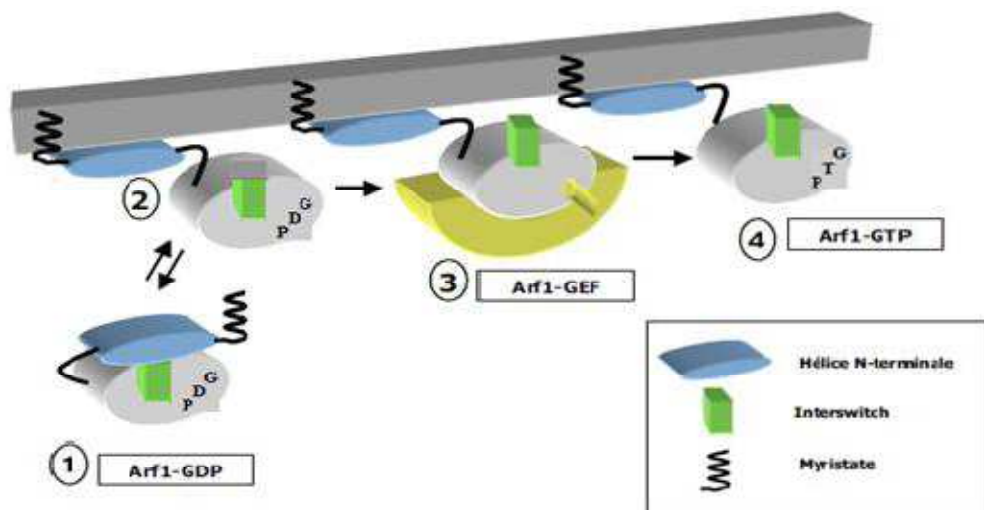


Figure 6 : Mécanisme d'activation d'arf1 par sa GEF en présence de la membrane

La forme inactive d'Arf1 liée au GDP passe d'une conformation où l'hélice N-Terminale est plaquée contre la protéine (conformation 1) et une forme fixée à la membrane (conformation 2) [5]. Le myristate renforce globalement l'interaction de la protéine avec la membrane, qu'elle soit sous forme GDP ou GTP [24]. L'Arf1-GDP fixée à la membrane s'associe avec sa GEF (conformation 3). Dans cette étape l'interswitch se place en position sortie, empêchant l'hélice N-Terminale fixée à la membrane de se plaquer à nouveau contre la protéine. La forme active liée au GTP reste fixée à la membrane avec l'interswitch sortie (conformation 4) [5].

V. Rôle de Mg^{2+} dans le mécanisme d'Arf1 :

Mg^{2+} a été établi comme un cofacteur essentiel pour les fonctions de la protéine liant le GTP. Pour de nombreuses petites protéines G, Mg^{2+} a été démontré nécessaire pour les liaisons entre les bases nucléiques (ou base azotée) et la guanine (Voir la figure 7) et l'hydrolyse de GTP (étape d'inactivation) [25].

En présence de Mg^{2+} , les petites protéines G présentent une affinité de liaison extrêmement élevée aux guanines nucléotides (comme les facteurs d'échange de nucléotides de guanine (GEFs)). L'activité GTPase intrinsèque de Ras devient indétectable lorsque le cofacteur Mg^{2+} est éliminé, ce qui indique que l'ion bivalent Mg^{2+} est absolument nécessaire pour l'hydrolyse du GTP dans les réactions de GTPase respectives [26,27].

Dans les cycles d'activation et d'inactivation des petites protéines G, Ras et ARF, les protéines GEFs favorisent la dissociation du GDP par déstabilisation de la liaison Mg^{2+} avec ces GTPases. [28,29]

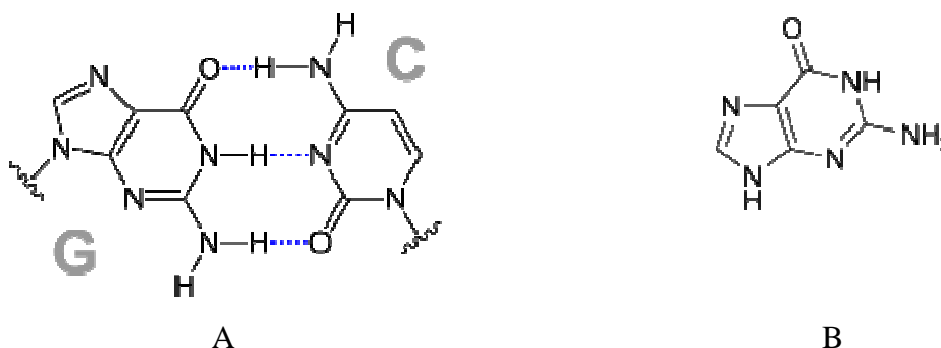


Figure 7 : A) Liaison de la guanine (G) avec une base nucléique (la cytosine (C)) ; B) la structure de la guanine (G)

Références bibliographiques

- [1] <http://sante-medecine>.
- [2] Gouadfel K, Badis Kh. contribution à l'évaluation des nouveaux facteurs pronostiques du cancer du sein et étude rétrospective sur une durée de 3 années. Biochimie appliquée. Université Mouloud Mammeri Tizi-Ouzou, 2013.
- [3] Kaheima A, Etude structurale et biochimique d'un facteur d'échange atypique d'Arf. Sciences agricoles. Thèse de doctorat. Université Paris Sud - Paris XI, 2012.
- [4] SCHONN JS, Mécanismes de la sécrétion régulée des hormones et des neurotransmetteurs, rôle des GTPase Rab3 et Rab27. Thèse de doctorat : Université Pierre & Marie Curie (Paris VI), 2003, 242.
- [5] Menetrey J, Etude structurale des petites protéines G : Rap2A dans un complexe non catalytique avec le GTP et Arf6 en complexe avec du GDP, Biochimie, Université Pierre et Marie Curie - Paris VI, 2000.
- [6] Boulay PL, L'étude du rôle d'ARF1 dans la migration et la prolifération des cellules du cancer du sein, Université de Montréal, 2011.
- [7] Wennerberg K, and Der CJ, Rho-family GTPases: it's not only Rac and Rho, Journal of cell science, 2004,117, 1301-1312
- [8] Bustelo XR, Sauzeau V, and Berenjano IM, GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. BioEssays: news and reviews in molecular, cellular and developmental biology, 2007, 29, 356-370
- [9] Lewis-Saravalli S, ARF1 contrôle la migration des cellules hautement invasives du cancer du sein via Rac1, Université de Montréal, 2012.
- [10] Weis K, Cell, 2003, 112: 441-451.
- [11] Zerial M, McBride H, Nat Rev Mol Cell Biol, 2001, 2(2):107-17.
- [12] Grosshans BL, Ortiz D, and Novick P, Proc Natl Acad Sci U S A, 2006, 103: 11821-7.
- [13] Donaldson JG, and Jackson CL, ARF family G proteins and their regulators: roles in membrane transport, development and disease, Nat Rev Mol Cell Biol, 2011, 12(6): 362-375.
- [14] Myers KR, and Casanova JE, Regulation of actin cytoskeleton dynamics by Arf family GTPases, Trends Cell Biol, 2008, 18(4): 184-192.
- [15] Kahn RA, and Gilman AG, J Biol Chem, 1984, 259(10)
- [16] Ben-Tekaya H, Kahn RA, and Hauri HP, Mol Biol Cell, 2010, (23): 4130-4140.

- [17] Donaldson JG, Radhakrishna H, Peters P, Cold Spring Harb Symp Quant Biol, 1995, 60: 229-234
- [18] Donaldson JG, J BiolChem, 2003, 278: 41573-41576.
- [19] Whitney JA, et al., Cell, 1995, 83(5): 703-13.
- [20] Pasqualato S, Renault L, Cherfils J, EMBO Rep, 2002, 3(11): 1035-1041.
- [21] Daher Z, Rôle des GTPases ARF dans la migration des cellules endothéliales et la sécrétion du NO, Université de Montréal, 2009.
- [22] Cherfils J et Pacaud P, M/S : médecine sciences, 2004, 20 : 393-396.
- [23] Nie Z, Hirsch DS, Randazzo PA, Curr Opin Cell Biol, 2003, 15(4): 396-404.
- [24] Vincent M, Hélices amphipathiques et transport vésiculaire, Biochimie, Université Nice Sophia Antipolis, 2009.
- [25] Baolin Z, Yaqin Z, Zhi-xin W, Yi Zheng, J. Biol. Chem, 2000, 25299–25307.
- [26] John J, Rensland H, Schlichting I, Vetter I, Borasio GD, Goody RS, and Wittinghofe AJ, *Biol. Chem.*, 1993 , 268: 923–929
- [27] Burstein ES, and Macara IG, *Biochem J*, 1992, 282: 387–392
- [28] Boriack-Sjodin PA, Margarit SM, Bar-Sagi D, and Kuriyan J, *Nature*, 1998, 394, 337–343
- [29] Beraud-Dufour S, Robineau S, Chardin P, Paris S, Chabre M, Cherfils J, and Antonny B, *EMBO J*, 1998, 17 : 3651–3659

CHAPITRE II

LES PRINCIPALES APPROCHES DE LA MODÉLISATION MOLÉCULAIRE

I. Introduction

La modélisation moléculaire a pour but de prévoir la structure et la réactivité des molécules ou de systèmes de molécules et de manipuler les modèles des structures en deux ou trois dimensions.

Les méthodes de la modélisation moléculaire comprennent : les méthodes de la chimie théorique et celles du drug design.

II. Les méthodes de la chimie théorique

II.1- Les méthodes de la chimie quantiques

II.1.1- Equation de Schrödinger

La chimie quantique définit la structure moléculaire comme un noyau autour duquel gravitent des électrons. Ces derniers sont décrits par leur probabilité de présence en un point et représentés par des orbitales [1]. Les équations de la chimie quantique sont basées sur la résolution de l'équation de SCHRÖDINGER qui s'écrit pour les états stationnaires [2] :

$$H\Psi = E\Psi \quad (1)$$

Où :

Ψ : Fonction d'onde de la molécule.

H : Hamiltonien total d'une molécule comportant N noyaux et n électrons, son expression est donnée par la formule (2).

E : Energie totale d'une molécule.

La résolution exacte de cette équation n'est possible que pour l'atome d'hydrogène et les systèmes mono-électriques. Pour les systèmes polyélectriques, on fait appel aux méthodes d'approximation, pour la résolution approchée de l'équation (1).

$$\hat{H} = \sum_{j=1}^N -\frac{\hbar^2}{2M_j} \Delta_j + \sum_{i=1}^n -\frac{\hbar^2}{2m_e} \Delta_i + \sum_{j=1}^N \sum_{j'=1, j' > j}^N Z_j Z_{j'} \frac{e^2}{r_{jj'}} - \sum_{j=1}^N \sum_{i=1}^n Z_j \frac{e^2}{r_{ij}} + \sum_{i=1}^n \sum_{i'=1, i' > i}^n \frac{e^2}{r_{i'i'}} \quad (2)$$

I

II

III

IV

V

Où I : est l'énergie cinétique des N noyaux atomiques (indexés j)

II : est l'énergie cinétique des n électrons (indexés i)

III : est l'énergie de répulsion entre les noyaux

IV : est l'énergie d'attraction noyau-électron

V : est l'énergie de répulsion entre les électrons

Δ_i et Δ_j sont les laplaciens de l'électron i et du noyau j . Δ_i est donné par la relation :

$$\Delta_i = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \quad (3)$$

m_e : est la masse de l'électron

M_j , Z_j et Z_j' : sont respectivement la masse du noyau j , les numéros atomiques des atomes j et j' .

$r_{ii'}$: est la distance entre l'électrons i et i' .

$r_{jj'}$: est la distance entre le noyau j et j' .

r_{ij} : la distance entre un électron i et un noyau j .

e : la charge de l'électron.

Avec h : $\hbar = \frac{h}{2\pi}$ la constante de Planck.

L'équation (1) ne peut être résolue analytiquement que pour des petits systèmes, d'où la nécessité d'introduire des approximations et de résoudre le système numériquement [3]

II.1.2- Approximation de Born-Oppenheimer (BO)

L'approximation de Born-Oppenheimer est utilisée pour simplifier l'équation de Schrödinger. Elle se base sur le fait que les noyaux possèdent une masse beaucoup plus importante que celle des électrons. Ainsi, leur mouvement est considéré comme étant très lent par rapport à celui des électrons [5]. Cette hypothèse implique que l'énergie cinétique des noyaux est négligée et le terme d'interaction entre les noyaux est constant. [5]

Dans le cadre de cette approximation, l'hamiltonien moléculaire peut se diviser en une partie nucléaire H_N et une partie électronique H_e . La partie électronique est donnée par la relation :

$$\hat{H}_e = \sum_{i=1}^n h^c(i) + \sum_{i=1}^n \sum_{i' > 1}^n \frac{e^2}{r_{ii'}} \quad (4)$$

Avec :

$$h^c(i) = \sum_{i=1}^n -\frac{\hbar^2}{2m_e} \Delta_i + \sum_{j=1}^N Z_j \frac{e^2}{r_{ij}} \quad (5)$$

$h^c(i)$ est un hamiltonien de cœur relatif à l'électron i , c'est un operateur mono-électronique qui décrit le mouvement d'un des n électrons dans le champ des N noyaux. Le dernier terme de l'équation (5) empêche la séparation de l'équation de Schrödinger en n équations mono électroniques d'ou la nécessité d'introduire une autre approximation et qui concerne l'expression de la fonction d'onde.

II.1.3- Approximation orbitale

La fonction d'onde électronique doit tenir compte du spin de l'électron, donc il est nécessaire d'ajouter une quatrième coordonnée pour définir l'état du moment de spin [6].

$$\chi(i) = \varphi(i)\alpha(i) \quad (6)$$

$$\bar{\chi}(i) = \varphi(i)\beta(i) \quad (7)$$

Les deux fonctions de spin d'un électron peuvent prendre des valeurs $+1/2$ et $-1/2$ selon l'orientation de l'électron.

Pour un système ayant n électrons, la fonction d'onde est :

$$\Psi(1,2,\dots,n) = \chi_1(1)\chi_2(2)\chi_3(3)\dots\chi_n(n) \quad (7)$$

Cette fonction d'onde ne prend pas en compte l'indiscernabilité des électrons car elle n'obéit pas au principe d'exclusion de Pauli [7]

Pour prendre en compte cette indiscernabilité et assurer l'antisymétrie, la fonction d'onde est écrite sous la forme d'un déterminant de Slater [6] :

$$\Psi (1, 2 \dots n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) \dots & \chi_n(1) \\ \chi_1(2) & \chi_2(2) \dots & \chi_n(2) \\ \vdots & \vdots & \vdots \\ \chi_1(n) & \chi_2(n) \dots & \chi_n(n) \end{vmatrix} \quad (8)$$

$\frac{1}{\sqrt{n!}}$: est le facteur de normalisation, n étant le nombre d'électrons.

II.1.3- Approximation LCAO

L'approximation LCAO (*Linear Combination of Atomic Orbitals*) est basée sur le fait que la liaison chimique se forme suite au recouvrement de fonctions d'onde appartenant à deux atomes suffisamment rapprochés [10].

II.1.4- Les différentes méthodes de la chimie quantiques

a- Les méthodes ab-initio

Le terme latin ab initio signifie –depuis le début. Il désigne les calculs qui dérivent des principes théoriques (calculs de mécanique quantique approchés) sans prendre en considération les données expérimentales. Les calculs ab initio au niveau Hartree-Fock, et surtout post Hartree-Fock, sont très demandeurs en temps de calcul. Ce temps augmente rapidement avec le nombre d'électrons, ou plus exactement avec le nombre de fonctions de base. Pour remédier à ce problème les méthodes semi-empiriques ont été développées [9,10].

Les méthodes semi-empiriques (SE) [11-12] sont des techniques de résolution de l'équation de Schrödinger de systèmes à plusieurs électrons. Contrairement aux méthodes *ab initio*, les méthodes semi-empiriques négligent une grande partie des intégrales en les corrélant à des données expérimentales [13]. Ces méthodes ne considèrent que les électrons de valence qui sont décrits par les OA de type Slater. Cette approximation est basée sur le fait que les électrons impliqués dans la liaison chimique sont ceux de la couche de valence. Ces méthodes cherchent d'une part, à diminuer considérablement le nombre d'intégrales moléculaires à calculer et d'autre part, à éliminer complètement le calcul des intégrales biélectroniques, à trois et quatre centres. Actuellement, il existe un grand nombre de méthodes SE. Les méthodes semi empiriques les plus couramment utilisées sont MNDO (Modified

Neglect of Diatomic Differential Overlap) [Dewar en 1977.], AM1 (Austin Model 1) [Dewar en 1985] et PM3 (*Parametric Method 3*) [14].

b- Les méthodes de la théorie fonctionnelle de la densité (DFT)

Ces méthodes sont basées sur le théorème de Hohenberg Kohn [15] qui dit que l'énergie d'un système dans son état fondamental est une fonctionnelle de la densité électronique $\rho(r)$ ainsi que toute densité $\rho'(r)$, autre que la densité réelle, conduit nécessairement à une énergie supérieure.

Contrairement aux méthodes avec une fonction d'onde complexe à $3N$ -dimensions, la DFT cherche une simple fonction à trois dimensions : la densité électronique totale ρ [16,17]. Ainsi, l'énergie d'une molécule peut être déterminée de la densité électronique au lieu d'une fonction d'onde.

II.2- Les méthodes de la mécanique moléculaire

II.2.1- Définition de la mécanique moléculaire (MM)

La Mécanique Moléculaire est basée sur l'approximation de Born-Oppenheimer selon laquelle les noyaux sont beaucoup plus lents que les électrons et ces noyaux sont donc implicitement traités.

La MM est une méthode empirique où les noyaux sont représentés par des sphères, et les liaisons par des ressorts. La distance d'équilibre entre deux atomes liés est calculée soit par des méthodes quantiques, soit déterminée par des méthodes expérimentales (rayons X, RMN...). L'avantage principal de la MM par rapport aux autres méthodes est la rapidité des calculs et aussi la possibilité de traiter des systèmes moléculaires très grands grâce à une simplification des équations du champ de forces.

II.2.2- Champ de force

Un champ de force est constitué de plusieurs fonctions d'énergie potentielle qui dérivent les interactions intramoléculaires entre atomes liés et non liés. L'énergie potentielle est une somme de termes décrivant ces interactions :

$$E = \underbrace{\sum E_{\text{liaisons}} + \sum E_{\text{angles}} + \sum E_{\text{dièdres}} + \sum E_{\text{dièdres impropres}}}_{E_{\text{covalente}}} + \underbrace{\sum E_{VDW} + \sum E_{\text{électrostatique}}}_{E_{\text{non covalente}}}$$

Où :

E_{liaison} : est l'énergie de liaison covalente entre deux atomes.

E_{angle} : est l'énergie associée aux angles formés par deux liaisons consécutives.

$E_{\text{dièdres}}, E_{\text{dièdres impropres}}$: est l'énergie associée aux angles dièdre et dièdres impropres.

E_{VDW} : est l'énergie des interactions non-covalentes de Vander Waals.

$E_{\text{électrostatique}}$: est l'énergie des interactions électrostatiques.

Plusieurs champs de forces ont donc été développés afin d'étudier les différentes classes de molécules organiques et de biomolécules comme :

- MM2 est le premier champ de force développé par Allinger et *col.* [18-19-20]. Il a été conçu au début pour les molécules simples (alcane, alcène, alcyne non conjugués, amine...), mais ses versions améliorées MM3 (1989) [21] et MM4 (1996) [22] lui permettent de traiter des molécules organiques de plus en plus complexes
- MMFF (Merck Molecular Force Field) [23]
- CFF (Consistent *Force* Field) [24]
- CVFF (Consistent Valence Force Field) [25].
- SPASIBA (Spectroscopic Potentiel Algorithm for Simulating Biomolecular conformational Adaptability): élaboré par Gérard Vergoten et al. (1995)

On trouve aussi ceux qui sont utilisés dans le traitement de molécules biologiques, telles que les acides nucléiques et les protéines comme :

- GROMOS (GRONigen MOlecular Simulation) [26]
- AMBER (Assisted Model Building and Energy Refinement) [27,28]
- CHARMM (Chemistry at HARvard Macromolecular Mechanics) [29]
- OPLS : (Optimised potentials for Liquid Simulation): développé par W. L Jorgensen et J. Tirado Rives [30]

II.2.3- Les différentes méthodes de la mécanique moléculaire

a- La méthode de steepest descent

La méthode steepest descent (plus grande pente) se base sur le calcul de la première dérivée de la fonction énergétique par rapport aux coordonnées géométriques. C'est une méthode qui converge difficilement vers le minimum d'énergie, mais elle converge très rapidement lorsque la conformation initiale de la structure est éloignée du minimum énergétique. Pour cette raison, elle est souvent utilisée en première étape dans un processus de minimisation, dans le but d'accélérer la convergence du système [31].

b- La méthode du gradient conjugué

Cette méthode, fondée sur le même principe que la précédente " *steepest descent* ". L'intérêt de cet algorithme est d'éviter un comportement oscillatoire autour du minimum et d'accélérer la convergence. Lorsque la conformation initiale est loin du minimum la convergence devient difficile [32].

c- Méthode de Newton-Raphson

La méthode Newton-Raphson utilise la dérivée seconde des gradients pour localiser les minima d'énergie. C'est une méthode qui converge très bien au voisinage du minimum d'énergie. C'est pourquoi il est préférable d'utiliser au préalable la méthode SD pour s'approcher dans un premier temps d'un minimum d'énergie.

d- Recuit Simulé

La méthode de recuit simulé, développée par Kirkpatrick (1983), autorise la fonction d'énergie à augmenter momentanément afin de franchir des barrières d'énergie pour retomber dans un minimum plus profond.

Le recuit simulé permet, en plus d'explorer l'espace conformationnel d'une molécule et d'étudier ses propriétés dynamiques. Il est bien adapté aux molécules ayant un grand nombre de degrés de liberté.

e- Monté Carlo (MC)

La méthode de Monte Carlo (MC) fut développée par Von Neumann, Ulam et Metropolis, à la fin de la seconde guerre mondiale, pour l'étude de la diffusion des neutrons dans un matériau fissile. [33]

La probabilité d'une configuration particulière d'énergie potentielle P_i est alors proportionnelle à :

$$P_i = e^{\frac{-E_i}{k_B T}}$$

Avec :

E_i : L'énergie d'un état i

k_B : La constante de Boltzmann

T : La température

III- Les méthodes de drug design

Le *drug design* est une approche qui vise à rechercher des molécules thérapeutiques par conception *in silico* et en se basant sur des informations expérimentales. L'approche est basée soit sur le ligand, dite « ligand-based », soit sur le récepteur, dite « receptor-based ou structure-based » [34]. La première utilise des outils tels que le QSAR (*Quantitative Structure-Activity Relationships*) et ne nécessite pas la connaissance de la cible biologique. Quant à la deuxième, elle est appliquée lorsqu'une structure tridimensionnelle (3D) fiable du récepteur est disponible. Dans nos travaux, nous avons utilisé l'approche receptor-based que nous présentons ci-dessous.

III.1- Le QSAR

Le principe des méthodes QSAR (*Quantitative Structure Activity Relationship*) consiste à mettre en place une relation mathématique, reliant des propriétés moléculaires microscopiques appelées descripteurs, à une grandeur expérimentale (activité biologique, toxicité, affinité pour un récepteur), pour une série de composés chimiques similaires.

III.2- Le docking

C'est une approche basée sur la cible biologique, appelée également l'arrimage moléculaire. Cette technique permet de placer une petite molécule, dite ligand, dans le site

actif d'une macromolécule, dite récepteur. L'objectif du docking moléculaire consiste à prédire à la fois la conformation active (ou viable) et l'orientation relative d'une molécule au sein du récepteur. Il faut noter que dans une simulation de docking, le récepteur et le ligand peuvent être considérés rigides ou flexibles. Dans le deuxième cas, le ligand est considéré flexible mais la flexibilité permise au récepteur est limitée exclusivement aux chaînes latérales des résidus du site actif ; on parle alors d'un docking flexible, réservé au docking ligand-protéine.

Dans nos travaux rapportés dans ce mémoire, nous avons utilisé la suite de programmes de docking flexible ; Autodock4 [35].

IV- Le logiciel Autodock

IV.1- Présentation

Autodock est un des cinq logiciels de docking les plus cités dans la littérature. C'est une suite d'exécutables capable de docker des ligands de nature organique dans une macromolécule. Dans la plupart des cas, le ligand correspond à une petite molécule et la macromolécule correspond à une protéine. Ce programme réalise cette simulation en utilisant des grilles de potentiels, calculées par AutoGrid. La version utilisée dans ce travail, permet de traiter la protéine partiellement flexible. Cependant, les chaînes latérales des résidus sélectionnés dans la région étudiée (généralement le site actif) sont considérées flexibles. Nous allons détailler les principes et le fonctionnement de la version 4 de ce logiciel : AutoDock 4.

IV.2- AutoDockTools (ADT)

AutoDockTools (ADT) est l'interface graphique du logiciel AutoDock (figure 8). Il a été développé par la même équipe qui a créé le programme AutoDock. ADT permet de visualiser la localisation précise de l'espace à l'intérieur duquel vont se faire les calculs et de préparer les fichiers de données nécessaires pour AutoGrid et AutoDock ainsi que de visualiser et d'analyser les résultats de docking.

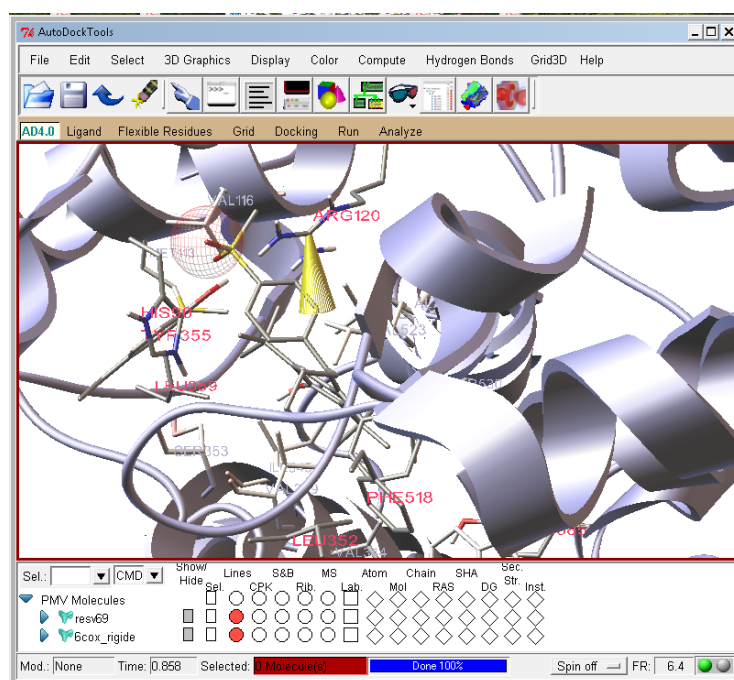


Figure 8 : Interface graphique d'AutoDock

IV.3- Principe d'utilisation d'AutoDock

IV.3.1. Préparation du récepteur

Pour permettre la flexibilité à certaines parties de la protéine, les résidus du site actif désirés être flexibles doivent être sélectionnés. La protéine ainsi préparée est enregistrée dans deux fichiers distincts : un regroupant les résidus flexibles du site actif et l'autre, la partie rigide de la protéine. Les deux fichiers doivent être au format *pdbqt* qui reprend le format *pdb* (*Protein Data Bank*) en y ajoutant une colonne pour les charges.

IV.3.2- Description de la macromolécule sous la forme de grilles

Une grille de point est une boîte définie à la périphérie de la macromolécule dans laquelle des conformations de docking pour un ligand donné sont recherchées. C'est une matrice tridimensionnelle qui englobe l'intégralité ou une région intéressante du récepteur étudié (**Figure 9**). Son intérêt est de faciliter le positionnement du ligand dans la macromolécule et d'augmenter la vitesse d'évaluation de l'énergie d'interaction ligand/récepteur. Pour chaque atome du ligand, AutoGrid calcule l'énergie d'interaction entre un atome sonde placé en un point de la matrice et l'ensemble de la macromolécule. L'espacement entre les points peut aller de 0,2 Å à 1,0 Å. Une grille supplémentaire est

calculée pour le potentiel d'interaction électrostatique avec pour une charge sonde (ponctuelle) +1 eV (Figure 2). La grille doit englober le site d'intérêt mais le temps de calcul pour l'explorer sera fonction de ses dimensions. Le volume explorable est donc limité par le temps et la puissance de moyens de calcul dont on dispose.

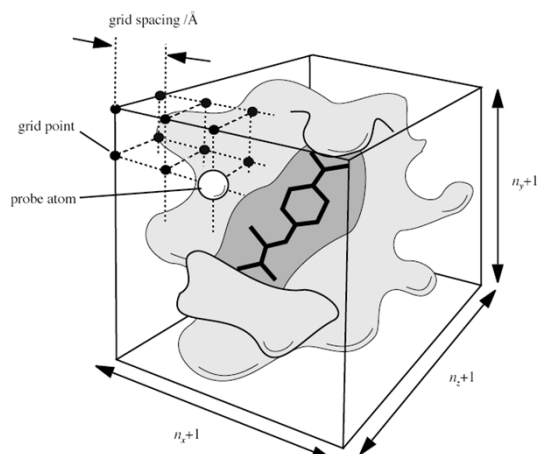


Figure 9: Schéma d'une grille de potentiel englobant le site actif d'une macromolécule.

Les grilles de potentiels sont calculées par le programme AutoGrid selon les instructions du fichier de données qu'on lui fournit. Ce dernier est un fichier texte dont l'extension est *gpf* (*Grid Parameters File*), il doit contenir les paramètres suivants :

- les noms des fichiers du ligand et de la macromolécule, aux formats *pdbqt*.
- la position et les dimensions de la boîte ainsi que l'espacement entre les points de la grille.

IV.3.3- Préparation des ligands

La préparation des ligands pour le docking moléculaire se fait avec l'interface graphique ADT. Après l'addition des hydrogènes, il est nécessaire d'attribuer des charges partielles. Les charges disponibles dans AutoDock sont celles de Gasteiger et de Kollman. La flexibilité du ligand est simulée par la rotation des groupements autour des liaisons non rigides. Cela revient à sélectionner les degrés de liberté qui correspondent aux libres rotations des groupements autour des liaisons simples. Le fichier de coordonnées atomiques du ligand doit être enregistré au format *pdbqt*. AutoDock fait ensuite varier les différents degrés de liberté du ligand (axes de translation, de rotation et angles dièdres) à l'intérieur de la grille.

IV.3.4- Recherche des conformations de docking

AutoDock permet d'explorer l'espace conformationnel avec plusieurs méthodes de recherche conformationnelle parmi lesquelles on peut citer : algorithme génétique darwinien (AG), un algorithme génétique lamarckien (AGL) et la technique Monte-Carlo appliquée avec recuit simulé (RSMC).

Le programme AutoDock fait une recherche de solutions de docking en fonction des paramètres qu'on lui transmet par l'intermédiaire du fichier *dpf* (*Dock Parameters File*). Les paramètres d'un calcul d'AutoDock sont :

- les noms des fichiers contenant les ligands et les grilles de potentiels à utiliser
- les états initiaux des ligands (position, orientation et conformation aléatoire ou précisée)
- la méthode de recherche à utiliser (RSMC, AG ou AGL) avec les paramètres correspondants.

IV.3.5- Paramètres d'un algorithme génétique

Les paramètres de l'algorithme génétique sont :

- Le nombre de cycles à exécuter (*Number GA Runs*) c'est le nombre de solutions souhaitées.
- La taille de la population (*Population Size*). La valeur standard est de 150 mais peut être modifiée en fonction du nombre de degrés de liberté et de l'étendue de la boîte à explorer.
- Le nombre maximum d'évaluations d'énergie (*Maximum Number of evals*). C'est un paramètre d'arrêt du docking utile lorsque le processus met trop de temps, ou ne parvient pas, à converger.
- Le nombre maximum de générations qui est un autre paramètre d'arrêt (*Maximum Number of generations*). Ces deux derniers paramètres sont à adapter conjointement en fonction de la taille de la population, du nombre de degrés de liberté, de l'étendue de la boîte à explorer et du degré de convergence à atteindre.
- Le niveau de sélection élitiste (*Maximum Number of top individu that automatically survive*) : le nombre des meilleurs individus qui survivent automatiquement à la génération suivante. C'est un paramètre défini par l'utilisateur, il indique combien des

meilleurs individus survivent automatiquement à la génération suivante. Par défaut, cette valeur est 1 : le meilleur uniquement.

- Le taux de mutations (*Rate of gene Mutation*). Cet opérateur joue un rôle d'optimisateur local pour les AGs purs (darwinien) en opérant des variations de faible amplitude sur les gènes, chose que le croisement fait très difficilement. La valeur standard est 0,02.
- Le taux de croisements (*Rate of Crossover*). Le croisement est l'opérateur d'exploration globale, sa probabilité doit être élevée pour que l'AG ait ce caractère global. La valeur standard est 0,8.
- Les valeurs assimilables à la moyenne et la variance de la distribution de Cauchy pour la mutation des gènes (*Mean of Cauchy and Variance of Cauchy distributions for gene mutation*). Les valeurs standard sont 0 et 1 respectivement.
- Le nombre de générations précédentes N dans lesquelles on recherche l'individu le plus mal adapté pour calculer le facteur de sélection proportionnelle. La valeur optimale est de 10 générations.

A la fin de chaque cycle d'exploration, AutoDock enregistre la meilleure solution de docking du cycle. Le fichier de résultats contiendra donc autant de solutions que de cycles exécutés. Ce fichier textuel porte l'extension *dlg*.

IV.3.6- Analyse et évaluation des résultats

Les grilles de potentiel préalablement calculées par AutoGrid sont utilisées pour évaluer l'énergie d'un conformère du ligand à une position donnée dans la macromolécule et celle du docking du complexe (Figure 10). Le caractère désigné en rouge sur la figure 3 donne accès à l'affichage des paramètres quantitatifs concernant le modèle de docking considéré.

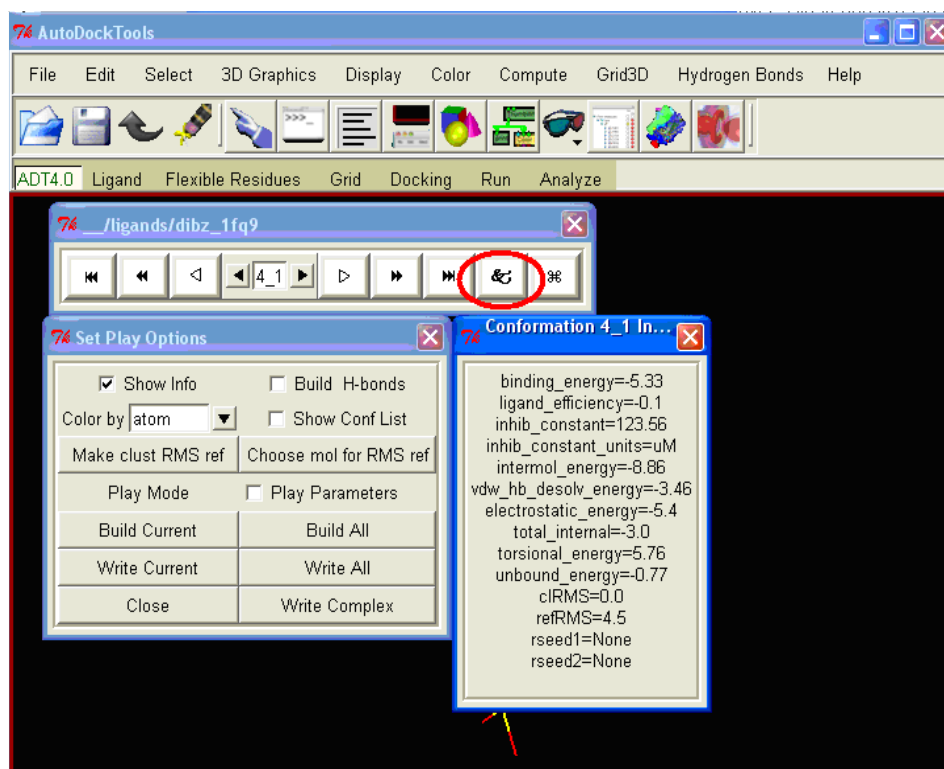


Figure 10 : Les différents outils d'analyse de résultats de docking disponibles dans ADT.

Référence bibliographique :

- [1] Cours de post-graduation; option; chimie théorique: assure par Pr. S. Mekelleche. 2000-2001
- [2] Gautier MOROY, thèse de doctorat, Université de Reims Champagne-Ardenne, France, 2005.
- [3] Burrau O., Klg. Danske Videnskab. Selskab. 1927. 7, 1.
- [4] Schrödinger E., Ann. Phys. Leipzig, 1926, 79, 361, 489, 734.
- [5] Born M., Oppenheimer JR., Ann. Physik. 1927, 84: 457.
- [6] Slater JC., Phys. Rev., 1929, 34: 1293.
- [7] Pauli W., Physik Z., 1925, 31: 765.
- [8] Stewart, I. P. Review in Computational Chemistry. In , Vol. 1 ; K.B. Lipkowitz and D.B. Boyd ed., VCH: New York, 1990; Chapter 2, pages 45-81.
- [9] Zerner, M. C. Review in Computational Chemistry. In; K.B. Lipkowitz and D.B. Boyd ed., VCH : New York, 1991; Chapter 8, pages 313-365.
- [10] (a) Stewart JJP, Review in Computational Chemistry, Vol. 1. (b) Lipkowitz KB and Boyd DB, VCH : New York, 1990, Chapter 2, 45–81.
- [11] Lipkowitz KB. and Boyd DB. ed., VCH: New York, 1991; Chapter 8, 313–365.
- [12] Parr RJ, J Chem Phys, 1952, 20: 239.
- [13] Stewart JJP., J. Comput. Chem., 1989, 10: 209-220.
- [14] Hohenberg, P. Kohn, W. Phys. Rev. 1964, 136, B864-b871.
- [15] Amant ST, A. In Review in Computational Chemistry. In , Vol. 7; K.B. Lipkowitz and D.B. Boyd ed., VCH : New York, 1996.
- [16] Kohn, W. J. Phys. Chem. 1996, 100, 12974.
- [17] Allinger N.L, Adv. Phys. Org. Chem, 1976, 13, pp.1.
- [18] Allinger N.L, J. Am. Chem. Soc, 1977, 99, 8127.
- [19] Burkert U, N.L. Allinger, Molecular Mechanics, ACS Monograph 177. American Chemical Society, Washington DC, 1982.
- [20] Allinger N.L, Yuh Y.H, J- H. Li, J. Am. Chem. Soc, 1989, 111, 8522.
- [21] Allinger N.L, Chen K, J- H. Lii, J. Comp. Chem, 1996, 17, 642.
- [22] Čížek J., Adv. Chem. Phys. 1969, 14: 35.
- [23] Lifson S., Warshel A., J Chem Phys, 1968, 49: 5116–5129.
- [24] Dauber-Osguthorpe P., Roberts VA., Osguthorpe DJ., Wolff J., Genest M., Hagler AT., Proteins: Structure, Function and Genetics, 1988, 4: 31-47.
- [25] Shimanouchi T, Pure Appl. Chem, 1963, 7, 131.

- [26] Hermans J., Berendsen HJC., van Gunsteren WF., Postma JPM., Biopolymers 1984, 23:1513 -1518.
- [27] Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz Jr. KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JC, Kollman PA, J Am ChemSoc 1995, 117:5179-5197.
- [28] Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner PKJ, J Am ChemSoc 1984, 106:765-784.
- [29] Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M, J Comp Chem 1983, 4:187-217.
- [30] Jorgensen.W.L and Rives.J.T, J. Am. Chem. Soc, 1988, 110, 1657.
- [31] Vergoten G., Cours de Modelisation moleculaire, professeur à l'Université des Sciences et Technologies de Lille.
- [32] Dudeck MJ., Scheraga HA., J. Comput.Chem, 1990, 11: 121.
- [33] Allalen M., « Molecular Dynamics and Monte-Carlo Simulations of CoPtalloys », Mémoire de Mgister, Université Tizi Ouzou, Algerie, 2002.
- [34] PatrickG. L, Chimie Pharmaceutique. Bruxelles, de boeck, 2003
- [35] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK and Olson AJ, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. Journal of Computational Chemistry, 1998, 19: 1639-1662

CHAPITRE III

ETUDE PAR DOCKING MOLECULAIRE DU RÔLE DE L'ION Mg^{2+} DANS LA STABILITE DES COMPLEXES GDP/Arf1 ET GTP/Arf1

I- Introduction

Dans ce chapitre, nous nous proposons d'étudier le rôle de l'ion Mg^{2+} sur la liaison des deux ligands GTP et GDP à la protéine Arf1. Il est bien de rappeler que cet ion a été établi comme cofacteur essentiel pour les fonctions de la protéine liant le GTP [1]. Cette étude est réalisée à l'aide d'une technique de la modélisation moléculaire basée sur la cible biologique, en l'occurrence le docking moléculaire. Notre objectif est d'étudier le mode de liaison de ces deux ligands, GTP et GDP, avec les résidus du site de liaison de l'Arf1 en présence et en absence de l'ion Mg^{2+} .

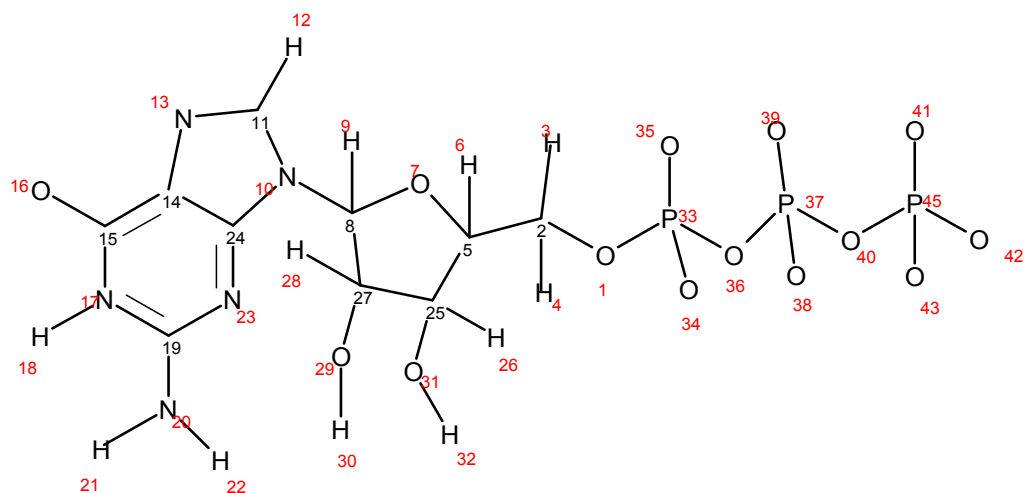
II- Méthodologie

II.1- Structures de départ

II.1.1- Préparation des structures de la protéine

Les coordonnées de la protéine Arf1 sont obtenues à partir d'une structure cristallographique disponible dans la banque de données des protéines (PDB : Protein Data Bank) [2]. Nous avons choisi les deux structures codées 1HUR et 1O3Y, de résolution égale à 2.0 Å [3] et 1.5 Å [4], respectivement. Ces deux structures correspondent aux complexes guanosine triphosphate (GTP)/Arf1 et guanosine diphosphate (GDP)/Arf1. Les structures chimiques des ligands GDP et GTP sont données dans la Figure 11.

A



B

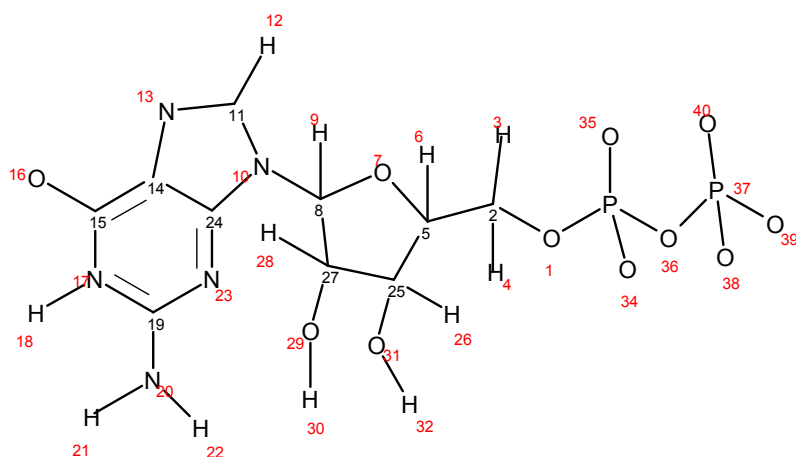


Figure 11 : (A) Structure de GTP et (B) Structure de GDP.

II.1.2- Préparation des ligands

Les structures chimiques des deux ligands étudiés, GDP et GTP, sont obtenues à partir des structures cristallographiques codées respectivement 1HUR et 1O3Y. Cette manipulation est réalisée à l'aide du logiciel AutodockTools [5]. Une vérification complète de ces structures a été effectuée avant de les soumettre aux calculs de docking. Le fichier contenant les coordonnées atomiques de chaque ligand est enregistré en format pdbqt.

II.2. Méthodologie utilisée

A l'aide de l'interface graphique AutodockTools (ADT), nous avons préparé et déterminé les paramètres optima que nous allons utiliser dans le cadre de ce travail

II.2. 1- Préparation des structures du récepteur

Nous avons sélectionné une seule chaîne, dépourvue des molécules d'eau et du ligand, pour la préparation de la structure de la cible des deux structures 1HUR et 1O3Y. Une fois le ligand retiré, les hydrogènes manquants sont ajoutés et les charges de Gasteiger sont assignées [6]. Les acides aminés entourant les deux ligands sont considérés dans notre étude comme des résidus englobant le site actif. Sachant que le programme Autodock4 [7] utilise un docking semi-flexible, nous avons séparé chacun des deux systèmes étudiés en une partie flexible, contenant les résidus du site actif, et une partie rigide, contenant le reste de la protéine. Deux fichiers de la protéine contenant chacun une partie flexible ou rigide sont enregistrés en format pdbqt pour chaque système. Dans le but d'étudier l'effet de l'ion Mg²⁺ sur les interactions entre les ligands et la protéine, nous avons effectué deux calculs pour chacun des deux systèmes étudiés, l'un en présence de Mg²⁺ et l'autre en absence de cet ion.

II.2.2- Localisation du site actif

Avant de retirer le ligand de chaque complexe, nous avons pris le soin de relever tous les résidus qui sont à une distance inférieure à 3 Å, du centre du ligand. Les résidus sélectionnés sont :

Structure	Résidus						
1HUR	ALA160	ALA28	ALA27	ASP129	THR32	THR31	GLY29
1O3Y	ASP129	GLY70	THR31	THR32	THR45	THR48	LYS30

Nous avons considéré ces acides aminés, englobant le site actif, flexibles dans nos calculs de docking.

II.2. 3- Démarche suivie dans les calculs de docking

a- Paramètres de calcul du grid

Le site actif ou la zone à explorer du récepteur, pour chaque système, est englobée dans une grille tridimensionnelle dont les paramètres sont définis de manière à permettre la libre rotation et translation des deux ligands étudiés. Ainsi, chaque ligand est docké dans cet espace délimitant le site de liaison dans l'Arf1. Nous présentons les paramètres utilisés dans cette étape dans le tableau suivant :

Tableau 1 : paramètres des grilles utilisés dans notre travail

Structure	Centre de la grille Suivant (x, y, z)	Volume de la grille	L'espacement entre les points
1HUR	58.417, 21.616, 9.655	28×42×46	0.375
1O3Y	-0.201, -6.730, 17.932	50×50×24	0.375

Il est bien de noter que l'espacement entre les points de la grille correspond au quart de la longueur d'une simple liaison C-C.

b- Paramètres de calculs de docking

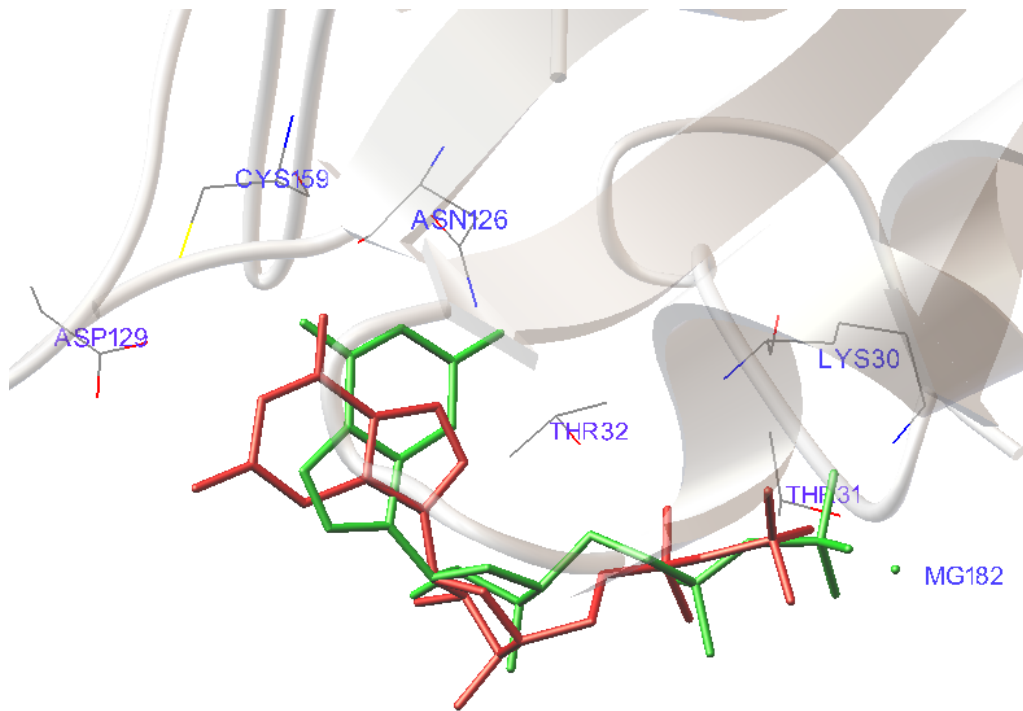
Le protocole de docking moléculaire, adopté dans ce travail, consiste à docker chaque ligand, considéré flexible, dans le site actif de l'Arf1. Les paramètres d'arrêt ont été repoussés à 2500.000 évaluations d'énergie et à 27000 générations. Dans le but d'avoir des résultats satisfaisant, nous avons retenu 50 conformations pour chaque ligand. Chaque calcul de docking a consisté en une série de 50 cycles, utilisant l'algorithme génétique lamarckien (LGA). Par ailleurs, nous avons pris les autres paramètres de simulations par défaut dans le logiciel Autodock.

III- Résultats et discussion

III.1- Validation des paramètres de calcul utilisés

Nous avons effectué cette validation en comparant les conformations des deux ligands étudiés, GDP et GTP, issues du docking et les conformations expérimentales issues des complexes 1HUR et 1O3Y. Nous présentons, dans la figure 12, la superposition des deux conformations RX issues du docking pour chaque système. Cette figure montre que les deux conformations se superposent bien; ce qui conforte l'utilisation des paramètres établis dans la suite de nos calculs de docking.

A



B

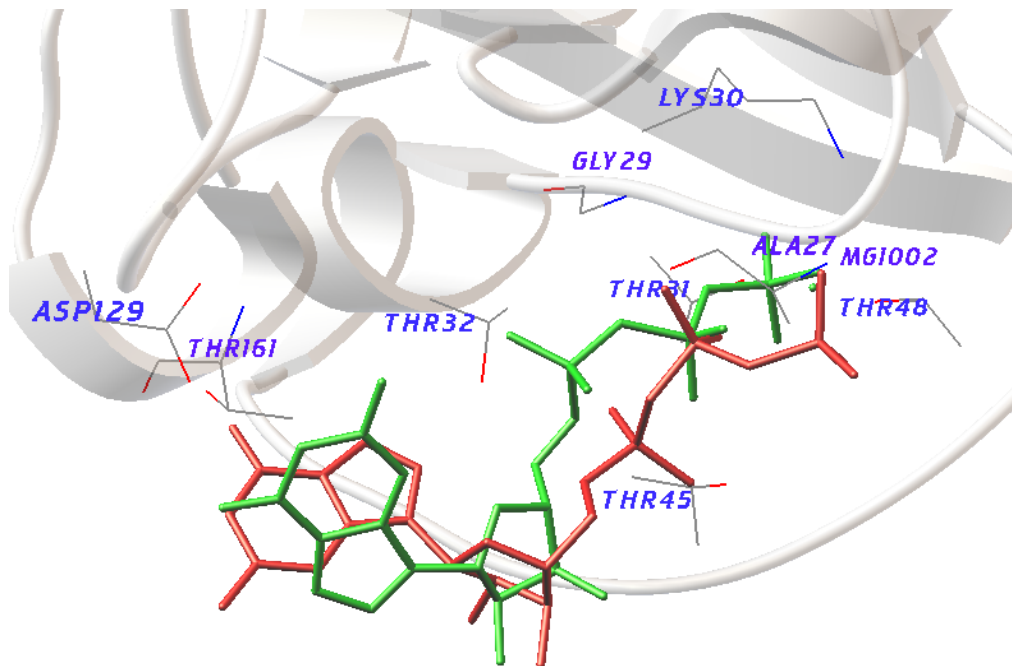
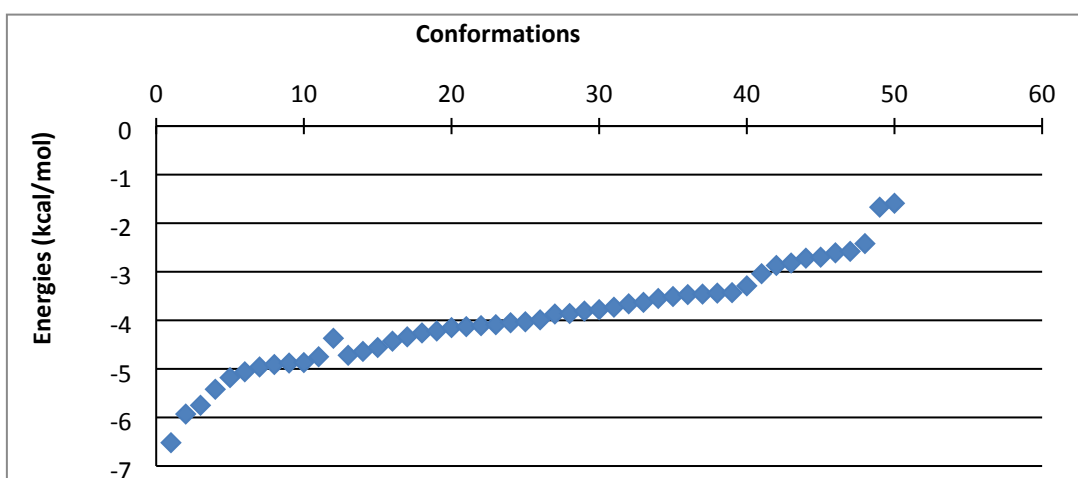


Figure 12 : superposition des deux conformations de A) GDP et B) GTP dans le site actif de l'Arf1 : les conformations issues du docking sont colorées en vert et les conformations cristallographiques en rouge.

III.2- Etude énergétique des conformations obtenue

A l'aide du logiciel MGLTools, nous avons réalisé une analyse détaillée des 50 conformations générées par nos calculs pour chaque système. Dans un premier temps, nous avons classé ces conformations par ordre d'énergie de liaison ligand / récepteur croissante. Cette étape nous a permis de repérer la conformation la plus stable pour chaque système. Les deux graphes que nous représentons dans la figure (3) montrent la classification des conformations obtenues par les calculs de docking pour le ligand GTP en absence et en présence de Mg^{2+} selon l'énergie croissante.

A



B

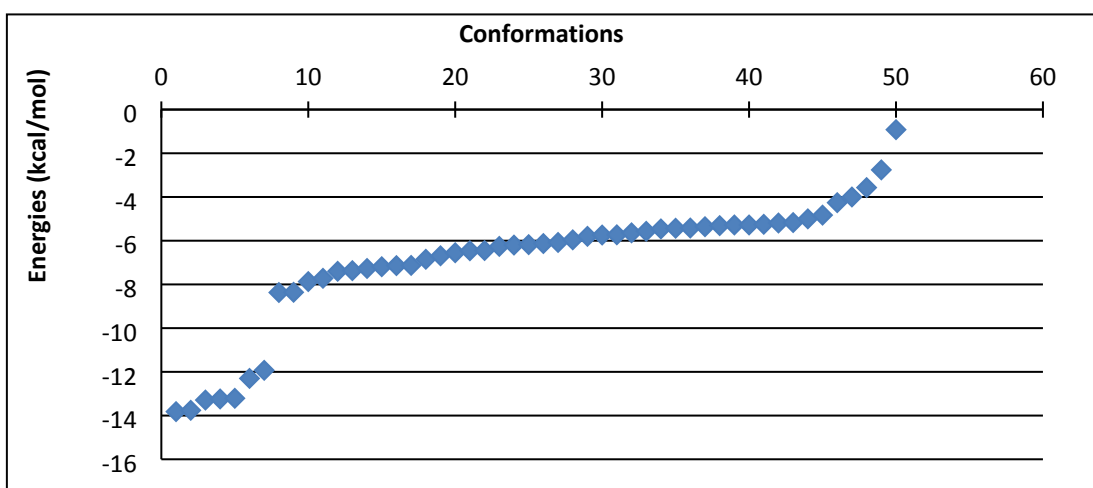
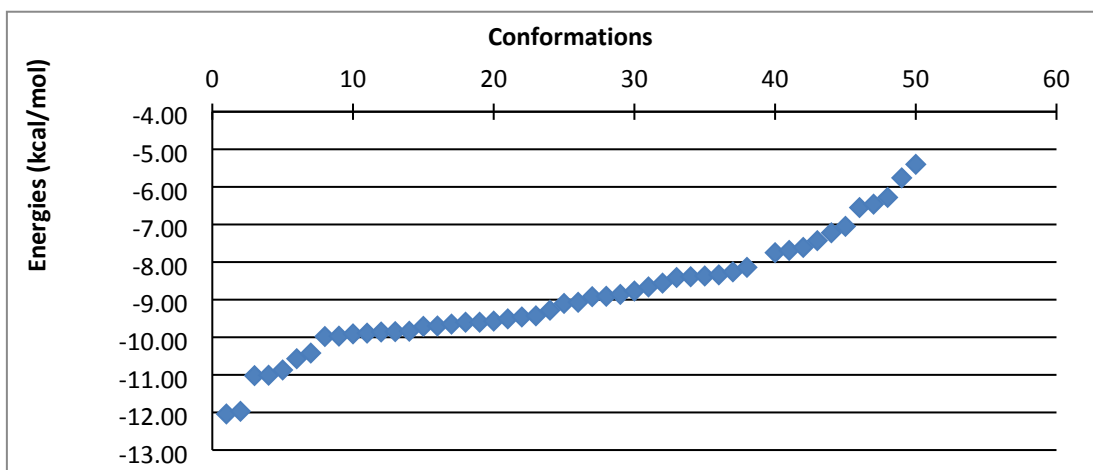


Figure 13 : Classification des conformations obtenues par les calculs de docking pour le ligand GTP A) en absence et B) en présence de Mg^{2+} selon l'énergie croissante

Les deux graphes précédents montrent que le complexe GTP/Arf1 le plus stable en absence de Mg^{2+} a une énergie de - 6,52 kcal/mol tandis qu'en présence de Mg^{2+} l'énergie du complexe le plus stable est de -13,82 kcal/mol. Ce résultat est en parfait accord avec les résultats des travaux de l'équipe de (Baolin Zhang) [1]. Rappelons que cette équipe a démontré que l'ion Mg^{2+} est nécessaires pour les liaisons entre les bases nucléiques (ou base azotée) et la guanine et l'hydrolyse de GTP (étape d'inactivation) [1].

Nous représentons dans la figure (4) la classification des conformations obtenues par les calculs de docking pour le ligand GDP en absence et en présence de Mg^{2+} selon l'énergie croissante.

A



B

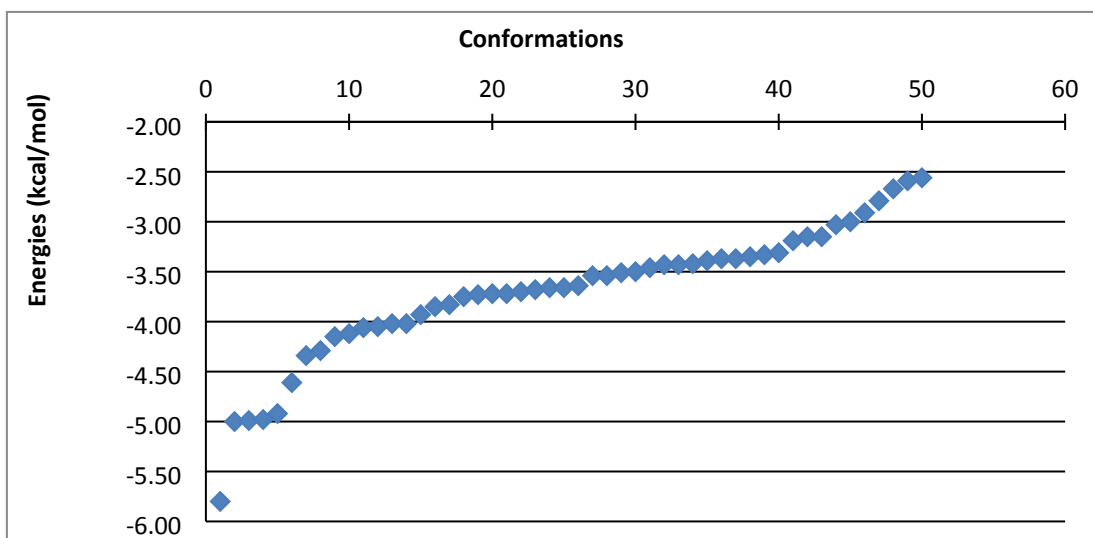


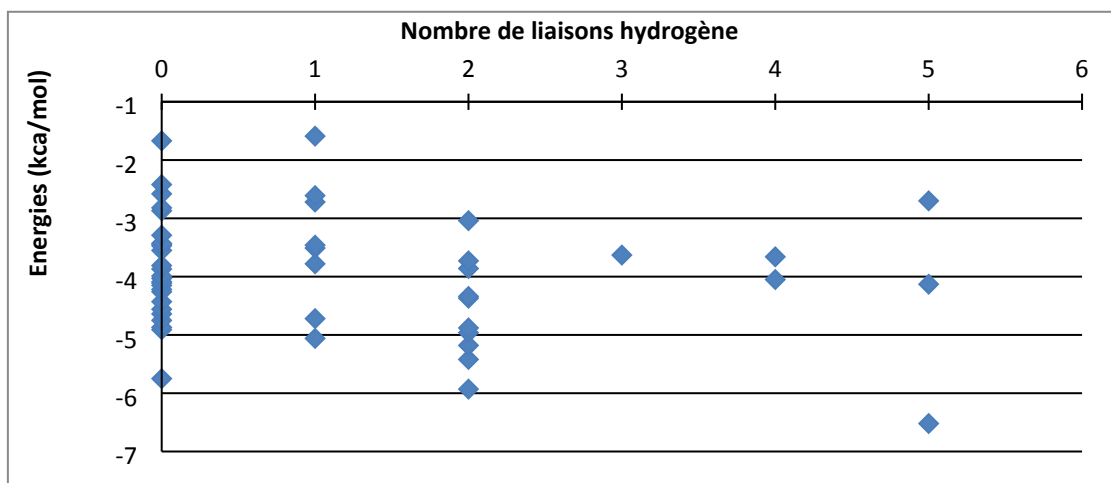
Figure 14 : Classification des conformations obtenues par les calculs de docking pour le ligand GDP A) en présence et B) en absence de Mg^{2+} selon l'énergie de liaison croissante

Ces deux graphes montrent que le complexe GDP/Arf1 le plus stable en présence de Mg^{2+} a une énergie de -12.04 kcal/mol tandis qu'en absence de Mg^{2+} l'énergie du complexe le plus stable est de -5,8 kcal/mol. Cette variation dans l'énergie de liaison entre les deux systèmes montre le rôle important de l'ion Mg^{2+} dans les liaisons entre GDP et la protéine Arf1.

III.3- Etude des liaisons Hydrogène

L'analyse des différentes conformations obtenues par nos calculs, montre que les ligands étudiés, GDP et GTP, forment des interactions électrostatiques de type liaison hydrogène avec les résidus du site actif de la protéine. La figure suivante montre le nombre de liaison hydrogène entre le ligand GTP et les résidus du site actif de la protéine en présence et en absence de Mg^{2+} .

A



B

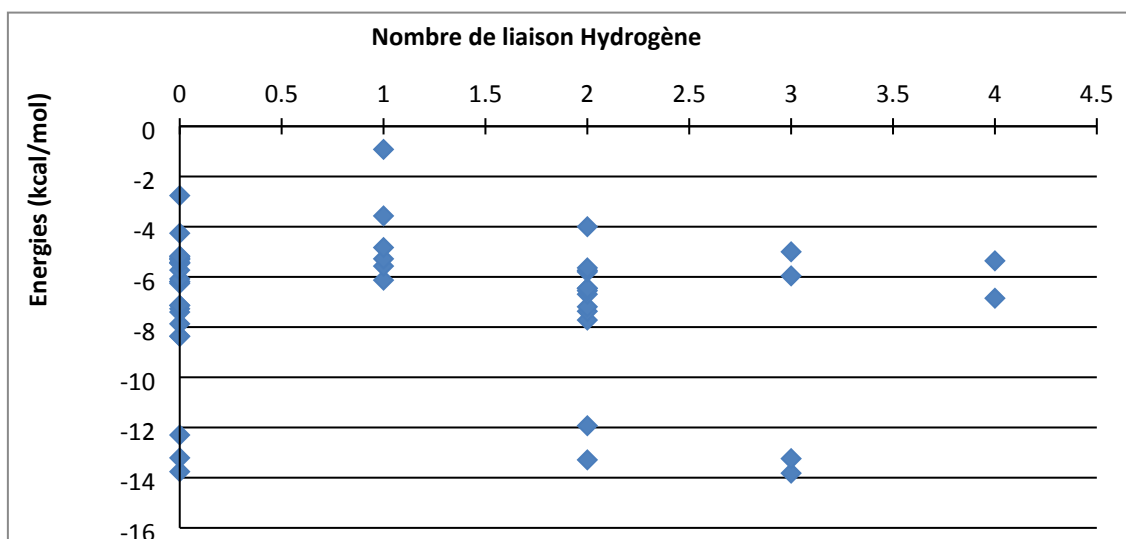
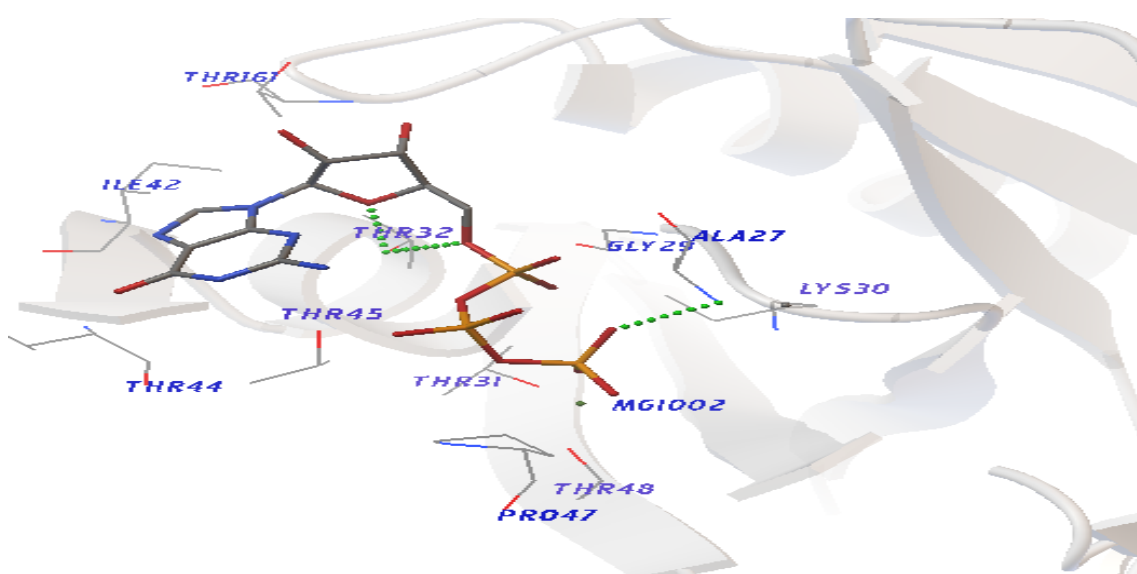


Figure 15 : Classification des conformations obtenues par les calculs de docking pour le ligand GTP A) en présence et B) en absence de Mg^{2+} selon l'énergie et le nombre de liaison hydrogène

L'analyse de ces deux graphes montre que les conformations les plus stables obtenues en présence de Mg^{2+} ($E = -13,82$ kcal/mol) et en absence de Mg^{2+} ($E = -6,52$ kcal/mol) présentent des interactions électrostatiques de type liaisons hydrogène avec les résidus du site actif de la protéine. Nous présentons, dans la figure 6 les positionnements des conformations les plus stable obtenues pour le ligand GTP dans le site de liaison de la protéine Arf1. Cette figure montre qu'en présence de Mg^{2+} , GTP forme trois liaisons hydrogène avec les résidus du site actif. Deux avec l'atome d'oxygène de la chaîne latérale du résidu THR32 ($d = 2.532$ Å et $d = 2.610$ Å) et une avec l'atome d'azote du backbone du résidu ALA27 ($d = 2.874$ Å).

En absence de Mg^{2+} , GTP forme cinq interactions de type liaison hydrogène. Deux de ces interactions sont formées entre ce ligand et les atomes d'azote de la chaîne latérale des résidus ASN126 ($d = 3.057 \text{ \AA}$) et LYS30 ($d = 2.940 \text{ \AA}$). Deux autres interactions sont formées entre GTP et les atomes d'oxygène des chaînes latérales des résidus THR31 ($d = 2.816 \text{ \AA}$) et THR45 ($d = 3.034 \text{ \AA}$). La cinquième interaction est formée entre ce ligand et l'atome d'azote du backbone du résidu GLY29 ($d = 2.680 \text{ \AA}$). Une autre interaction de type liaison hydrogène faible a été constatée avec l'atome d'azote de la chaîne latérale du résidu PRO45 ($d = 3.326 \text{ \AA}$).

A



B

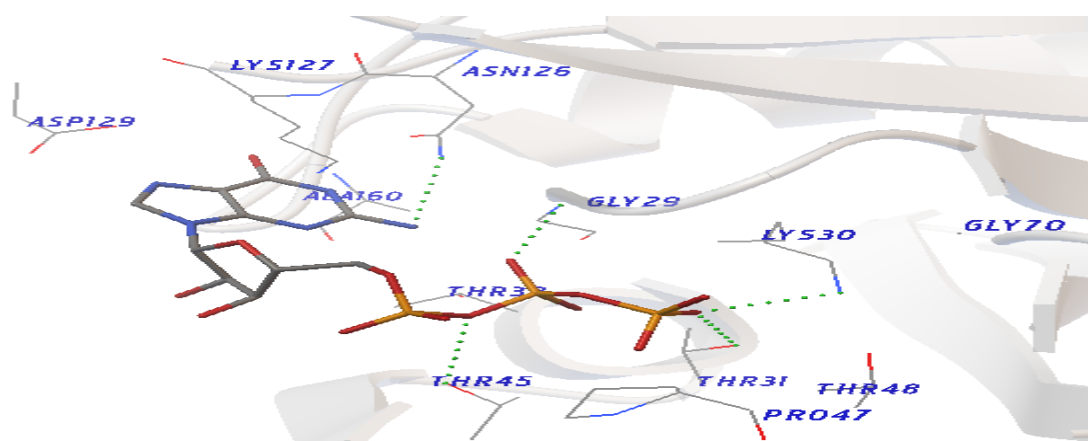


Figure 16 : visualisation des conformations pour le système GTP/Arf1 obtenue (A) en présence de Mg^{2+} et (B) en absence de Mg^{2+} . La protéine, colorée en gris, est représentée en mode cartoon, le ligand GTP en tiges et les résidus du site actif en ligne.

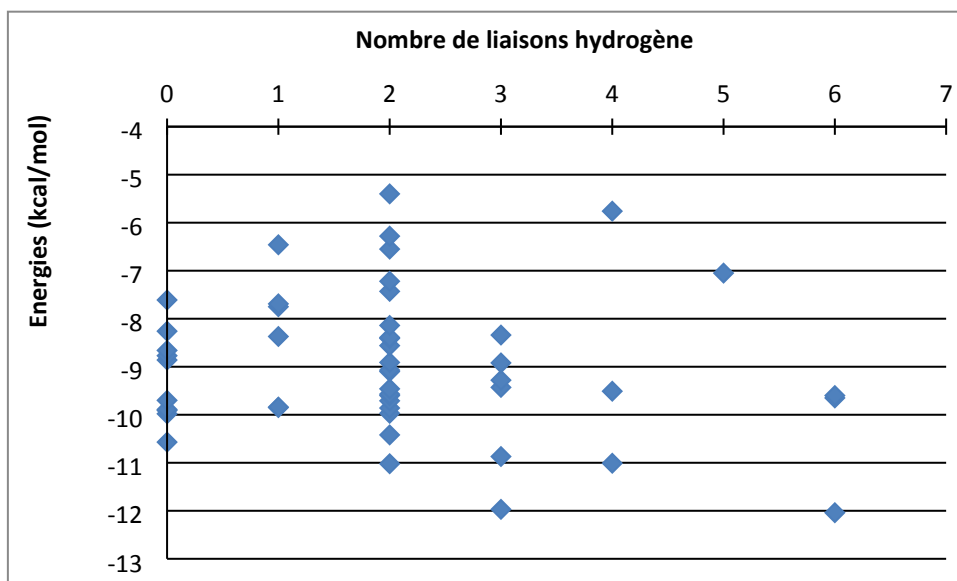
Nous résumons dans le tableau suivant les interactions qui ont lieu entre le ligand GTP et les résidus du site actif en présence et en absence de Mg²⁺. Par ailleurs, les résidus tapissant le site de liaison du ligand GTP sont conservés (tableau 2).

Tableau 2 : les énergies de liaison, les acides aminés constituant le site actif et les liaisons hydrogène entre le ligand GTP et les résidus du site actif de l'Arf1 en présence et en absence de Mg²⁺, les résidus conservés sont colorés en vert.

	Liaisonne H			AA de la protéine à moins de 4Å°	Energie (Kcal/mol)
	Atome du ligand GTP	Atome d'AA	Distance (Å°)		
GTP en présence de Mg ²⁺	O41	N(ALA27)	2.874	THR48, THR45, THR32,	-13.82
	O7	O(THR32)	2.610	THR31, LYS30, ALA27,	
	O1	O(THR32)	2.532	ALA28, GLY29, ILE42, VAL43, THR44, ILE46, PRO47, THR161, MG1002	
GTP en absence de Mg ²⁺	O 42	O (THR31)	2.816	ASP129, THR48, THR45,	-6.52
	O 42	N(LYS30)	2.940	THR32, THR31, LYS30,	
	O39	N(GLY29)	2.680	ASP26, ALA27, ALA28,	
	O1	N(THR45)	3.326	GLY29, ILE46, PRO47,	
	N20	N(ASN126)	3.057	ASN126, LYS127, CYS159, ALA160, THR161	

La figure suivante montre le nombre de liaisons hydrogène entre le ligand GDP et les résidus du site actif de la protéine en en présence et en absence de Mg²⁺.

A



B

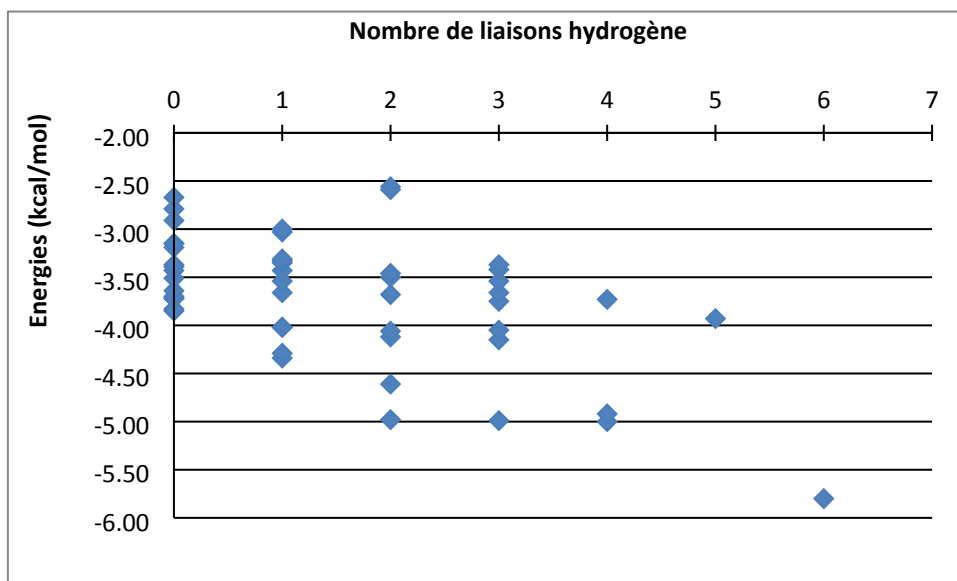
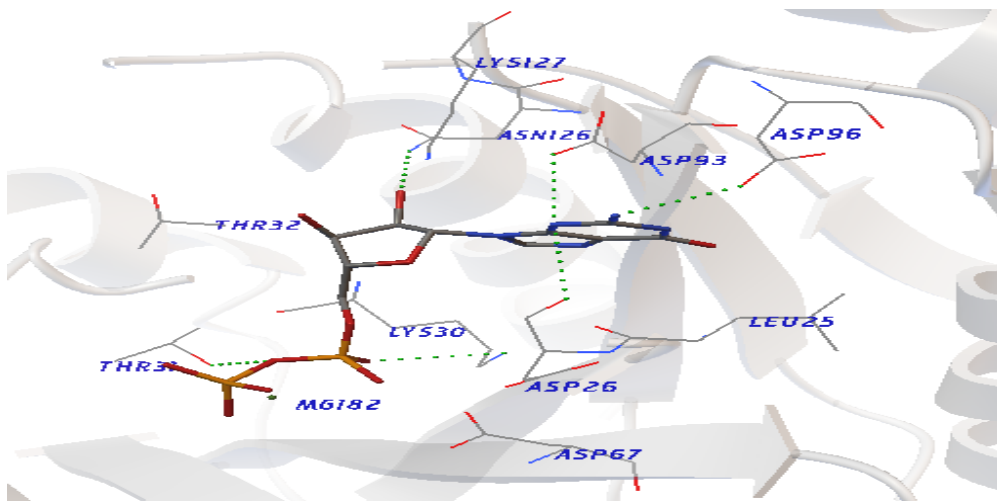


Figure 17 : Classification des conformations obtenues par les calculs de docking pour le ligand GDP A) en présence B) en absence de Mg^{2+} selon l'énergie et le nombre de liaison hydrogène

L'analyse de ces deux graphes montre que les conformations les plus stables obtenues en présence de Mg^{2+} ($E = -12,04$ kcal/mol) et en absence de Mg^{2+} ($E = -5,8$ kcal/mol) présentent respectivement six liaisons hydrogène avec les résidus du site actif de la protéine.

La visualisation des conformations les plus stables obtenues pour GDP en présence et en absence de Mg^{2+} dans le site de liaison de la protéine Arf1 (figure 8), montre que ce ligand forme plusieurs interactions électrostatiques avec les acides aminés du site actif de l'Arf1. L'analyse de ces conformations montre qu'en présence de Mg^{2+} , GDP forme six liaisons hydrogène avec les résidus du site actif. Quatre de ces liaisons ont lieu entre ce ligand et les atomes d'oxygène des chaînes latérales des acides aminés ASP96 ($d = 2.819 \text{ \AA}$), ASP93 ($d = 2.862 \text{ \AA}$), ASP26 ($d = 3.055 \text{ \AA}$) et THR31 ($d = 2.724 \text{ \AA}$). Deux autre interactions de type liaison hydrogène ont lieu entre le ligand GDP et l'atome d'azote de la chaîne latérale de LYS30 ($d = 3.162 \text{ \AA}$) et l'atome d'azote du backbone de ASN126 ($d = 2.797 \text{ \AA}$).

A



B

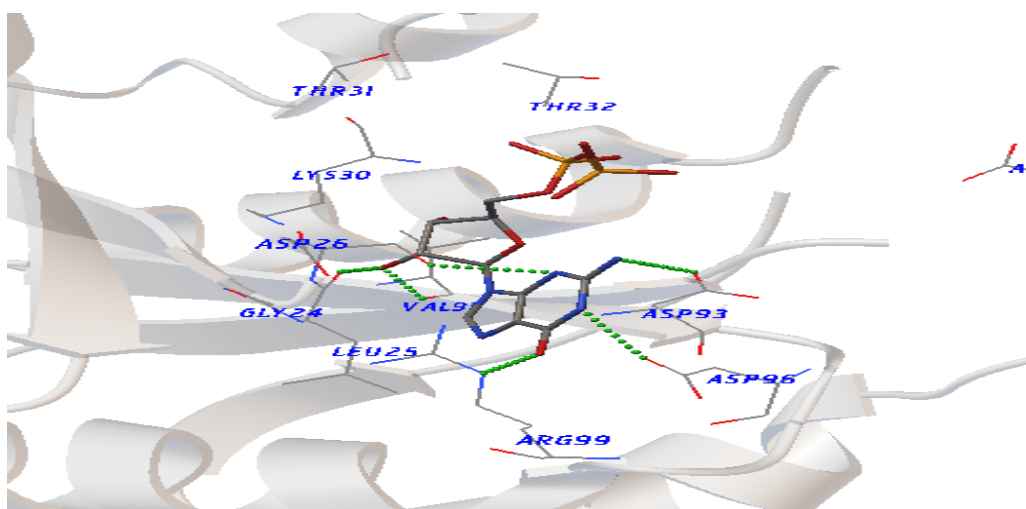


Figure 18: visualisation des conformations pour le complexe GDP/Arf1 obtenue (A) en présence de Mg^{2+} et (B) en absence de Mg^{2+} . La protéine, colorée en gris, est représentée en mode cartoon, le ligand GTP en tiges et les résidus du site actif en ligne.

Six autres liaisons hydrogène sont formées entre le ligand GDP et les acides aminés du site de liaison de la protéine Arf1 en absence de Mg^{2+} (figure 8B). Trois interactions de type liaison hydrogène ont été constatées entre ce ligand et les atomes d'oxygènes des chaînes latérales des résidus ASP96 ($d = 2.559 \text{ \AA}$), ASP93 ($d = 2.926 \text{ \AA}$) et ASP26 ($d = 3.119 \text{ \AA}$). Deux de ces interactions sont formées entre ce ligand et les atomes d'oxygène du backbone des acides aminés LEU25 ($d = 2.819 \text{ \AA}$), VAL91 ($d = 2.501 \text{ \AA}$). Une autre liaison hydrogène a été constatée entre GDP et l'atome d'azote de la chaîne latérale du résidu ARG99 ($d = 2.753 \text{ \AA}$). Par ailleurs, les résidus entourant le ligand GDP sont conservés.

Nous résumons dans le tableau suivant les différentes interactions qui ont lieu entre le ligand GDP et les résidus du site actif en présence et en absence de Mg^{2+} .

Tableau 3: les énergies de liaison, les acides aminés constituant le site actif et les liaisons hydrogène entre le ligand GDP et les résidus du site actif de l'Arf1 en présence et en absence de Mg^{2+} , les résidus conservés sont colorés en vert.

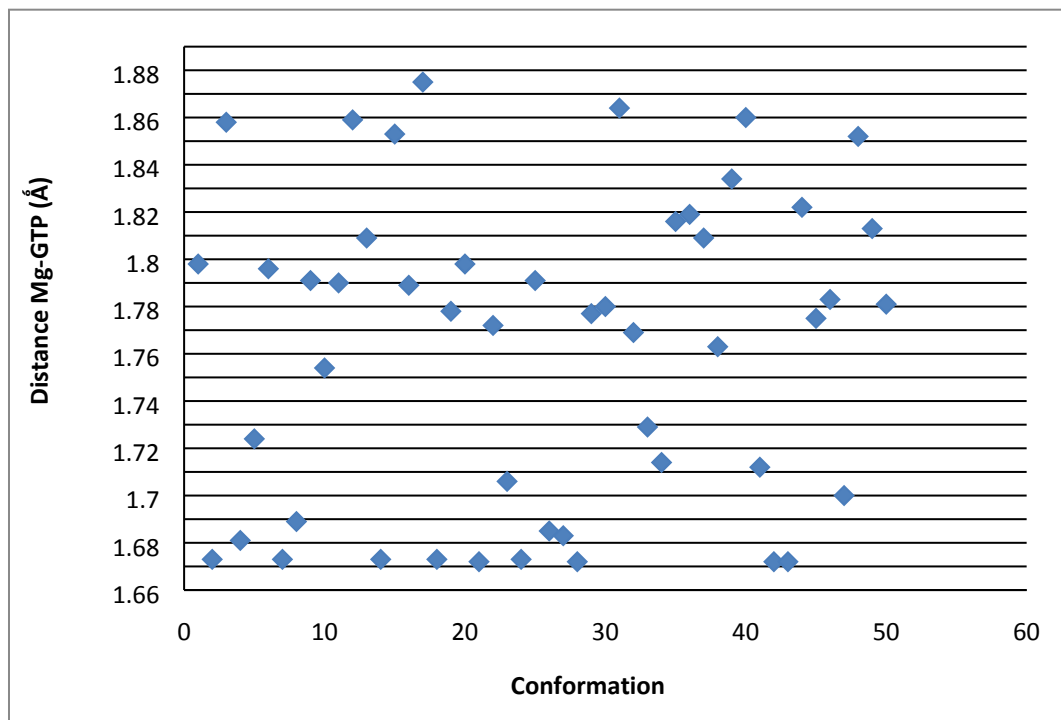
	Liaisonne H			AA de la protéine à moins de 4 Å°	Energie (Kcal/mol)
	Atome de ligand GDP	Atome d'AA	Distance (Å°)		
GDP en présence de Mg ²⁺	N 23 N20 O29 O36 N23 O34	O (ASP93) O(ASP96) N(ASN126) O(THR31) O(ASP26) N(LYS30)	2.819 2.862 2.797 2.724 3.055 3.162	THR32, THR31, LEU25, ASP26, LYS30, GLU54, ASP67, VAL91, VAL92, ASP93, ASP96, GLU99, ASN126, LYS127, MG182	-12.04
GDP en absence de Mg ²⁺	N20 N17 O29 O29 O16 N23	O(ASP93) O(ASP96) O(VAL91) O(LEU25) N(ARG99) O(ASP26)	2.926 2.559 2.501 2.819 2.753 3.119	THR32, GLY24, LEU25, ASP26, LYS30, VAL91, VAL92, ASP93, ASP96, ARG99, LYS125	-5.8

III.4- Etude des distances interatomiques entre l'ion Mg²⁺ et les ligands GTP et GDP

Nous reportons sur les deux graphes suivants les distances interatomiques entre l'ion Mg²⁺ et les deux ligands GTP et GDP. Les conformations sur ces deux graphes sont classées par ordre d'énergie croissante. L'analyse de ces résultats montre que les distances interatomiques varient de 1,669 Å à 1,879 Å pour le ligand GTP et de 2,26 Å à 2,835 Å pour le ligand GDP. La distance entre Mg²⁺ et la conformation la plus stable (E = -13,82 kcal/mol) de GTP est de 1,793 Å tandis que dans le cas de GDP (E = -12.04 kcal/mol) cette distance est de 2,264 Å. Il est bien de noter que les distances entre Mg²⁺ et les deux ligands GTP et GDP dans les structures expérimentales sont respectivement 2,205 Å et 2,426 Å. Ces résultats,

cohérents les données expérimentales, favorisent la stabilité des complexes GTP/Arf1 et GDP/Arf1.

A



B

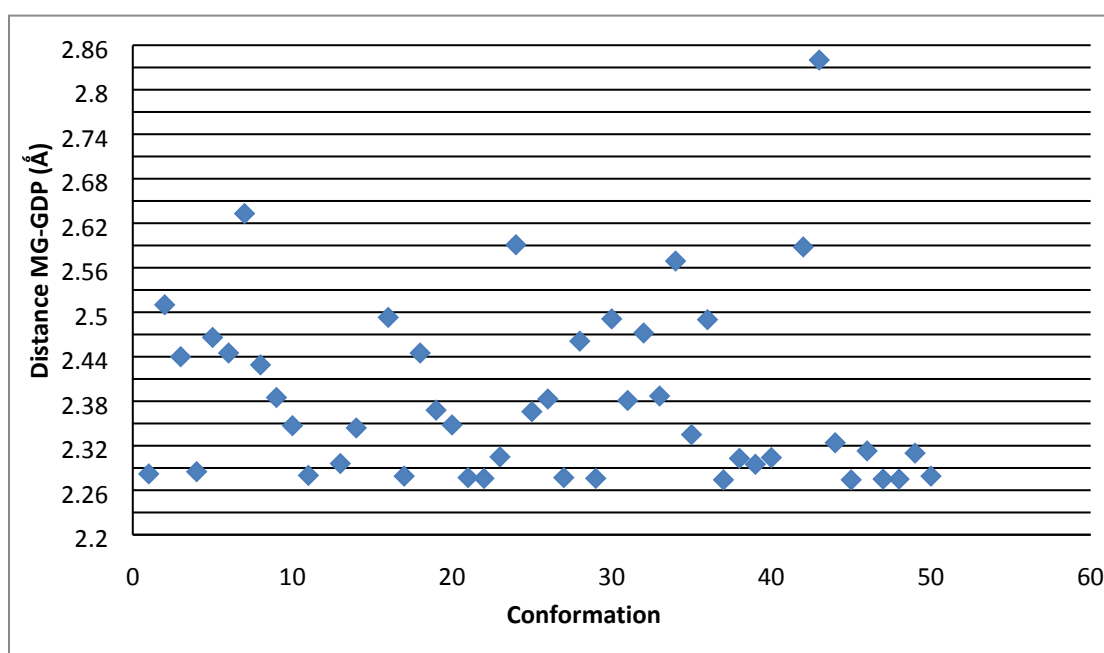


Figure 19: les énergies de liaison en fonction des distance entre A) Mg^{2+} et GTP B) Mg^{2+} et GDP

IV- Conclusion

Cette étude, menée au moyen du docking moléculaire, consiste à comprendre et à étudier le rôle de l'ion Mg^{2+} sur la liaison des deux ligands GTP et GDP à la protéine arf1. Les structures de la protéine étudiée (1HUR et 1O3Y) sont sélectionnées dans la littérature et comportent deux monomères et les ligands GDP et GTP respectivement. Une seule chaîne a été sélectionnée pour chaque structure. Les calculs de docking moléculaire ont été effectués en présence et en absence de l'ion Mg^{2+} pour voir l'effet de ce dernier sur la liaison entre ces deux ligands et la protéine Arf1. La validation des paramètres de calcul a été effectuée en comparant les conformations des deux ligands GTP et GDP issues du docking à celles issues des structures expérimentales (1O3Y et 1HUR) respectivement.

L'analyse des résultats montre que, pour les deux ligands GTP et GDP, les conformations obtenues par les calculs de docking en présence de l'ion Mg^{2+} sont plus stables que celles obtenues en absence de cet ion. De ce fait ce résultat est en parfait accord avec les résultats de la littérature.

La visualisation des conformations obtenues par les calculs de docking montre que les deux ligands étudiés forment des interactions électrostatiques de type liaison hydrogène avec les acides aminés du site actif en présence et en absence de Mg^{2+} . Le nombre de liaisons hydrogènes formées varie en fonction du ligand et de la présence ou de l'absence de l'ion Mg^{2+} . Cette étude a révélé l'importance des résidus 'tapissant' le site actif, en particulier ALA27, THR31, THR32, LYS30, GLY29, THR45, ASN126, ASP93, ASP96, ASP26, VAL91, LEU25 et ARG99, pour la liaison de ces deux ligands à la protéine Arf1.

L'analyse des distances interatomiques entre les deux ligands GDP et GTP et l'ion Mg^{2+} montre l'effet favorisant de cet ion à la stabilité des complexes GTP/Arf1 et GDP/Arf1.

En conclusion, nous pouvons dire que la présente étude, réalisée au moyen des outils de la modélisation moléculaire, a montré que l'ion Mg^{2+} favorise l'interaction des ligands GTP et GDP aux résidus tapissant le site actif de la protéine en particulier ASP93, ASP96 et THR31.

Références bibliographiques

[1] Baolin Z, Yaqin Z, Zhi-xin W, Yi Zheng. *J. Biol. Chem.*, 2000, 25299–25307.

[2] <http://www wwpdb.org/>.

[3] Amor JC, Harrison DH, Kahn RA, Ringe D., Structure of the human ADP-ribosylation factor 1 complexed with GDP, *Nature.*, 1994, 372:704-8.

[4] Shiba T, Kawasaki M, Takatsu H, Nogi T, Matsugaki N, Igarashi N, Suzuki M, Kato R, Nakayama K, Wakatsuki S, Molecular mechanism of membrane recruitment of GGA by ARF in lysosomal protein transport, *Nat. Struct. Biol.*, 2003, 10: 386-393.

[5] Michel F, Sanner M, Python: A Programming Language for Software Integration and Development., *J. Mol. Graphics Mod.*, 1999, 17:57-61.

[6] Gasteiger J, Marsali M, *Tetrahedron*, 1980, 36: 3219.

[7] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ, *Journal of Computational Chemistry*, 1998, 19: 1639-1662.

Conclusion générale

Les résultats de recherche discutés dans ce mémoire ont permis de mettre en lumière le rôle de l'ion Mg^{2+} sur la liaison des deux ligands GTP et GDP à la protéine arf1 qui est impliquée dans le cancer du sein.

Le cancer du sein est le cancer le plus fréquent chez la femme, le nombre de cas est en constante augmentation. Cette maladie représente plus de 50% des cancers gynécologiques chez les femmes en Algérie. Sa prise en charge est très complexe et l'évaluation de la thérapie anti-cancer du sein est une étape importante pour une décision thérapeutique adéquate. La mise en évidence du rôle de la protéine Arf1 dans le développement de cette pathologie a permis une nouvelle recherche pour une thérapie ciblée.

Arf1 est une petite protéine localisée au niveau de l'appareil de Golgi où elle exerce plusieurs fonctions homéostatiques comme le transport vésiculaire, la synthèse des lipides membranaires et la ségrégation des chromosomes, un événement clé de la division cellulaire. L'Arf1, comme toutes les protéines G, existe sous deux états conformationnels différents : un état inactif lié au GDP et un état actif lié au GTP. Contrairement aux autres membres de cette famille, l'arf1 est surexprimée dans les cellules invasives du cancer du sein et elle est impliquée dans la migration et la prolifération des cellules cancéreuses. Ainsi, Arf1 constitue une cible potentielle dans le traitement anti-cancer du sein.

L'ensemble de nos travaux de recherche bibliographique, regroupés dans le premier chapitre, ont permis de définir le rôle moléculaire d'Arf1 dans divers processus biologiques et de récolter des informations d'ordre expérimentales nécessaires à la réalisation de cette tâche. Les résultats de cette recherche soulignent l'importance de l'étude des petites GTPases et l'impact du débalancement de leur état d'activation et d'expression dans ce cancer.

Le deuxième chapitre est réservé aux méthodes de la modélisation moléculaire ainsi que les méthodes de drug design en particulier le docking moléculaire qui est une approche basée sur la cible biologique. Il consiste à placer une petite molécule, dite ligand, dans le site actif d'une macromolécule, dite récepteur.

Dans le chapitre III, nous avons étudié le rôle de l'ion Mg^{2+} sur la liaison des deux ligands GTP et GDP à la protéine Arf1. Cette étude est réalisée à l'aide d'une technique de la modélisation moléculaire basée sur la cible biologique, en l'occurrence le docking moléculaire. Les calculs de docking moléculaire ont été effectués en présence et en absence de l'ion Mg^{2+} pour voir l'effet de ce dernier sur la liaison entre ces deux ligands et la protéine Arf1. L'analyse des résultats montre que, pour les deux ligands GTP et GDP, les conformations obtenues par les calculs de docking en présence de l'ion Mg^{2+} sont plus stables que celles obtenues en absence de cet ion. De ce fait ce résultat est en parfait accord avec les résultats de la littérature.

L'analyse des conformations, obtenues par notre travail de docking, montre que les deux ligands étudiés forment des interactions électrostatiques de type liaison hydrogène avec les acides aminés du site actif en présence et en absence de Mg^{2+} . Le nombre de liaisons hydrogènes formées varie en fonction du ligand et de la présence ou de l'absence de l'ion Mg^{2+} . Cette étude révèle l'importance des résidus 'tapissant' le site actif, en particulier ALA27, THR31, THR32, LYS30, GLY29, THR45, ASN126, ASP93, ASP96, ASP26, VAL91, LEU25 et ARG99, pour la liaison de ces deux ligands à la protéine Arf1.

L'étude des distances interatomiques entre l'ion Mg^{2+} et les ligands GTP et GDP montre l'effet favorisant de cet ion à la stabilité des complexes GTP/Arf1 et GDP/Arf1. Ces résultats, sont parfaitement cohérents avec les données expérimentales.

En conclusion, nous pouvons dire que la présente étude, réalisée au moyen des outils de la modélisation moléculaire, a montré que l'ion Mg^{2+} favorise l'interaction des ligands GTP et GDP aux résidus tapissant le site actif de la protéine en particulier ASP93, ASP96 et THR31. Cette connaissance évoque de nouvelles questions de recherche qu'on désire exploiter dans le futur, en utilisant d'autre méthode telles que la dynamique moléculaire ou même les méthode de la QM/MM qui consistent à traiter, de manière quantique, le ligand et les résidus du site actif de la protéine.