



**HAL**  
open science

# Étude par modélisation moléculaire de l'effet allergène des antibiotiques de la famille des $\beta$ - lactamines, tant sur le plan immédiat que retardé

Julie-Anne Chemelle

► **To cite this version:**

Julie-Anne Chemelle. Étude par modélisation moléculaire de l'effet allergène des antibiotiques de la famille des  $\beta$ - lactamines, tant sur le plan immédiat que retardé. Médecine humaine et pathologie. Université Claude Bernard - Lyon I, 2010. Français. NNT: 2010LYO10318 . tel-00860118

**HAL Id: tel-00860118**

**<https://theses.hal.science/tel-00860118>**

Submitted on 10 Sep 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE L' UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

Ecole Doctorale Interdisciplinaire Sciences-Santé

DIPLOME DE DOCTORAT

(arrêté du 7 août 2006)

présentée et soutenue publiquement

*le 6 décembre 2010*

par

**M<sup>lle</sup> CHEMELLE Julie-Anne**

TITRE :

**Etude par modélisation moléculaire de l'effet allergène des antibiotiques de la famille des bêta-lactamines, tant sur le plan immédiat que retardé.**

Directeur de thèse : Dr. TERREUX Raphaël

JURY :

Pr. CARRUPT Pierre Alain	Rapporteur
Pr. DEMOLY Pascal	Rapporteur
Pr. PALLARDY Marc	Rapporteur
Pr. GOOSSENS An	Examineur
Dr. LAVERY Richard	Examineur
Dr. TERREUX Raphaël	Directeur de thèse
Pr. NICOLAS Jean -François	Co-Directeur de thèse
Dr. ROZIERES Aurore	Invité



## UNIVERSITE CLAUDE BERNARD LYON 1

Président de l'Université	M. Lionel COLLET
Vice-Président du Conseil d'Administration	M. Joseph LIETO
Vice-Président du Conseil Scientifique	M. Jean-François
MORNEX	
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Daniel SIMON

### Composantes de l'Université Claude Bernard Lyon 1

#### SANTE

UFR de Médecine Lyon RTH Laennec Directeur :	M. Pierre COCHAT
UFR de Médecine Lyon Grange Blanche Directeur :	M. Xavier MARTIN
UFR de Médecine Lyon Nord Directeur :	M. Jérôme ETIENNE
UFR de Médecine Lyon Sud Directeur :	M. François-Noël GILLY
Institut des Sciences Pharmaceutiques et Biologiques Directeur :	M. François LOCHER
UFR d'Odontologie Directeur :	M. Olivier ROBIN
Institut des Techniques de Réadaptation Directeur :	M. Yves MATILLON
Département de formation et centre de recherche en Biologie Humaine Directeur :	M. Pierre FARGE

#### SCIENCES

UFR de Biologie Directeur :	M. Hubert PINON
UFR de Chimie et Biochimie Directeur :	Mme Hélène PARROT
UFR de Mathématiques Directeur :	M. Marc CHAMARIE
UFR de Physique Directeur :	Mme Sonia FLECK
UFR Sciences de la Terre Directeur :	M. Pierre
HANTZPERGUE	
Observatoire de Lyon Directeur :	M. Bruno GUIDERDONI

#### SCIENCES ET TECHNOLOGIES

UFR de Génie Electrique et des Procédés Directeur :	M. Guy CLERC
UFR d'Informatique Directeur :	M. Samir
AKKOUCHE	
UFR de Mécanique Directeur :	M. Hamda BEN
HADID	
UFR de Sciences et Techniques des Activités Physiques et Sportives (STAPS) Directeur :	M. Claude
COLLIGNON	
Institut des Sciences et Techniques de l'Ingénieur de Lyon (ISTIL) Directeur :	M. Joseph LIETO
I.U.T. A Directeur :	M. Christian
COULET	
I.U.T. B Directeur :	M. Roger
LAMARTINE	
Institut des Sciences Financières et d'Assurance (ISFA) Directeur :	M. Jean Claude AUGROS

## ***REMERCIEMENTS***

Ce travail de thèse a été réalisé au sein du laboratoire de Bioinformatique : Structures et Interactions de l'Institut de Biologie et Chimie des Protéines, en collaboration avec l'unité INSERM U851.

Je tiens à remercier tout particulièrement les Professeurs Pierre-Alain Carrupt, Pascal Demoly et Marc Pallardy d'avoir accepté d'être les rapporteurs de ces travaux.

Je remercie également les Professeurs An Goossens, Richard Lavery, Jean-François Nicolas et le Docteur Raphaël Terreux de me faire l'honneur d'apprécier ce travail.

J'adresse mes remerciements aux Professeurs Gilbert Deléage et Richard Lavery qui m'ont accueillie au sein du laboratoire.

Merci à toute l'équipe du Laboratoire de Bioinformatique Structures et Interactions pour leurs conseils, leurs compétences et leur bonne humeur.

Mes remerciements vont aussi à nos différents collaborateurs, et notamment le Docteur Marie Baeck, le Docteur Aurore Rozières, M<sup>lle</sup> Cathy Nihm...

Enfin, un grand merci au Docteur Raphaël Terreux, mon directeur de thèse, pour son aide, son soutien, son dynamisme scientifique et sa sympathie.

<b><i>Sommaire</i></b>	<b>5</b>
<b><i>Table des figures, tableaux et schémas</i></b>	<b>9</b>
<b><i>Abréviations</i></b>	<b>13</b>
<b><i>Chapitre 1 : Des bêta-lactamines aux hypersensibilités allergiques</i></b>	<b>15</b>
I- Introduction	15
II- Les bêta-lactamines	17
II-1- Les bêta-lactamines sont des antibiotiques :	17
II-2- Historique des bêta-lactamines:	19
II-2- Classification :	20
II-2-1- Les pénicillines :	21
II-2-2- Les céphalosporines :	23
II-2-3- Les carbapénèmes :	25
II-2-4- Les monobactames :	26
II-2-5- Les inhibiteurs de bêta-lactamases :	28
II-3- Mécanisme d'action des antibiotiques de type bêta-lactamine :	29
II-3-1- Principe :	29
II-3-2- La paroi bactérienne :	30
II-3-3- Déroulement :	31
II-3-3-1- Pénétration dans la bactérie :	31
II-3-3-2- Activité antibactérienne :	31
II-3-4- Résistance aux bêta-lactamines :	33
II-3-4-1- Imperméabilisation :	33
II-3-4-2- Altération des PBP :	33
II-3-4-3- Bêta-lactamases :	33
III- Hypersensibilités allergiques médicamenteuses :	35
III-1- Hypersensibilités allergiques et non allergiques :	35
III-2- Immunogénicité du médicament :	37
III-2-1- Haptène, Prohaptène, P-i concept :	37
III-2-2- Les épitopes :	39
III-3- Initiation et déroulement d'une HSA :	40
III-4- Les différents mécanismes allergiques	42
III-4-1- Hypersensibilités allergiques de type I :	42
III-4-2- Hypersensibilités allergiques de type II :	46
III-4-3- Hypersensibilités allergiques de type III :	46
III-4-4- Hypersensibilités allergiques de type IV :	47
III-5- Régulation de la réponse d'hypersensibilité allergique :	48
III-6- Facteurs de risques :	50
III-7- Diagnostic de l'hypersensibilité allergique médicamenteuse :	50
IV- Hypersensibilités allergiques aux bêta-lactamines	52
IV-1- Immunogénicité des bêta-lactamines :	53
IV-1-1-Déterminant majeur :	54
IV-1-2-Déterminants mineurs :	54
IV-2-Réactions immédiates IgE médiées :	56
IV-2-Réactions retardées :	59

## **Chapitre 2 : La modélisation moléculaire** **63**

I- Evolution des différentes méthodes utilisées pour découvrir de nouveaux principes actifs :	65
II- Du modèle aux méthodes :	68
II-1- Les origines expérimentales du modèle :	68
II-1-1- La Résonance Magnétique Nucléaire 2D :	69
II-1-2- La cristallographie de rayons X:	69
II-2- Les modèles théoriques :	70
II-3- Les méthodes de simulation :	71
II-3-1- Les méthodes quantiques :	72
II-3-2- Les méthodes de mécanique moléculaire :	73
II-3-3- Les méthodes mixtes :	77
III- Apport de la modélisation moléculaire dans la découverte de nouveaux composés :	78
III-1-Approche sans récepteur :	78
III-1-1-Analyse conformationnelle et flexibilité moléculaire :	78
III-1-2-Relations structure - activité quantitative (QSAR) :	79
III-1-3- Les pharmacophores :	82
III-1-4- La méthode de De novo design sans récepteur :	83
III-2- Approche avec récepteur	84
III-2-1- La méthode d'arrimage (Docking)	84
III-2-2- Méthode du De novo design avec récepteur :	85
III-2-3- La méthode de « Multiple Copy Simultaneous Search (MCSS) » :	86
III-2-4- Calcul d'énergie libre :	86
III-3- Modélisation des structures 3D à haut débit : le logiciel Modeome :	88
III-4-Le logiciel SuMo (Surfing the Molecules) :	89

## **Chapitre 3 : Le logiciel SuMo – Optimisation et nouveaux paramètres** **90**

I- Principe du logiciel SuMo	90
II- Concept d'origine	93
II-1- Description des liaisons hydrogène	94
II- 2- Description des autres fonctions chimiques et types géométriques :	96
II-3- Poids des objets :	97
II-4- Critiques du dictionnaire d'origine :	101
III- Nouvelle version de SuMo	102
III-1- Concept: un pharmacophore appliqué aux protéines	102
III-2- Nouveau dictionnaire :	104
III-2-1- Flexibilité des protéines :	104
III-2-2- Les nouveaux objets :	106
III-2-3- Poids des objets :	110
III-3- Conclusions et remarques :	111
IV- Tests comparatifs entre SuMo d'origine et SuMo actuel :	112
IV-1-Test des lectines :	112
IV-2-Test des protéases :	115
V- Critique globale et conclusions	119

<b>Chapitre 4 : Classification et réactivité moléculaire des bêta-lactamines</b>	<b>121</b>
I- Classification des bêta-lactamines - Relation Structure-Activité:	121
I-I- Création d'un dendrogramme pour aider au choix d'une molécule de remplacement chez un patient hypersensible :	121
I-2- Création de relations structure-activité quantitatives :	125
II- Réactivité moléculaire :	126
II-1- Etude par Mécanique Quantique :	128
II-3- Etude de la réaction de fixation sur l'albumine :	135
II-3-1 Docking :	136
II-3-2 Etude mixte Mécanique Quantique / Mécanique moléculaire :	142
II-4- Recherche d'autres protéines d'hapténisation : approche SuMo :	144
II-5- Discussion :	147
<b>Chapitre 5 : Hypersensibilité allergique immédiate aux bêta-lactamines</b>	<b>152</b>
I- Validation du dendrogramme :	152
II- Reconnaissance moléculaire des bêta-lactamines par une IgE :	158
II- 1- Modèle d'IgE :	159
II-2- Modèle du médicament :	160
II-3 Protéines d'hapténisation :	160
III- Analyse de l'albumine hapténisée :	161
III-1- Les bioconjugués :	161
III-1-1 Synthèse :	161
III-1-2 Analyse par modélisation moléculaire :	162
III-1-3- Interaction IgE-médicament :	168
III-2- Applications :	168
IV- Perspectives :	169
<b>Chapitre 6 : Hypersensibilité allergique retardée aux bêta-lactamines</b>	<b>171</b>
I- Différents acteurs pour différents scénarios :	172
I-1- Les modèles :	174
I-1-1- CMH :	174
I-1-2- Médicament :	175
I-1-3- Peptide :	175
I-1-4- TCR :	175
I-2- Les scénarios :	177
I-2-1- Mode classique :	177
I-2-2- P-i concept: Pharmacological interaction with the immune system:	179
I-2-2-1- Principe	179
I-2-2-2- Modélisation	180
I-2-3- "New P-i"	183
I-2-3-1- Principe	183
I-2-3-2- Modélisation	184
I-2-4- Absence de peptide	188
I-3- Conclusions :	189
II- Validation expérimentale :	191
II-1- Bioconjugués albumine – pénicilline :	192
II-2- Bioconjugués peptide-pénicilline	193
III- Perspectives	201



<i>Conclusion générale</i>	203
<i>Annexe</i>	205
<i>Références bibliographiques</i>	208

## **FIGURES, TABLEAUX, SCHEMAS**

Figure 1 : Noyau commun des pénicillines

Figure 2 : Structure des céphalosporines

Figure 3 : Structure des carbapénèmes

Figure 4 : Structure des monobactames

Figure 5 : Structure des inhibiteurs de bêta-lactamases

Figure 6 : Physiopathologie de l'HSAI

Figure 7 : Principales voies métaboliques de la pénicilline, in vivo, aux conditions physiologiques

Figure 8 : Structure d'une immunoglobuline de type E

Figure 9 : Structure des molécules de CMH de type 1 et 2

Figure 10 : Les différents termes de champ de forces Amber.

Figure 11 : Termes définissant l'énergie totale du système dans un système mixte QM/MM

Figure 12 : Alignement des 42 antibiotiques sur le noyau  $\beta$ -lactame

Figure 13 : Alignement des 42 antibiotiques sur le pharmacophore

Figure 14 : Dendrogramme obtenu par alignement rigide de 42 molécules antibiotiques

Figure 15 : Surfaces d'isoprépondérance des descripteurs sur une bêta-lactamine

Figure 16 a : Modèles utilisés pour l'étude par mécanique quantique

Figure 16b : Structure des intermédiaires réactionnels calculés par mécanique quantique

Figure 17 : Site actif de la PFP en interaction avec la molécule de cefuroxime

Figure 18 : Formules d'évaluation de l'affinité d'une molécule utilisées par la fonction consensus CScore

Figure 19 : Interaction entre l'amoxicilline et le site K190

Figure 20: Interaction entre l'amoxicilline et le site K212

Figure 21: Modèle 3D de l'amoxicilline fixée sur la lysine 190 de l'albumine.

Tableau 11 : Protéines obtenues par l'approche SuMo et ayant un site proche du site de fixation des bêta-lactamines de la PFP.

Figure 22 : Superposition du site de fixation des pénicillines de la PFP avec un site de la L-Ficoline

Figure 23 : Formule décrivant le calcul du pKa d'un acide aminé

Figure 24 : Albumine sérique humaine hapténisée par 10 molécules d'amoxicilline

Figure 25 : Surface d'accès au solvant de l'albumine sérique humaine libre et de l'albumine sérique humaine hapténisée en position 190

Figure 26 a : Surface d'accès au solvant colorée en fonction des charges partielles pour l'albumine sérique humaine modifiée en position 190

Figure 26 b : Surface d'accès au solvant colorée en fonction des charges partielles pour l'albumine sérique humaine libre

Figure 27 : Surface d'accès au solvant de l'albumine sérique humaine libre et de l'albumine sérique humaine hapténisée en position 212

Figure 28 a : Surface d'accès au solvant colorée en fonction des charges partielles pour l'albumine sérique humaine modifiée en position 212

Figure 28 b : Surface d'accès au solvant colorée en fonction des charges partielles pour l'albumine sérique humaine libre

Figure 29 : Complexe entre une molécule de HLA-DR1, le peptide hémagglutinine et le TCR

Figure 30 : Analyse par RMS-2D du complexe formé par HLA-DR1, le peptide hapténisé par l'amoxicilline et le TCR

Figure 31 : Conformation stable issue de la dynamique moléculaire du complexe HLA-DR1, peptide hapténisé par l'amoxicilline et TCR

Figure 32 : Analyse par RMS-2D du complexe formé par HLA-DR1, le peptide en interaction non covalente avec l'amoxicilline et le TCR

Figure 33 : Complexe formé entre HLA-DR1, peptide en interaction non covalente avec l'amoxicilline et TCR selon la configuration de la première sous-famille conformationnelle

Figure 34 : Vue schématique du New « pharmacological interaction with immune receptors » Concept

Figure 35 : Structure du complexe entre une molécule d'amoxicilline, le cuivre Cu<sup>2+</sup> et le dipeptide histidine-histidine

Figure 36 : Analyse par RMS-2D du complexe formé entre HLA-DR1, peptide modifié, cuivre, amoxicilline et TCR

Figure 37 : Conformation du complexe entre HLA-DR1, peptide modifié, cuivre, amoxicilline et TCR.

Figure 38 a : Analyse par RMS-2D du complexe HLA-DR1, amoxicilline, et TCR

Figure 38 b : Analyse par RMS-2D du complexe HLA-DR1, cuivre II, amoxicilline, et TCR

Figure 39 : Signature SuMo du site de référence (PFP)

Figure 40: Résultats de la comparaison du site de fixation des pénicillines de la PFP avec les modèles de la Serum Databank : recherche de protéines d'hapténisation.

Schéma 1a : Réaction de transpeptidation

Schéma 1b : Inhibition de la transpeptidation par la bêta-lactame

Schéma 2a : Inactivation de la transpeptidation

Schéma 2b : Inactivation de la  $\beta$ -lactame par une bêta-lactamine

Schéma 3 : Structure du TCR

Schéma 4 : Schéma général de la recherche d'une molécule active

Schéma 5: courbe énergétique classique des réactions chimiques.

Schéma 6 : Réaction chimique entre le noyau-P et la méthylamine

Schéma 7 : Diagrammes énergétiques des réactions impliquant les noyaux-P et -C et les modèles de lysine et sérine.

Tableau 1 : Description des liaisons hydrogène par les objets delta plus et minus pour chaque acide aminé

Tableau 2 : Objets SuMo d'origine : types géométriques et acides aminés porteurs

Tableau 3 : Description des acides aminés par les anciens objets : position et poids des objets

Tableau 4 : Analyse de la flexibilité de la protéine kinase C par dynamique moléculaire

Tableau 5 : Description des acides aminés par les nouveaux objets

Tableau 6 : Poids des nouveaux objets

Tableau 7 : Résultats comparatifs des trois tests SuMo sur le jeu des lectines

Tableau 8 : Résultats comparatifs des trois tests SuMo sur le jeu des protéases et isomérases

- Tableau 9 : Energies d'activation des réactions de liaisons des noyau-C et-P sur la méthylamine et le méthanol.
- Tableau 10 : CScore d'interaction entre l'amoxicilline (centrée sur la lysine numéro X) et l'albumine.
- Tableau 11 : Résultats de la comparaison entre le site de fixation de la pénicilline de la PBP et les protéines de la Serum Databank
- Tableau 12 : Résultat du calcul des pKa des acides aminés de type lysine de l'albumine sérique humaine
- Tableau 13 : Tableau des données cliniques de 42 patients souffrant d'HSA Immédiate, sur une batterie de 17 molécules
- Tableau 14 : Pourcentages de molécules positives par famille et par patient
- Tableau 15 : Neuf peptides issus des protéines candidates analysées par SuMo et reconnus par les molécules de CMH classe 2.
- Tableau 16 : Analyses des interactions mises en jeu lors d'une reconnaissance de l'amoxicilline par le système CMH/TCR
- Tableau 17 : Orientation des chaînes latérales des résidus peptidiques au sein de la niche à peptide
- Tableau 18 (Annexe) : Quarante deux structures de bêta-lactamines

## **ABREVIATIONS**

AA : Acide Aminé

ADF : Amsterdam Density Functional Software

ADN: Acide DésoxyriboNucléique

Å : Angström

ARNm et t : Acide RiboNucléique messenger et de transfert

Atm : Atmosphère

BCR : B Cell Receptor soit récepteur des lymphocytes B

BLAST : Basic Local Alignment Search Tool

CE : Combinatorial Extension

CD : Cellule Dendritique

CDR : Complementarity Determining Region soit région reconnaissant l'antigène

CHARMM : Chemistry at HARvard Molecular Mechanics

CMC/MD : Chemical Monte-Carlo / Molecular Dynamics

CMH : Complexe Majeur d'Histocompatibilité

CMI : Concentration Minimum Inhibitrice

CoMFA : Comparative Molecular Field Analysis

COSY : COrrrelation SpectroscopY

CPA : Cellule Présentatrice de l'Antigène

DFT : Density Functional Theory

DSSP : Dictionnaire de Structures Secondaires de Protéines

Fab : Fragment Antigen Binding pour Fragment de liaison l'Antigène

FASTA : FAST-ALL

Fc R : Récepteur du Fragment constant des immunoglobulines

FEP : Free Energy Perturbation pour perturbation d'énergie libre

GGA/PW : Approximation des Gradients Généralisée et l'algorithme de Perdew/Wang

HF : Hartree Fock

HLA : Human Leukocyte Antigen

HSA I et HSA R : HyperSensibilité Allergique Immédiate ou Retardée

HSA: Human Serum Albumine soit albumine sérique humaine

HTS: High-Throughput Screening

IFN : Interféron

Ig : Immunoglobuline

IL : InterLeukine

IRC : Intrinsic Reaction Coordinate soit Coordonnées Intrinsèques de la Réaction  
ITAM et IM : Immunoreceptor Tyrosine-based Activation Motif et Immunoreceptor Tyrosine-based Inhibition Motif  
K: Kelvin  
LCAO : Linear Combination of Atomic Orbitals soit Combinaison Linéaire d'Orbitales Atomiques  
LB : Lymphocyte B  
LT : Lymphocyte T  
MCSS : Multiple Copy Simultaneous Search soit recherche simultanée de copies multiples  
MM: Molecular Mechanic soit Mécanique Moléculaire  
MMFF : Merck Molecular Force Field  
NAMD : NANoscale Molecular Dynamics  
NOESY : Nuclear Overhauser Effect Spectroscopy  
PBP ou PFP : Penicillin-Binding Protein pour Protéine Fixant les Pénicillines  
PKC : Protéine Kinase C  
QM : Quantum Mechanic soit Mécanique Quantique  
QSAR : Quantitative Structure Analysis Relationship soit Relation Structure-Activité Quantitative  
RMN : Résonance Magnétique Nucléaire  
RMS : Root Mean Square ou distance quadratique moyenne  
RSA : Receptor Shape Analysis  
Screening : criblage  
SMX : SulfaMéthoXazole  
SUMO : SURfing the Molecules  
TAB : Test d'Activation des Basophiles  
TCR : T Cell Receptor soit récepteur des lymphocytes T  
TNF : Tumor Necrosis Factor  
VIH: Virus de l'Immunodéficience Humaine  
VMD : Visual Molecular Dynamics





# Chapitre 1 : Des bêta-lactamines aux hypersensibilités allergiques

## I- Introduction

La modélisation moléculaire constitue une approche scientifique novatrice, originale, qui permet de travailler *in silico* dans des domaines extrêmement variés, que ce soit dans un but très fondamental, ou beaucoup plus appliqué. Elle se base sur des modèles moléculaires théoriques et offre des possibilités d'études extrêmement variées, allant de l'interaction d'une petite molécule avec son récepteur, à la réaction chimique de fixation d'un ligand sur sa cible, ou l'analyse de la flexibilité d'un système biologique...C'est une technique récente en plein essor grâce aux développements informatiques actuels qui permet de renforcer de façon complémentaire les approches *in vivo* et *in vitro* des différentes études menées. Ainsi, en tant que modélisateurs, nous travaillons en étroite collaboration avec l'ensemble de la communauté scientifique et complétons le cercle de conception d'une molécule active basé sur la triade : étude *in silico*, validation *in vitro* et *in vivo*.

Dans ce travail de thèse, nous avons exploité les capacités de la modélisation moléculaire afin de comprendre les processus immunitaires impliqués dans les réactions d'hypersensibilité allergique dues aux antibiotiques de la famille des bêta-lactamines.

Les hypersensibilités allergiques médicamenteuses sont des pathologies mettant en jeu le système immunitaire et induites par la prise de médicaments (Johansson *et al.*, 2001). Elles représentent une part importante de la iatrogénie médicamenteuse. Il faut noter que très souvent, même si les réactions observées présentent des symptômes cliniques qui évoquent des hypersensibilités allergiques, elles ne sont pas médiées par le système immunitaire. Toutefois, les hypersensibilités allergiques dues aux médicaments sont responsables d'une mortalité, de surcoûts et d'une morbidité encore mal évalués, qui poussent de nombreux chercheurs et

scientifiques à mieux comprendre et tenter de soigner cette pathologie. Les mécanismes ainsi que les manifestations cliniques associées sont extrêmement variés. Celles-ci peuvent aller d'une simple urticaire localisée au choc anaphylactique parfois mortel et incluent l'ensemble des quatre types de réactions immunologiques de la classification de Gell et Coombs (Demoly et al, 2000). Sur un plan épidémiologique, et malgré le fait qu'il n'existe pas de réelles études épidémiologiques systématiques, on estime que les allergies aux médicaments représenteraient jusqu'à 30% des effets secondaires dus aux médicaments (Demoly et al., 2003). Toutefois, leurs diagnostics ne sont pas faciles à confirmer et il est de ce fait important d'améliorer nos connaissances fondamentales sur ces phénomènes mais aussi de les appliquer dans un but de pratique thérapeutique et diagnostique.

Aujourd'hui, la famille des bêta-lactamines est la classe d'antibiotiques la plus fréquemment utilisée (78% du marché des antibiotiques en France) et impliquée dans les allergies médicamenteuses. Ce type d'antibiotiques est souvent essentiel dans le traitement des maladies infectieuses mais la description croissante et fréquente de réactions anaphylactiques aux pénicillines dénonce un problème de santé publique. Tous âges confondus, environ 10 % des sujets traités par des bêta-lactamines rapportent des manifestations susceptibles d'évoquer une hypersensibilité allergique. La majorité des réactions présumées allergiques aux bêta-lactamines sont cutanées, susceptibles de survenir dans les minutes suivant la prise du médicament (urticaires...) ou dans les jours (toxidermies érythémateuses...).

Face à ce problème de santé publique, les travaux de recherche se multiplient mais la prédiction des allergies aux xénobiotiques (médicaments, cosmétiques et aliments en particulier) est très insuffisante. En effet, l'impact économique et social de ces affections est extrêmement fort. Il est donc essentiel de les identifier dès la première réaction allergique, afin d'éviter les conséquences graves d'une réexposition. Les tests cutanés actuels, bien qu'efficaces, nécessitent le déplacement du patient dans un centre spécialisé, voire son hospitalisation, et présentent un risque de sensibilisation important à prendre en considération (prédisposition à une future allergie).

Le but de ce travail présente un aspect fondamental qui vise à analyser au niveau moléculaire les interactions mises en jeu entre les différents acteurs impliqués dans les réactions

d'hypersensibilité allergique immédiate et retardée liées aux bêta-lactamines. D'autre part, nous visons un aspect plus pratique offrant une aide à la prescription pour le médecin allergologue et au diagnostic pour le biologiste.

La première partie de ce travail sera bibliographique et décrira tout d'abord les antibiotiques de la famille des bêta-lactamines, les réactions d'allergies médicamenteuses dues à ces molécules, puis les bases de la modélisation moléculaire. La seconde partie présentera les résultats obtenus avec dans un premier temps, l'étude de la réactivité moléculaire des bêta-lactamines, puis dans un second temps, les analyses concernant les hypersensibilités immédiates et retardées.

## **II- Les bêta-lactamines**

Les bêta-lactamines ( $\beta$ -lactamines) ou antibiotiques  $\beta$ -lactame sont une large classe d'antibiotiques qui comprennent les dérivés de la pénicilline, les céphalosporines, les monobactames, les carbapénèmes et les inhibiteurs des  $\beta$ -lactamases. Aujourd'hui, elles constituent la classe d'antibiotiques la plus utilisée dans le monde et sont la cause la plus fréquente d'allergies médicamenteuses.

### **II-1- Les bêta-lactamines sont des antibiotiques :**

Toutes les découvertes médicales effectuées au cours du XXe siècle sont importantes. Néanmoins l'antibiotique, dans la mesure où il a fait reculer le taux de mortalité en permettant de guérir certaines maladies infectieuses, est sans doute le médicament dont la découverte a le plus bouleversé la médecine et la démographie. Le premier antibiotique (de synthèse) a ouvert une voie nouvelle dans la lutte contre de nombreuses maladies qui étaient jusqu'alors considérées comme incurables. Les antibiotiques ont augmenté l'espérance de vie d'environ 15 ans. Comparativement, un médicament qui guérirait 100 % des cancers n'augmenterait l'espérance de vie des patients traités que de 5 ans.

Un antibiotique (du grec anti : « contre », et bios : « la vie ») est une substance d'origine naturelle ou synthétique possédant une action spécifique de blocage et/ou de destruction des

bactéries. Pour les autres micro-organismes, on utilisera les termes d'antifongique (lutte contre les champignons), et d'antiviral (lutte contre les virus).

Une substance antibiotique est donc un médicament qui empêchera la prolifération des bactéries et l'on parlera alors d'un "bactériostatique", ou les tuera de façon ciblée et l'on parlera d'un "bactéricide".

Il existe plus de 10 000 molécules antibiotiques mais une centaine seulement sont efficaces et utilisables à des fins thérapeutiques. La plupart des antibiotiques sont naturels et produits par les procaryotes (actinomycètes surtout ou eubactéries), les champignons (17 % des antibiotiques), les végétaux supérieurs, les animaux et les lichens. Les antibiotiques sont divisés en familles : les bêta-lactamines, les aminosides, les macrolides, les cyclines, les quinolones, les polypeptides, les glycopeptides, les phénicolés, les nitro-furanes. Leurs modes d'action sont variables. On distingue alors les antibiotiques agissant sur :

- La paroi bactérienne. Ces antibiotiques agissent sur des cibles extracellulaires et ne sont actifs que sur les germes en croissance. Les cellules au repos ne sont pas perturbées par l'action de ces molécules qui bloquent la synthèse de la paroi, provoquant ensuite la dégradation de la cellule sous l'effet de la pression osmotique interne. C'est notamment le cas des bêta-lactames dont nous reprendrons plus précisément le mécanisme d'action.
- La membrane des cellules. Ce mécanisme concerne la polymyxine, surfactant qui désorganise la bicouche phospholipidique membranaire. Cela conduit à la destruction de l'intégrité de la membrane entraînant la sortie des éléments hydrosolubles de la cellule. Cette molécule est efficace sur les cellules en croissance et au repos.
- L'ADN. Plusieurs types d'antibiotiques procèdent de cette façon : Les sulfamidés bloquent une enzyme impliquée dans la synthèse de l'acide folique qui est un cofacteur de la synthèse des bases puriques et pyrimidiques. Ainsi, la cellule meurt par carence en bases nucléiques. La mitomycine, dont la structure est asymétrique, et l'actinomycine de structure symétrique se fixent sur les brins de l'hélice d'ADN et établissent un pontage entre eux empêchant ainsi la réplication de l'ADN. Les quinolones et fluoroquinolones inhibent l'ADN gyrase (qui contrôle le surenroulement de l'ADN) empêchant la réplication de l'ADN bactérien.

- Le ribosome bactérien. C'est le cas de très nombreux antibiotiques. Les aminosides se fixent sur la petite sous-unité des ribosomes et empêchent la traduction de l'ARNm. Les phénicolés en fixant la grosse sous-unité du ribosome bactérien inhibent la formation de la liaison peptidique. Les cyclines bloquent l'élongation de la chaîne polypeptidique. Les macrolides et kétolides agissent dans le même sens mais s'attachent pour leur part à la grosse sous-unité. Enfin, la puromycine mime l'extrémité d'un ARNt et bloque l'élongation de la chaîne polypeptidique.

## II-2- Historique des bêta-lactamines:

Le premier antibiotique, identifié dès la fin du XIXe siècle par Ernest Duchesne, fut la pénicilline. Ses propriétés furent redécouvertes en 1928 par Sir Alexander Fleming qui s'aperçut que certaines de ses cultures bactériennes dans des boîtes oubliées avaient été contaminées par les expériences de son voisin de pailleuse étudiant un champignon : le *Penicillium notatum*. Ce chercheur écossais travailla pendant plusieurs années à essayer de purifier cet antibiotique (Galmiche, 1999). Ce n'est qu'en 1940 que deux autres chercheurs Florey et Chain, purifièrent la pénicilline. Elle ne fut toutefois exploitée en thérapeutique que vers 1940 et ne servait jusque-là qu'à nettoyer les boîtes de Petri de leurs bactéries. L'exploitation de mutants a permis d'augmenter le rendement de la production (Bryskier, 1996) de *Penicillium chrysogenum*. Dix ans plus tard, la benzylpénicilline (pénicilline G) était disponible pour l'utilisation clinique. Son usage fut toutefois limité par son instabilité au pH acide, sa susceptibilité à la destruction par une  $\beta$ -lactamase (pénicillinase), et sa relative inactivité antibactérienne contre les bacilles Gram négatif. En 1957, l'isolement du noyau commun, l'acide 6-aminopénicillanique, ouvre au développement de nombreuses pénicillines semi-synthétiques, et permet donc l'élargissement du potentiel d'utilisation : stabilité à pH acide, résistance aux  $\beta$ -lactamases, activité antibactérienne contre les bactéries Gram positif et Gram négatif.

Les céphalosporines se sont développées en parallèle, sous l'impulsion de Brotzu qui isola en 1948 le *Cephalosporium acremonium*. A partir de cette moisissure a été isolé le noyau commun des céphalosporines : l'acide 7-aminocéphalosporanique (7-ACA) à partir duquel les autres molécules de ce type furent synthétisées.

Les carbapénèmes, isolées à partir de souches de *Streptomyces*, et les monobactames à partir de bactéries non pathogènes, ont été découverts dans les années 1970.

## II-2- Classification :

Les bêta-lactamines possèdent toutes un noyau azétidine nitré à quatre sommets : le cycle bêta-lactame qui correspond à la partie efficace de la molécule. Des variations au niveau de la chaîne latérale naturelle ou greffée permettent de modifier les propriétés de la molécule antibiotique.

Malgré une structure de base commune, on distingue plusieurs familles de bêta-lactames en fonction de la nature du cycle accolé au noyau commun (Donowitz et al., 1988). Les pénames ou pénicillines, les céphèmes ou céphalosporines, les carbapénèmes et les monobactames. A ces quatre sous-familles, il faut ajouter enfin les inhibiteurs des  $\beta$ -lactamases qui n'ont pas d'activité antibactérienne intrinsèque mais permettent de prévenir la dégradation de la bêta-lactamine co-administrée.

Les pénicillines sont classées en fonction des différents substituants fixés à l'acide 6-aminopénicillanique, ce qui détermine des propriétés pharmacologiques et antibactériennes différentes.

Les céphalosporines sont classées par génération et en fonction de leur spectre d'action antibactérien.

Les monobactames sont actifs sur les bacilles Gram négatifs et résistent relativement bien aux  $\beta$ -lactamases. Ils sont inactifs sur les bacilles Gram positifs.

Les carbapénèmes sont des  $\beta$ -lactames à large spectre réservés aux infections résistantes et aux infections à *Enterobacter*.

## II-2-1- Les pénicillines :

Les pénicillines naturelles sont des molécules synthétisées par certains champignons microscopiques de la famille des *Penicillium*. Elles ont des propriétés antibiotiques bactériostatiques : elles empêchent la synthèse de la paroi bactérienne et bloquent donc la prolifération des bactéries. Ce sont les premiers antibiotiques obtenus à l'échelle industrielle et leur utilisation pour traiter les infections a entraîné une véritable révolution médicale au cours de la deuxième moitié du vingtième siècle (Weintraub, 2003).

Les pénicillines constituent une famille structurale très homogène. Elles possèdent en effet toutes un motif commun et une chaîne latérale qui porte les modifications structurales. La figure 1 présente la structure des pénicillines.

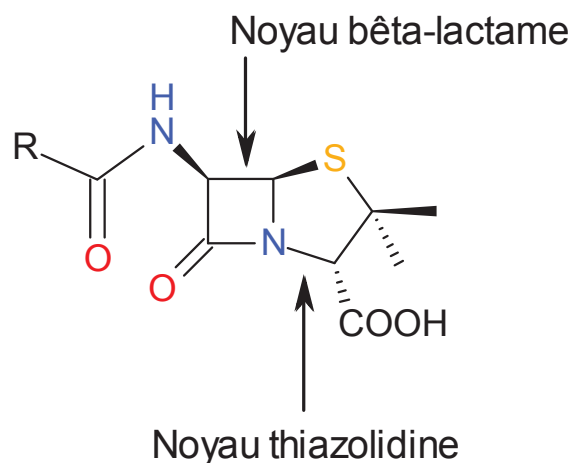


Figure 1 : Structure des pénicillines

Le cycle  $\beta$ -lactame associé à un noyau thiazolidine forme l'acide 6-aminopénicillanique qui représente la structure de base des pénicillines. L'acide 6-aminopénicillanique constitue le noyau de base des pénicillines : le noyau pénème. Il peut être substitué par acylation sur sa fonction aminée ce qui aboutira à des dérivés se distinguant par leur stabilité, leur pharmacocinétique, leur spectre d'action ainsi que leur résistance aux bêta-lactamases. De plus, la fonction carboxylique peut être transformée en carboxylate (donnant des composés plus solubles)

ou estérifiée (donnant alors des prodrogues). Notons toutefois que cette fonction acide appartient au pharmacophore et de ce fait est essentielle à l'activité antibactérienne.

Les pénicillines possèdent trois carbones asymétriques (RRS) et une modification entraîne une perte d'activité (formation en livre ouvert). Le tableau 18 présenté en annexe donne plusieurs structures de pénicillines.

Les pénicillines peuvent se diviser en cinq sous groupes :

- Les pénicillines sensibles aux pénicillinases avec tout d'abord la pénicilline G ou benzylpénicilline, obtenue par incorporation d'acide phénylacétique au milieu de culture. Elle est sensible à l'acidité gastrique et de ce fait constitue la forme intraveineuse de la pénicilline. Elle est active sur les streptocoques, méningocoques, gonocoques, pneumocoques, leptospires, tréponèmes, clotridies et *Corynebacterium dophtheriae*. On l'utilise pour des infections plus sévères où on ne peut s'en remettre à la pénicilline sous forme orale. Ensuite, la pénicilline V ou phénoxy méthylpénicilline est la forme orale de la pénicilline grâce à sa résistance à l'acidité gastrique. Elle provient de l'ajout d'acide phénoxyacétique au milieu de culture. Son spectre d'action est très proche de la pénicilline G.
- Les pénicillines résistantes aux pénicillinases : On parle aussi de pénicillines M. C'est le cas de la méticilline, l'oxacilline et la cloxacilline qui dérivent de l'acide 6-amino-pénicillanique et possèdent le même spectre d'action, soit celui des pénicillines conventionnelles. Elles sont indiquées surtout pour les infections à staphylocoques même si il existe des souches méti-résistantes.
- Les aminopénicillines, ou pénicillines A, avec principalement l'ampicilline et l'amoxicilline, sont caractérisées par un spectre d'action élargi touchant les bactéries Gram positif : streptocoques et pneumocoques et Gram négatif : *E.Coli*, *Proteus mirabilis*, *salmonella*, *shigella*... qui ne produisent pas de pénicillinases. Elles sont administrées par voie orale du fait de leur acido-résistance.
- Les pénicillines anti-pseudomonales, avec les carboxypénicillines (ticarcilline) et les uréidopénicillines (pipéracilline et mezlocilline). Leur spectre d'action concerne plusieurs



germes Gram négatif dont *Pseudomonas aeruginosa*, *proteus*, *enterobacter*. Ils n'existent que sous forme injectable en milieu hospitalier.

- Les amidinopénicillines dont le spectre d'action est étroit et utilisé contre certaines entérobactéries. On trouvera le mécillinam et le pivmécillinam, qui en est la prodrogue.

Afin de maintenir une activité face aux bactéries productrices de  $\beta$ -lactamases et donc d'élargir le spectre, certaines pénicillines sont associées à des inhibiteurs de  $\beta$ -lactamases, tels que l'acide clavulanique et le tazobactam. La pénicilline G et les pénicillines anti-pseudomonales s'administrent par voie parentérale. Les autres pénicillines peuvent être administrées par voie orale. De manière générale les pénicillines sont éliminées par voie rénale, et présentent une demi-vie courte. Le temps pendant lequel la concentration sérique se situe au-delà de la Concentration Minimale Inhibitrice constituant le paramètre d'efficacité essentiel, les pénicillines nécessitent une administration fréquente. L'effet secondaire principal des pénicillines, et le plus redouté, est l'allergie de type immédiat médiée par les IgE.

## II-2-2- Les céphalosporines :

Les céphalosporines, dont un certain nombre se trouve dans le tableau 18 en annexe, diffèrent chimiquement des pénicillines par le remplacement du cycle thiazolidine par un cycle dihydrothiazine. Elles possèdent en effet pour structure commune l'acide 7-amino céphalosporanique, présentée sur la figure 2. Par rapport à l'acide 6-aminopénicillanique, ce noyau possède un carbone supplémentaire, mais la distance séparant la fonction carboxylique de l'amide cyclique reste conservée. Un élément important est la possibilité de substitution en C3 par des groupements électrocapteurs. Ceux-ci permettront une meilleure délocalisation des électrons au niveau du cycle bêta-lactame, rendant en principe les céphalosporines plus actives vis-à-vis des transpeptidases en comparaison des pénicillines. Ceci ne se traduit cependant pas toujours par un avantage clinique.

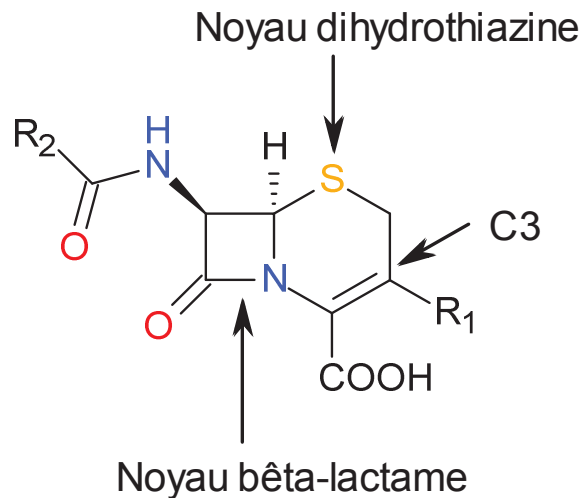


Figure 2 : Structure des céphalosporines

Grand nombre de céphèmes orales sont administrées sous forme de prodrogue afin d'augmenter leur résorption. Les céphalosporines sont habituellement classées en fonction de leur spectre d'activité, leur résistance aux  $\beta$ -lactamases, ainsi que de leur date de commercialisation. Ainsi, on distingue les céphalosporines de première, deuxième, troisième et quatrième génération :

- Céphalosporines de première génération : Elles présentent un spectre antibactérien commun avec celui des pénicillines M et A et sont surtout utilisées contre les cocci Gram positif, à l'exception des entérocoques et de certains staphylocoques résistants. Citons notamment le céfaclor, la céfadoxil, la céfazoline, ou la céfatrizine qui constituent un premier choix dans la prophylaxie chirurgicale.
- Céphalosporines de deuxième génération comprenant la céfuroxime, le céfamandole et la céfoxitine et caractérisées par une meilleure résistance aux  $\beta$ -lactamases et un spectre d'action plus large (*Haemophilus influenzae* et certaines enterobactéries), une activité à faible concentration ainsi qu'une bonne diffusion tissulaire. Elles sont utilisées dans un grand nombre d'infections, notamment respiratoires, urinaires et ostéoarticulaires.
- Céphalosporines de troisième génération telles que le céfotaxime, la ceftriaxone, la ceftazidime. Elles sont moins efficaces que celles de première génération sur les germes

Gram positif, mais beaucoup plus que celles de deuxième génération sur les entérobactéries. Certaines de ces molécules sont particulièrement actives contre le bacille pyocyanique (*Pseudomonas aeruginosa*). Leur prescription est généralement réservée aux infections sévères dues à des bacilles Gram négatif. De même, elles sont utilisées avec un certain succès dans le cas des infections nosocomiales. Elles sont actives à plus faible concentration et résistent mieux aux bêta-lactamases.

- Céphalosporines de quatrième génération avec la cefpirome et la céfépime qui pourraient remplacer les céphalosporines de 3e génération pour le traitement des infections à germes résistants. Elles possèdent un spectre large, leur activité est améliorée sur les germes Gram positifs et sont plus stables face aux bêta-lactamases.

De nombreuses céphalosporines ne sont pas résorbées par voie orale et doivent donc être administrées par voie parentérale. Comme pour les pénicillines, les céphalosporines présentent une courte demi-vie (à l'exception de la ceftriaxone), et doivent être administrées de façon régulière (Wise, 1990). Ceci est d'autant plus important que le principal paramètre d'efficacité est le temps au delà duquel la concentration sérique est maintenue supérieure à la Concentration Minimale d'Inhibition.

### II-2-3- Les carbapénèmes :

Dans les années 70, des molécules apparentées à l'acide olivanique et comportant un noyau carbapénème, ont été isolées de cultures d'actinomycètes. Les carbapénèmes dérivent de la thiénamycine qui présente certaines différences structurales notables par rapport aux antibiotiques plus anciens. Ainsi, elles possèdent une chaîne hydroxyéthyle en position C-6 du cycle  $\beta$ -lactame à la place de la chaîne amino-acyle. Le deuxième cycle contient une double liaison et est dépourvu de soufre, contrairement aux pénicillines et aux céphalosporines. Les substituants du noyau  $\beta$ -lactame sont donc des chaînes alkyles. De plus, ces substituants sont en relation stéréochimique trans alors que le  $\beta$ -lactame des antibiotiques classiques est caractérisé par une substitution cis. Ces aspects structuraux se retrouvent dans tous les carbapénèmes. La figure 3 ci-dessous représente la structure commune des carbapénèmes.

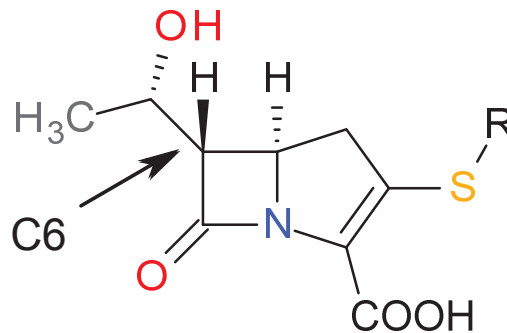


Figure 3 : Structure des carbapénèmes

Les carbapénèmes ( tableau 18, annexe) notamment représentées par l'imipénème et le méropénème (Fish *et al.*, 1997) possèdent des CMI très basses vis-à-vis de la majorité des Gram (+), des Gram (-) et des anaérobies, y compris la quasi totalité des souches productrices de bêta-lactamases (Bradley *et al.*, 1999). Ceci leur permet d'agir efficacement dans les infections causées par des germes résistants à toutes les autres bêta-lactames. Il est donc raisonnable de réserver leur usage pour le traitement d'infections nosocomiales (surtout polymicrobiennes) sévères (infections urinaires, respiratoires, gynécologiques, des articulations et des os). Notons que l'imipénème est souvent associé à la cilastatine qui en inhibe le catabolisme rénal très rapide (dégradation en quelques minutes par la déhydropeptidase I épithéliale des tubules contournés proximaux) et permet ainsi de maintenir un taux antibactérien suffisant dans l'urine. L'imipénème possède le plus large spectre de toutes les bêta-lactames.

## II-2-4- Les monobactames :

Les monobactames sont des  $\beta$ -lactames monocycliques, initialement découvertes en 1977 dans des surnageants de cultures de bactéries non pathogènes (*Acetobacter*, *Gluconobacter*, *Pseudomonas*), à la différence des autres  $\beta$ -lactames qui ont elles été isolées dans des cultures de levures. Les monobactames comportent un unique cycle : le noyau azétidine substitué par une fonction sulfonate ( $\text{SO}_3^-$ ). Leur structure monocyclique les différencie des pénicillines et des céphalosporines présentant une structure bicyclique. Toutefois, le groupement sulfonate des monobactames possède une orientation dans l'espace équivalente à celle du carboxylate des

autres bêta-lactames, ce qui explique un mode d'action semblable. Les composés naturels présentant une faible activité, des monobactames de synthèse ont été mis au point en greffant diverses chaînes latérales au noyau 3-amino-monobactamique. La figure 4 présente la structure des monobactames, le tableau 17 donne la structure de l'aztréonam :

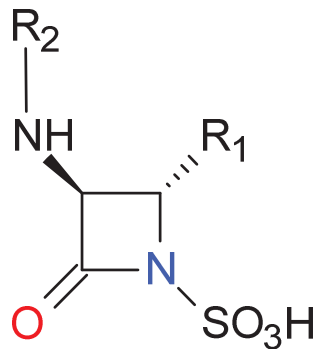


Figure 4 : Structure des monobactames

Les monobactames sont inactifs sur les bactéries Gram (+) et les anaérobies. Par contre, ils sont très actifs sur les entérobactéries et sur *Pseudomonas aeruginosa*. Le groupement amino-2-thiazol confère à l'aztréonam, principal antibiotique de ce groupe, son action sur les bactéries à Gram négatif tandis que la chaîne latérale protège le médicament contre l'inactivation enzymatique. L'activité anti-Gram (-) de l'aztréonam est globalement comparable à celle des céphalosporines de troisième génération et notamment à celle de la ceftazidime. Actuellement, seul l'aztréonam, produit par synthèse totale afin d'obtenir une molécule dotée à la fois d'un spectre élargi vers les Gram (-) difficiles et d'une bonne résistance aux  $\beta$ -lactamases, est employé (Hellinger *et al.*, 1999). Notons cependant l'existence de  $\beta$ -lactamases plasmidiques à large spectre capables d'inactiver l'aztréonam. Ce monobactame constitue une alternative de choix pour le traitement des infections graves aux bactéries Gram (-) chez les patients allergiques aux autres types de  $\beta$ -lactames ou ne pouvant recevoir d'aminoglycosides (par exemple, les insuffisants rénaux). L'aztréonam présente une bonne sécurité d'emploi. En principe, il n'y a pas de résistance croisée avec les autres  $\beta$ -lactames. L'élimination de ce médicament étant rénale, il faut ajuster les posologies en cas d'insuffisance rénale. Son administration se fait par voie parentérale.

## II-2-5- Les inhibiteurs de bêta-lactamases :

Rappelons que plusieurs bactéries se défendent contre les pénicillines à l'aide d'enzymes appelées bêta-lactamases. Ces enzymes coupent le cycle  $\beta$ -lactame, partie cruciale de la molécule, rendant l'antibiotique inactif. Ainsi, les inhibiteurs de  $\beta$ -lactamases sont employés afin de préserver l'activité des bêta-lactamines auxquelles ils sont associés. Ces dérivés de l'acide clavulanique et de l'acide pénicillanique ne possèdent qu'une très faible activité anti-bactérienne intrinsèque mais jouent le rôle de substrat-suicide capable de se lier de manière irréversible à la  $\beta$ -lactamase, empêchant ainsi son action ultérieure sur les antibiotiques. Notons que ces molécules présentent des similitudes avec les pénicillines afin d'être reconnues par les enzymes bactériennes. La figure ci-après montre la structure générale de ces inhibiteurs.

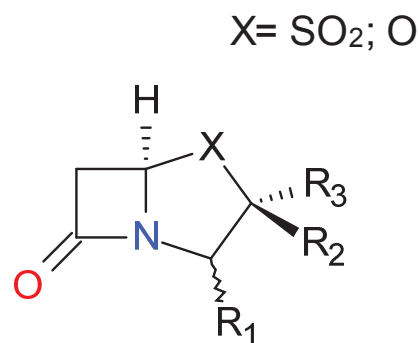


Figure 5 : Structure des inhibiteurs de bêta-lactamases

Citons trois inhibiteurs de  $\beta$ -lactamases : l'acide clavulanique, le tazobactam et le sulbactam (tableau 18 en annexe).

- L'acide clavulanique est produit par *Streptomyces clavuligerus* et dérive du noyau clavame qui se distingue du noyau péname du fait d'un atome d'oxygène qui remplace celui de soufre. Il sera associé à l'amoxicilline (utilisation thérapeutique très large) ou la ticarcilline.

- Le sulbactam qui présente un noyau pénème possède une activité intrinsèque antibiotique sur quelques germes mais ne sera employé qu'en association avec une pénicilline. Il existe un ester de sulbactam-ampicilline dont l'hydrolyse dans l'organisme libérera le sulbactam et la pénicilline.
- Le tazobactam possède lui aussi un noyau pénème et dérive de l'acide pénicillanique. Il présente une activité proche de l'acide clavulanique et du sulbactam et sera associé à la pipéracilline.

## II-3- Mécanisme d'action des antibiotiques de type bêta-lactamine :

Les  $\beta$ -lactamines sont des antibiotiques inhibiteurs de la synthèse de la paroi bactérienne, que ce soit des Gram positif ou négatif. Ces composés n'exercent de ce fait leur action bactéricide que si les bactéries sont en phase de croissance active et en cours de synthèse de leur paroi. Ces antibiotiques possèdent une activité temps-dépendant ce qui signifie que leur action bactéricide dépend de la durée pendant laquelle les concentrations en médicament demeurent supérieures à la concentration minimale bactéricide ou Concentration Minimale Inhibitrice.

### II-3-1- Principe :

Le mécanisme d'action des  $\beta$ -lactamines est fondé sur la liaison de l'antibiotique aux enzymes participant à la synthèse de la paroi bactérienne, les PFP (Protéines Fixant la Pénicilline) encore appelées PBP (Penicillin Binding Proteins). La conséquence de cette interaction moléculaire est l'inhibition de la synthèse et du remodelage du peptidoglycane, composant essentiel de la paroi bactérienne. Ainsi, les transpeptidases bactériennes, PFP intervenant dans la dernière étape de la synthèse du peptidoglycane, vont reconnaître le cycle bêta-lactame du fait de son analogie structurale avec leur substrat naturel. En effet, les  $\beta$ -lactamines sont des analogues structuraux du substrat naturel dipeptidique D-ala D-ala qui termine les chaînes polypeptidiques des précurseurs du peptidoglycane et vont donc se lier avec les enzymes responsables de la réticulation du peptidoglycane : les transpeptidases intervenant au moment de l'élongation (assemblage des chaînes peptidiques par formation de ponts interpeptidiques). Notons dès à

présent que la liaison à une PFP se fait de façon covalente et irréversible. Les bêta-lactamines sont donc des « substrats suicide ». Ainsi, après que l'antibiotique s'est fixé à la PFP, la réaction de transpeptidation est inhibée, la synthèse du peptidoglycane bloquée et la cellule meurt rapidement. Toutefois, l'activité bactéricide des pénicillines n'est pas liée à la rupture de la paroi fragilisée sous l'effet de la pression osmotique car pour qu'il y ait mort cellulaire, il faut la coopération active de la bactérie qui utilise sa propre machinerie enzymatique suicide : les peptidoglycanes-hydrolases (autolysines).

### II-3-2- La paroi bactérienne :

La paroi est un polymère complexe entrecroisé, le peptidoglycane, constitué de polypeptides et de polysaccharides. Le peptidoglycane confère à la bactérie sa forme et sa rigidité ce qui lui permet de résister à la forte pression osmotique intra-cytoplasmique. Il contient une alternance de sucres aminés, de la N-acétylglucosamine et de l'acide N-acétylmuraminique, qui forment un réseau entourant complètement la bactérie. Il est renforcé par de courtes chaînes d'acides aminés fixées sur des molécules d'acide N-acétylmuraminique. Les chaînes d'acides aminés forment également des ponts reliant ensemble les couches successives. La composition en acides aminés de ces chaînes varie suivant le type de bactérie. La synthèse du peptidoglycane s'effectue en trois stades et met en jeu plus de 30 étapes enzymatiques :

- Dans une première phase, se produisent les événements cytoplasmiques qui conduisent à la synthèse de la N-acétylglucosamine et de l'acide N-acétylmuraminique lié à une unité pentapeptidique comprenant le dipeptide D-alanyl-D-alanine.
- La deuxième phase concerne les étapes membranaires qui correspondent au transport des précurseurs au travers de la membrane cytoplasmique, grâce à un transporteur lipidique.
- Enfin, la troisième phase s'effectue en deux étapes et correspond à l'incorporation des précurseurs dans la paroi en formation.

Le peptidoglycane de la bactérie est en perpétuel remaniement car il résulte d'un équilibre dynamique fragile entre activité de synthèse (PFP) et d'hydrolyse à bas niveau (autolysines) et



que toute rupture de cet équilibre, par inhibition des PFP, puis activation des autolysines, est néfaste à la bactérie.

### II-3-3- Déroutement :

#### II-3-3-1- Pénétration dans la bactérie :

Les  $\beta$ -lactames ne doivent pas pénétrer à proprement parler dans la bactérie, mais rejoindre leur cible au niveau de la face interne de la paroi (espace périplasmique). Cet accès est direct et facile pour les bactéries à Gram positif dont le peptidoglycane est relativement perméable aux bêta-lactamines (sauf celui de *E. faecalis*). Par contre, concernant les Gram(-), les  $\beta$ -lactames doivent emprunter les passages ménagés par les porines au niveau de la membrane externe pour rejoindre l'espace périplasmique. Ainsi, chez les bactéries à Gram négatif, ces propriétés de perméabilité vont contribuer à définir le phénotype naturel de résistance propre à chaque espèce bactérienne.

#### II-3-3-2- Activité antibactérienne :

Comme nous l'avons vu, elle résulte de l'inhibition de la synthèse du peptidoglycane par fixation covalente aux protéines liant la pénicilline. Ces dernières sont des enzymes impliquées dans la synthèse du peptidoglycane et sont des transpeptidases et des carboxypeptidases. Elles sont ancrées dans la membrane cytoplasmique par leur extrémité carboxylique, et émergent sur la face externe de cette membrane dans l'espace périplasmique par leur extrémité  $\text{NH}_2$  porteuse du site actif enzymatique.

- Arrêt de la croissance bactérienne :

En mimant la structure tridimensionnelle de la séquence D-Ala-D-Ala, les bêta-lactames se comportent comme des inhibiteurs de la transpeptidase. En présence d'antibiotique, cette enzyme, qui agit comme une acyl-transférase à sérine active, exerce son action hydrolytique sur la molécule antibiotique, conduisant à la formation d'un complexe enzyme-produit qui ne se dissocie pas car l'enzyme est liée de façon covalente au produit. Ce mécanisme est appelé

inhibition-suicide (l'enzyme elle-même catalyse la transformation de l'antibiotique en un composé hautement réactif qui se lie à l'enzyme de façon irréversible).

Les résultats qui seront ici présentés mettront en évidence le fait que la fixation de la bêta-lactamine sur la PFP peut se faire sur d'autres types d'acides aminés que la sérine, et notamment la lysine.

Les schémas suivants exposent l'action de la transpeptidase et le principe antibactérien des  $\beta$ -lactamines (Université catholique de Louvain, Cours en ligne, Pharmacologie et pharmacothérapie des anti-infectieux)

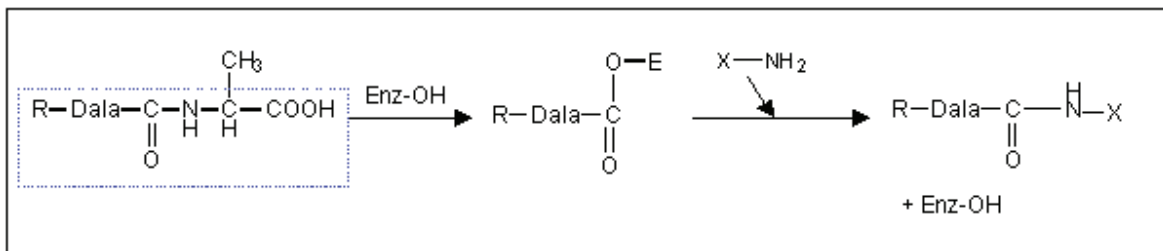


Schéma 1a : Réaction de transpeptidation

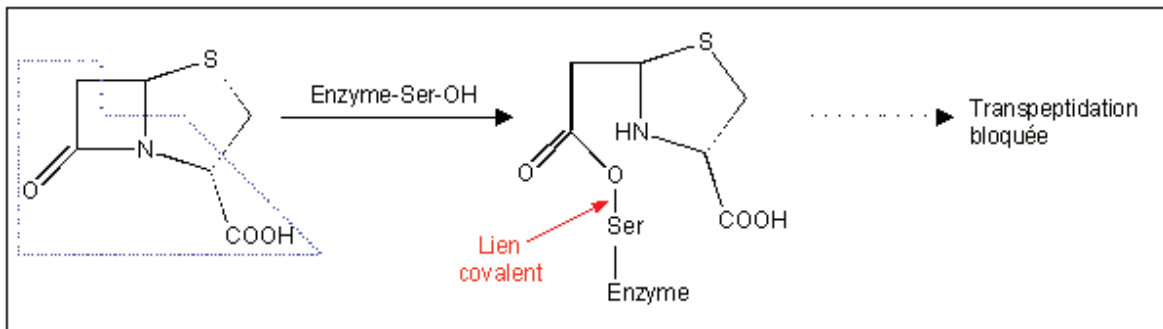


Schéma 1b : Inhibition de la transpeptidation par la bêta-lactame

Notons en pointillé l'analogie structurelle entre la séquence D-Ala-D-Ala et la  $\beta$ -lactamine.

L'inhibition des PFP essentielles entraîne un arrêt de croissance (effet bactériostatique). Trois effets morphologiques différents peuvent être observés selon les fixations préférentielles

d'une bêta-lactamine sur la PFP : la filamentation, la formation de sphéropastes suivie d'une lyse bactérienne, et la formation d'éléments ovoïdes.

Concernant les inhibiteurs de bêta-lactamases, elles n'ont pas de rôle contre les bactéries mais servent de substrats pour les bêta-lactamases qui ne pourront plus exercer leur effet sur les antibiotiques.

- Effet bactéricide :

Il fait suite à l'arrêt de la croissance de la paroi bactérienne et met en jeu le système des autolysines, qui dégradent le peptidoglycane des bactéries en phase de croissance. Ainsi la rupture de l'équilibre dynamique entre la construction du peptidoglycane et son réarrangement par les autolysines va se faire en faveur des autolysines entraînant la mort de la bactérie.

#### II-3-4- Résistance aux bêta-lactamines :

##### II-3-4-1- Imperméabilisation :

Une imperméabilisation par altération des porines (liée à une mutation chromosomique) se rencontre chez les Gram (-) et confère une résistance aux  $\beta$ -lactames trop hydrophiles pour diffuser à travers la membrane externe (céphalosporines, pénicillines à large spectre).

##### II-3-4-2- Altération des PBP :

Ce phénomène concerne surtout les bactéries Gram (+). Une altération des PBP peut réduire leur affinité pour les  $\beta$ -lactames (Spratt, 1994). Ce type de mécanisme, développé au niveau de la PBP2, se retrouve dans les souches de staphylocoques dites "résistantes à la méthicilline". Cette altération entraîne une résistance à l'ensemble de la classe des  $\beta$ -lactames et a donc une grande importance clinique.

##### II-3-4-3- Bêta-lactamases :

La production par les bactéries d'enzymes hydrolysant l'antibiotique constitue le mécanisme de résistance le plus courant. Ces enzymes sont secrétées au dehors dans le cas des

Gram (+), mais maintenues dans l'espace périplasmique dans le cas des Gram (-). Il s'agit de protéases à sérine active, qui se lient aux  $\beta$ -lactames avec plus d'affinité que les PBP. Comme celles-ci, elles hydrolysent la liaison amide du cycle  $\beta$ -lactame, pour former une acyl-enzyme, mais la différence majeure entre les 2 types d'enzymes concerne la vitesse avec laquelle l'acyl-enzyme est hydrolysée. En effet, si les PBP ne sont capables d'hydrolyser qu'une  $\beta$ -lactame par heure, de telle sorte que l'acyl-enzyme apparaît dans ce cas comme un intermédiaire stable, les  $\beta$ -lactamases les plus efficaces en hydrolysent 1000 par seconde, rendant l'antibiotique totalement inactif et régénérant l'enzyme pour une nouvelle réaction d'hydrolyse.

Les schémas 2a-b font le parallèle entre l'inhibition de la transpeptidase et celle de l'antibiotique (Université catholique de Louvain, Cours en ligne, Pharmacologie et pharmacothérapie des anti-infectieux)

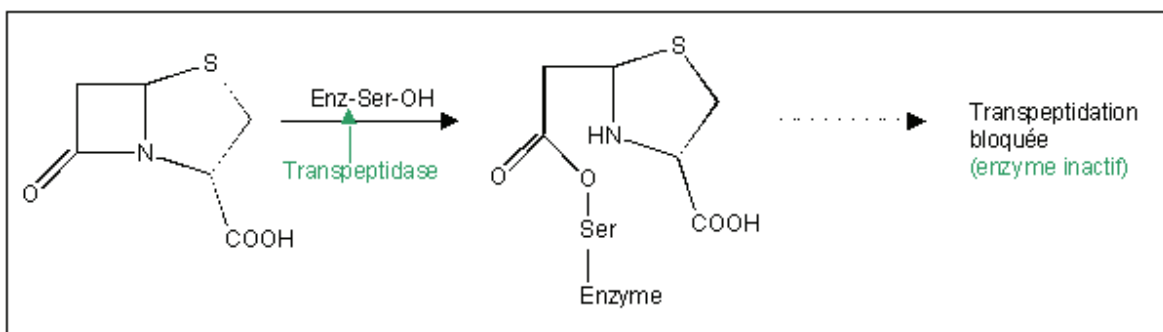


Schéma 2a : Inactivation de la transpeptidation

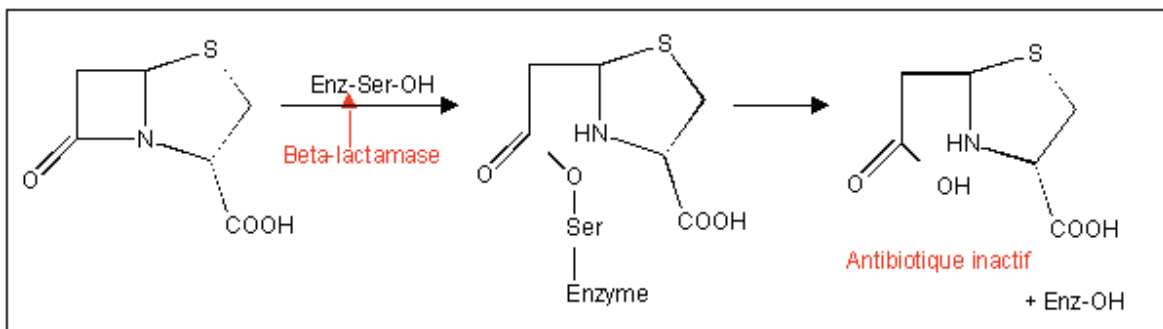


Schéma 2b : Inactivation de la  $\beta$ -lactame par une bêta-lactamine

L'importance du mode de résistance associé à la production de  $\beta$ -lactamases a conduit à la synthèse de molécules résistantes à ces enzymes. Les  $\beta$ -lactamases, cependant, se caractérisent par une plasticité extraordinaire, de telle sorte qu'elles finissent toujours par réussir à hydrolyser

les nouveaux antibiotiques. Une autre stratégie consiste dès lors à mettre au point des inhibiteurs de  $\beta$ -lactamases (sulbactam, tazobactam, acide clavulanique). Ces molécules, de structures voisines aux  $\beta$ -lactames, lient de façon irréversible la  $\beta$ -lactamase, empêchant son action ultérieure sur l'antibiotique coadministré.

## **III- Hypersensibilités allergiques médicamenteuses :**

### **III-1- Hypersensibilités allergiques et non allergiques :**

Très souvent, la distinction entre l'hypersensibilité allergique et l'hypersensibilité non allergique est complexe car les manifestations cliniques peuvent être extrêmement proches. Toutefois, ce que l'on appelle aussi, « pseudo-allergie », « intolérance médicamenteuse » ou encore « fausse allergie » diffère fondamentalement des « vraies allergies » puisque dans une hypersensibilité non allergique, aucun mécanisme immunologique n'est mis en jeu.

Les Hypersensibilités Allergiques font appel à des phénomènes immunologiques avec implication de lymphocytes T et d'anticorps spécifiques du médicament responsable des manifestations cliniques. De nombreux médicaments peuvent induire des réactions d'allergies mais les anti-inflammatoires non stéroïdiens (dont l'aspirine) et les bêta-lactamines sont les plus souvent impliqués.

Ces réactions, que l'on peut définir comme un phénomène immunitaire démesuré et inapproprié, se déroulent toujours en deux phases. La première correspond à la sensibilisation du patient vis-à-vis du médicament. On observe alors la mise en place d'une immunité adaptative avec production d'anticorps et de lymphocytes T spécifiquement dirigés contre le médicament. La seconde phase arrive lors d'un nouveau contact avec le même médicament. Le patient développera alors la réaction d'hypersensibilité allergique marquée par des manifestations cliniques diverses.

La phase de sensibilisation est modulée par de nombreux facteurs tels que la structure chimique du médicament, sa voie d'administration mais aussi des pathologies intercurrentes et des facteurs génétiques (David, 1996)

En 1945, Gell et Coombs ont classé les HSA en quatre types qui rendent compte de la symptomatologie présentée ainsi que des mécanismes immunologiques impliqués.

On distingue les **HSA de type I** liées aux immunoglobulines de type E, de **type II** liées aux immunoglobulines de type G, de **type III** impliquant les complexes immuns et de **type IV** liées aux lymphocytes T CD4+ et CD8+ cytotoxiques.

En 2001, l'Association Européenne d'Allergologie et Immunologie Cliniques (EAACI) propose de classer ces HSA en fonction de l'implication ou non des IgE. Ainsi, on distinguera les réactions d'hypersensibilité allergique IgE médiées et les réactions d'hypersensibilité allergique non IgE médiées. (Pouyanne, Haramburu et al., 2000).

Les HSA de type 1 (HSA Immédiate) et 4 (HSA Retardée) de la classification de Gell et Coombs sont les plus souvent observées. Dans le premier cas, les symptômes apparaissent dans les minutes qui suivent la prise médicamenteuse, alors que dans le second cas, ils apparaissent plusieurs heures après la prise.

Les médicaments sont susceptibles d'induire les quatre types d'hypersensibilité et une même molécule pourra être impliquée dans plusieurs de ces mécanismes.

Différents moyens permettent la mise en évidence de ces hypersensibilités et la différenciation entre phénomènes allergiques et non allergiques s'appuiera sur des tests cliniques et biologiques. A ce jour, la mise en évidence des IgE et des LT spécifiques du médicament fait appel d'une part à des tests cutanés à lecture précoce (nommés prick-tests et intra-dermo-réactions) et à lecture tardive (intra-dermo-réactions et patch-tests). Sur un plan biologique, il est possible d'analyser les HSA Immédiates grâce aux tests de dégranulation ou d'activation des basophiles (TAB).

Notons que pour des patients souffrant d'hypersensibilité non allergique, les tests cutanés seront négatifs.

Malgré les études menées dans le but d'améliorer la capacité des tests à diagnostiquer formellement les réactions allergiques, il faut rappeler que l'imputabilité du médicament n'est pas toujours aisée et qu'il est donc extrêmement important d'enrichir et perfectionner ces tests diagnostiques.

## III-2- Immunogénicité du médicament :

Pour qu'il y ait mise en jeu du système immunitaire spécifique, il faut que l'organisme reconnaisse le médicament, appelé allergène, comme étant un antigène (capable d'être spécifiquement reconnu par les anticorps). Le médicament doit également être immunogène, c'est-à-dire capable d'induire une réaction immunitaire. Tout antigène n'est pas forcément immunogène. Ceci est particulièrement vrai pour les médicaments.

### III-2-1- Haptène, Prohaptène, P-i concept :

Les médicaments de nature protéique et ceux possédant un poids moléculaire élevé, telles que les hormones, pourront se comporter comme des antigènes complets.

Pour les substances de faible poids moléculaire (inférieur à 1000 daltons), ce qui est le cas de la majorité des médicaments, elles doivent se lier de façon covalente à une protéine endogène porteuse, encore appelée protéine d'hapténisation, pour induire une réponse de type B ou T (Pichler et al. 1998, Britschgi et al. 2001, Pichler 2002, Lu et al. 2003)

Un **haptène** est un composé chimique réactif dans les conditions physiologiques. C'est le cas des antibiotiques de la famille des bêta-lactamines qui se fixent de manière covalente et irréversible aux acides aminés de protéines porteuses, en particulier sur des lysines et des sérines. (Padovan et al., 1997). Le complexe protéine-médicament sera immunogène et donc reconnaissable par le système immunitaire. Il sera capté par des cellules présentatrices qui vont effectuer un apprêtement intracellulaire au cours duquel la protéine sera dégradée en peptides qui seront ensuite présentés au sein de la niche à peptide du CMH. Le peptide qui a fixé l'antigène médicamenteux sera reconnu par les lymphocytes T et B. Ceci explique que la réponse immunologique pourra présenter de nombreuses formes cliniques. (Pichler et al., 2006). Il existe

des haptènes dits forts car capables d'induire une sensibilisation chez une majorité d'individus et des haptènes faibles qui ne sensibilisent que peu de patients. Notons que la plupart des haptènes sont faibles et nécessitent une exposition prolongée avant de sensibiliser la personne, contrairement aux haptènes forts (tel que le DiNitroFluoroBenzène) pour lesquels une seule exposition peut suffire.

On parlera d'**un pro-haptène** si l'antigène nécessite une phase de métabolisation via des enzymes avant de se fixer covalamment aux protéines cibles. L'industrie pharmaceutique s'est penchée sur la conception de molécules chimiquement inertes afin d'éviter que les médicaments se comportent comme des haptènes. Toutefois, dans certains cas, la métabolisation du produit va aboutir à des dérivés qui seront eux réactifs et pourront se fixer aux protéines d'hapténisation. (Park et al. 1998, Naisbitt et al. 2000, Pichler 2002). Le sulfaméthoxazole est un très bon exemple tandis que les bêta-lactamines ne seront jamais métabolisées et donc ne seront pas des pro-haptènes.

On parlera de **pré-haptène** si le composé chimique subit avant sa fixation à une protéine, des modifications sans l'intervention d'enzymes (Lepoittevin 2006) mais liées à différents facteurs tels que la chaleur, la lumière, l'oxydation. Un bon exemple est la paraphénylènediamine.

D'autre part, l'équipe de Werner Pichler (Pichler et al., 2006) a également étudié la possibilité que certains médicaments ne nécessitent pas de liaison covalente au peptide présent dans la niche à peptide du CMH pour induire une réponse immunitaire. Ainsi, ces médicaments interagiraient de manière directe et réversible (non covalente) avec le CMH et le TCR, entraînant une stimulation des ces cellules selon un mode comparable à l'activation pharmacologique classique d'autres récepteurs de l'organisme. Dans ce cas, ni le métabolisme, ni l'apprêtement par une CPA n'interviennent. On parle alors du Pi-concept, pour Pharmacological interaction with the Immune receptor. On peut comparer ce mécanisme d'interaction non covalente entre l'haptène, le CMH et le LT à l'activation des LT par des superantigènes ou le nickel (Lu et al. 2003). Ce phénomène a été observé notamment pour la lidocaïne, le sulfaméthoxazole, la carbamazépine.



### III-2-2- Les épitopes :

Une molécule antigénique sera reconnue par le système immunitaire au niveau de ce que l'on appelle l'épitope ou déterminant antigénique. Il correspond à la partie de l'antigène qui interagit spécifiquement avec le paratope d'un anticorps ou d'un lymphocyte B (au niveau de son récepteur : BCR) ou T (au niveau de son récepteur : TCR). Un antigène présentera ses épitopes aux cellules et structures immunitaires et sera reconnu comme appartenant au non-soi. De ce fait, l'interaction entre épitope et paratope est la base d'une réponse immunitaire adaptative et spécifique.

L'épitope peut impliquer les quatre niveaux de structure d'une protéine (primaire : séquence en acides aminés, secondaire : repliement en hélices alpha ou brins bêta, tertiaire : agencement en domaines et quaternaire : polymérisation). Il pourra être séquentiel, c'est-à-dire linéaire, ou conformationnel donc discontinu et formé de la juxtaposition dans l'espace de zones de la molécule antigénique.

Généralement les épitopes sont constitués de 8 à 25 acides aminés dans le cas d'une protéine, et de 5 à 6 oses dans le cas d'un polysaccharide. On distingue les épitopes B et les épitopes T.

Les **épitopes B** sont exposés à la surface de tous les antigènes et interagissent directement avec les anticorps ou les récepteurs des lymphocytes B. Ils peuvent être séquentiels ou discontinus et sont retrouvés sur les protéines, les lipides et les polysaccharides. L'anticorps interagira de manière non covalente avec l'épitope B (ce qui implique une accessibilité de l'épitope, une proximité directe et une complémentarité) au niveau de deux sites de liaisons (nommés Fragment antigen binding) qui sont hypervariables. Leur taille varie de 8 à 20 acides aminés.

Les **épitopes T** sont reconnus par un lymphocyte T et impliquent un complexe à trois partenaires : l'antigène, le CMH et le TCR. L'antigène doit donc posséder un site de liaison pour le TCR (il s'agit de l'épitope T) et un pour la molécule de CMH. Notons que tous les antigènes ne présentent pas forcément un épitope T. Lorsque l'antigène est pris en charge par une cellule présentatrice de l'antigène (CPA), il sera dégradé puis présenté au sein de la niche à peptides du

CMH aux récepteurs TCR. Ces épitopes de nature principalement protéique (mais aussi parfois lipidique) sont séquentiels. S'ils sont présentés par les molécules CMH de classe I, leur taille sera de 9 acides aminés, tandis que les molécules de classe II interagissent avec des séquences de 11 à 25 acides aminés.

Lorsque deux antigènes possèdent des épitopes similaires, ils peuvent être reconnus par le même récepteur spécifique (TCR ou BCR) et dans ce cas, on parlera de réactivité croisée. Ces antigènes peuvent appartenir à la même famille chimique (ce sera le cas des réactions croisées entre différentes bêta-lactamines) ou non (réactions croisées entre les allergènes de la pomme et ceux du bouleau par exemple).

Au cours de ce travail, nous avons particulièrement étudié les phénomènes de réactions croisées, fréquemment retrouvées pour les bêta-lactamines. Notons que le noyau bêta-lactame autant que les chaînes latérales peuvent jouer le rôle d'épitope.

### III-3- Initiation et déroulement d'une HSA :

Le système immunitaire maintient l'intégrité de notre organisme. Pour cela il met en œuvre de nombreuses réactions qui ont pour but de protéger nos cellules. On distingue l'immunité non spécifique, qui permet une action rapide et immédiate et fait intervenir des cellules responsables de la phagocytose, et l'immunité spécifique à médiation humorale (due aux anticorps ou immunoglobulines) et à médiation cellulaire (due aux lymphocytes T), qui nécessite un délai de quelques jours pour se mettre en place. Après l'élimination de la substance étrangère, notre organisme garde en mémoire les cellules nécessaires pour éliminer ce même antigène. Immunité non spécifique et spécifique fonctionnent en synergie : pour qu'un antigène soit pris en compte par le système immunitaire spécifique, il doit lui être présenté par une cellule du système immunitaire non spécifique. Cette cellule devient alors une cellule présentatrice d'antigène: CPA.

Dans un premier temps, l'antigène sera pris en charge par une CPA qui, après digestion et association aux molécules du CMH, présente les débris de l'antigène aux lymphocytes. La stimulation antigénique (reconnaissance spécifique entre le récepteur du LT et le déterminant antigénique) provoque la transformation lymphoblastique des lymphocytes T, ainsi que leur

multiplication et leur maturation en cellules effectrices (Lymphocytes CD8 et CD4 helper ou auxiliaires) ou cellules mémoires (CD4 mémoires). Les CD8 sont cytotoxiques et détruisent les cellules infectées. Les lymphocytes CD4 sont capables de produire diverses cytokines permettant l'activation des autres acteurs de l'immunité spécifique. La transformation lymphoblastique des LB, après reconnaissance spécifique de l'antigène, conduit à leur prolifération et leur activation en "cellules effectrices B", appelées plasmocytes. Ces derniers produisent des anticorps capables de reconnaître spécifiquement l'antigène et de le fixer, formant alors un complexe immunitaire détruit par la suite par des cellules telles que les macrophages. Lors d'une première agression par un antigène, les phagocytes qui jouent leur rôle de CPA captent l'antigène, l'internalise et le dégrade en peptides. Ceux-ci seront présentés au sein de la niche à peptides du CMH de classe II. Grâce à son TCR, un lymphocyte T CD4 spécifique de l'antigène sera recruté parmi un ensemble de LT CD4 naïfs. Le lymphocyte sera activé en lymphocyte T helper permettant une coopération avec les autres cellules effectrices. Il existe deux grandes voies possibles en fonction de différents facteurs (dont le type de cytokines produites).

La **balance Th1/Th2** est une notion qui date des années 1990. Mosmann a montré dans un modèle murin que les lymphocytes T CD4 + peuvent être divisés sur la base de leur profil de sécrétion cytokinique en deux sous populations : les Th1 et les Th2. Ces 2 types de lymphocytes T helper dérivent d'un précurseur commun qui sécrète de l'interleukine IL 2 mais ne sécrètent ni IL4 ni IFN gamma. La stimulation de ce précurseur induit sa différenciation en Th0 qui produit à la fois des cytokines de type 1 et de type 2. L'ambiance cytokinique de l'environnement va par la suite polariser la différenciation des Th0 en Th1 ou Th2. Le profil Th1 est sous le contrôle de l'IL2 et de l'IFN gamma notamment, tandis que l'IL4 permet une orientation phénotypique vers les Th2. Les deux populations de helper s'inhibent mutuellement et s'auto-amplifient en faveur de leur propre développement. La polarisation Th1/Th2 qui aboutira à un caractère de défense immunitaire globalement humoral ou globalement cellulaire, est sous la dépendance de différents facteurs dont l'environnement cytokinique : l'IFN gamma favorise la différenciation des lymphocytes T CD4 en clones Th1 et inhibe le développement de la population Th2, alors que l'IL4 favorise la différenciation des clones Th2. L'IL10 inhibe la production d'IFN gamma par les LT et favorise ainsi la réponse Th2. L'IL4 joue un rôle fondamental dans l'orientation du profil de cytokines vers le type Th2.

Dans le cas d'une **réponse immunitaire par voie humorale**, ce sont les LT helper de type 2 qui sont acteurs, on parle de la **voie Th2**, et l'élimination de l'antigène passera par la formation de complexes immuns avec des immunoglobulines spécifiques. En effet, les LTh2 produisent des cytokines de type Th2. Elle orientent la différenciation des LB, qui ont reconnu spécifiquement l'antigène grâce à leur BCR, en plasmocytes producteurs d'Ig spécifiques. Il y a donc coopération étroite entre lymphocytes B, T et CPA qui communiquent entre elles de manière directe grâce à des interactions membranaires, ou indirecte par des cytokines solubles telles que l'interleukine 1 sécrétée par les CPA et costimulatrice dans la présentation de l'antigène ou l'IL2, IL3, IL4, IL5 sécrétées par le LT helper et qui sont des facteurs d'activation, de prolifération et de différenciation des lymphocytes B.

Dans le cas d'une **réponse immunitaire par voie cellulaire (voie Th1)**, la CPA active un LT helper 1 producteur de cytokines de type 1 (cytokines pro inflammatoires : IL2, IFN gamma) qui conduira à l'éviction de l'antigène notamment par des LT CD8 à propriétés cytotoxiques et spécifiques de l'antigène mais aussi par des macrophages non spécifiques.

### III-4- Les différents mécanismes allergiques

#### III-4-1- Hypersensibilités allergiques de type I :

C'est le type le plus fréquent et le plus important du point de vue clinique. Les phénomènes qui en résultent apparaissent dans les minutes qui suivent le contact avec l'allergène. On parle également d'hypersensibilité allergique immédiate, médiée par les immunoglobulines de type E capables de se fixer sur les mastocytes tissulaires et sur les basophiles du sang circulant qui vont alors subir une dégranulation explosive à l'origine de la pathologie. Elle peut survenir chez des patients dit atopiques (c'est-à-dire présentant un terrain de prédisposition génétique aux allergies) ou sans antécédent spécifique.

Dans ce cas, l'antigène déclenche une réponse amplifiée selon un mode humoral (voie Th2).

Lors du premier contact avec l'agent responsable, les lymphocytes B spécifiques vont se multiplier, subir une transformation lymphocytaire et seront alors producteurs d'IgE spécifiques de l'antigène. Les mêmes événements que ceux impliqués dans la synthèse des autres classes d'immunoglobulines (Torres et al. 1998) contrôlent la synthèse d'IgE avec coopération entre les Cellules Présentatrices de l'Antigène, les lymphocytes T helper et les lymphocytes B. Mais, chez les sujets allergiques, on note un déséquilibre de la balance Th1/Th2, avec une élévation du taux de Th2. Le Complexe Majeur d'Histocompatibilité classe II présente, au sein de ce que l'on nomme la niche à peptides, le peptide issu de l'allergène, aux LT. L'allergie de type I est un modèle de profil Th2, caractérisé par la production d'interleukines 5, 4 et 13, fondamentales pour la synthèse d'IgE. IL4 et IL10 sont capables d'inhiber la différenciation des Th1. Dans un syndrome hyper IgE, l'injection d'IFN $\gamma$ , produit par les LTh1, provoque une diminution d'IgE (Reig Rincon et al. 2005). Les lymphocytes T et B, pour fonctionner ensemble, ont besoin du CD40 (à la surface du LT). Le CD23 (qui peut être soluble) est un récepteur de faible affinité pour l'IgE, il va avoir un effet synergique avec les IL4, IL5 et IL 13 pour orienter les LB vers la synthèse d'IgE.

Les IgE vont se fixer via leur fragment constant Fc sur les cellules de l'organisme, notamment les mastocytes et les basophiles qui présentent en surface un récepteur à haute affinité pour les IgE (Fc  $\epsilon$  RI). Lors d'un deuxième contact avec le même antigène, celui-ci va se lier sur les IgE spécifiques en surface des mastocytes et des basophiles, cellules les plus impliquées dans les phénomènes d'HSAI, dont les granulations cytoplasmiques métachromatiques sont remplies de médiateurs préformés (histamine, Tumor Necrosis Factor alpha) et néosynthétisés (cytokines proinflammatoires : IL-1, IL-3, Granulocyte Colony Stimulating Factor: GM-CSF, IL-4, IL-5, TNF, prostaglandines, leucotriènes). Les basophiles sont des cellules de type granulocytaire extrêmement activables par les allergènes : on observe une augmentation significative du nombre et du pourcentage des basophiles sanguins, particulièrement lors des périodes d'expositions aux allergènes. Il faut noter également la présence d'éosinophiles renfermant également de nombreux médiateurs capables d'entretenir ou aggraver les symptômes cliniques. Des éosinophiles hydrodenses, du fait de la perte de leurs granulations, sont observés au cours des maladies allergiques. Les plaquettes sont activées par le Platelet Activating Factor (PAF) au cours de la maladie asthmatique. Elles contiennent des facteurs vasoactifs comme la sérotonine. Les polynucléaires synthétisent de nombreux médiateurs lipidiques, comme les leucotriènes.

Suite à la liaison des IgE sur les récepteurs à haute affinité, il y aura dégranulation des mastocytes/basophiles qui libéreront leur contenu dans la circulation, provoquant ainsi des manifestations cliniques en trois temps : Une action immédiate liée à la libération de l'histamine responsable d'une vasodilatation locale, une bronchoconstriction, et d'autres manifestations (tachycardie, troubles digestifs...). Une action intermédiaire correspondant à la libération des prostaglandines et leucotriènes. Enfin, une action tardive d'auto-entretien de la réponse inflammatoire avec recrutement d'un infiltrat inflammatoire sous l'effet des cytokines. Les cellules immunitaires attirées au site d'inflammation (éosinophiles, basophiles, lymphocytes...) vont à leur tour libérer des cytokines entretenant l'inflammation et favorisant la production d'IgE.

Sur le plan physiopathologique, l'hypersensibilité de type I se déroule en deux phases :

- La phase de sensibilisation : l'exposition à l'allergène induit la synthèse d'IgE spécifiques de cet allergène. Les IgE se fixent sur les mastocytes et les basophiles et possèdent un tropisme à très forte affinité pour ces cellules. Il n'y a aucun symptôme clinique lors de cette phase.
- La phase de déclenchement : une réintroduction de l'allergène entraîne le pontage des IgE membranaires qui se traduit par une activation cellulaire et la libération de médiateurs responsables des manifestations allergiques qui peuvent aller jusqu'au choc anaphylactique.

Après le pontage de deux IgE, l'allergène recrute une série de protéines tyrosine-kinase (cascade de phosphorylation, déphosphorylation) qui vont induire une activation de la phospholipase C produisant de l'IP3 (inositol tris phosphate) et du DAG (diacylglycérol) d'où une augmentation du  $Ca^{2+}$  intracellulaire (au bout de 2 minutes), une mise en action des tubules de la cellule puis une dégranulation du mastocyte et du basophile (Sanz et al. 2003, Sanz et al. 2005). La figure 6 ci-dessous explique le pontage de deux IgE :

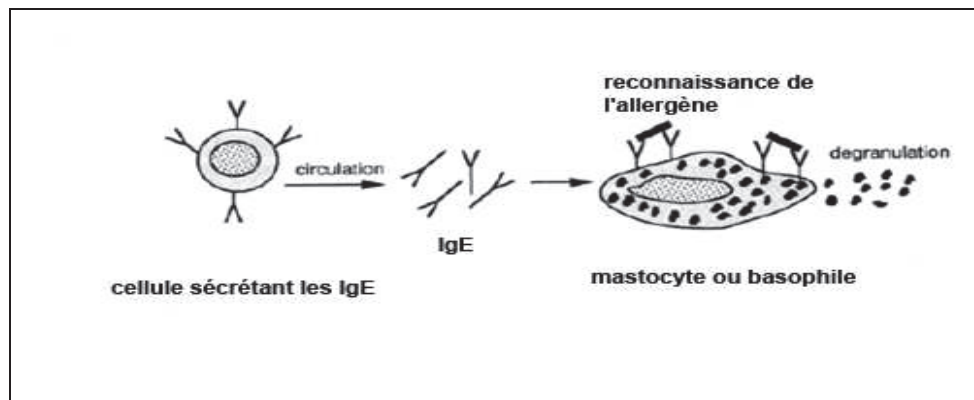


Figure 6 : Physiopathologie de l'HSAI

Sur le plan clinique, les symptômes sont généralement localisés au niveau du site d'activation des mastocytes (manifestations oculaires avec uvéite et conjonctivite, respiratoires avec asthme, toux, rhinite, cutanées avec dermatite atopique, urticaire et œdème de Quincke (Baldo et al. 2000)., digestives avec vomissements, diarrhées et douleurs abdominales), mais, en cas d'introduction de l'allergène par voie systémique, on peut observer jusqu'au choc anaphylactique, avec chute brutale de la tension et pouvant entraîner la mort du patient allergique. Les chocs anaphylactiques impliquent plus souvent les pénicillines que les céphalosporines responsables elles de réactions cutanées principalement (Lin, 1992). Ils résultent de la libération rapide et massive des médiateurs de l'HSAI par les mastocytes et basophiles. Ils peuvent aller de l'urticaire aiguë généralisée ou locale à la mort brutale en moins de cinq minutes. Les symptômes du choc, parfois annoncé par des prodromes tel qu'un malaise général, sont plus fréquemment cutanéomuqueux. Le plus souvent, on aura un rush urticarien généralisé, dans certains cas un angio-œdème qui pourra diffuser jusqu'au larynx (on parle de l'œdème de Quincke potentiellement mortel). Les signes seront fréquemment également respiratoires avec des crises d'asthme. Une vasodilatation généralisée signera une atteinte circulatoire avec tachycardie instabilité de la tension, stigmates de syncope...

Il faut noter que la plupart des médicaments sont susceptibles d'induire ce type d'hypersensibilité.

### III-4-2- Hypersensibilités allergiques de type II :

Ces réactions encore appelées cytotoxiques mettent en jeu les immunoglobulines de type G produites au cours de la phase de sensibilisation orientée sur la voie humorale. Elle démarre 4 à 6 h après le contact avec l'allergène. Les antigènes exogènes se fixent sur les cellules sanguines circulantes (hématies, plaquettes) pour former un complexe anticorps/antigène qui conduira à la destruction de la cellule porteuse via le système du complément avec ou non intervention des cellules Natural Killer.

Le plus souvent, ces réactions aboutissent à des cytopénies : anémies hémolytiques, thrombopénies, leucopénies, néphrites interstitielles (Benveniste 1986, Lachmann 1975).

Les antigènes impliqués pourront être endogènes c'est-à-dire naturellement exprimés à la surface des cellules (exemple : les récepteurs à la thyroestimuline) ou exogènes (exemples : bactérie virale, médicaments)

### III-4-3- Hypersensibilités allergiques de type III :

Ici encore, la voie mise en jeu est de type humoral. Elles sont dues à la formation de complexes immuns suite à la rencontre entre des antigènes solubles et leurs anticorps spécifiques, le plus souvent des IgG. Ces complexes s'accumulent en différents sites de l'organisme induisant des lésions via plusieurs mécanismes : activation du complément, stimulation plaquettaire, stimulation des monocytes, macrophages et neutrophiles.

Sur le plan clinique, on observera des néphropathies, des vascularites, et dans les cas les plus graves, un syndrome de maladie sérique.

Les médicaments mis en cause sont notamment certaines pénicillines, l'allopurinol, des anti-inflammatoires non stéroïdiens.



### III-4-4- Hypersensibilités allergiques de type IV :

On parle aussi d'hypersensibilité allergique retardée car les manifestations cliniques n'apparaissent que 24 à 48 heures après la réintroduction de l'allergène chez un patient sensibilisé, ou à médiation cellulaire puisque ici la réponse est de type cellulaire (voie Th1) et ce sont les lymphocytes T qui sont les principaux effecteurs.

La physiopathologie des HSA de type IV est bien décrite pour les réactions cutanées qui se définissent par une lésion d'eczéma au site de contact avec l'allergène. Parmi les HSAR cutanées, la dermatite allergique de contact est la forme la plus étudiée. Ces réactions se déroulent en deux phases :

- La phase de sensibilisation (ou d'initiation) : Lors du premier contact avec le médicament, celui-ci va être pris en charge par les cellules de l'immunité innée situées dans la peau (cellules de Langerhans et dendritiques du derme). Ces cellules vont migrer, jouer leur rôle de CPA et présenter l'allergène aux LT du ganglion drainant. Ces LT spécifiques sont effecteurs (et sont des LTCD8 cytotoxiques dans le cas d'HSAR cutanée) et régulateurs (ce sont des LTCD4+) et vont ensuite recirculer en périphérie. Cette phase est asymptomatique.
- La phase de révélation : Lors d'un deuxième contact avec le même antigène (rappelons que si l'on considère le médicament comme haptène, il sera alors lié à des protéines), les LT spécifiques cytotoxiques sont recrutés dans la peau et activés. Ils seront alors responsables, par leur sécrétion d'IFN gamma de la mise en place de l'inflammation (avec recrutement de cellules inflammatoires non spécifiques : macrophages et neutrophiles) et de la lésion cutanée. Les LTCD4+ Th1 régulateurs exerceront leur rôle en sécrétant des cytokines anti-inflammatoires.

Dans la plupart des cas, ce type d'hypersensibilité sera caractérisé par des manifestations cutanées même lorsque l'administration se fait par voie systémique. L'aspect clinique est variable et pourra aller des rashes ou exanthèmes maculo-papuleux (EMP) qui sont les plus fréquents à des réactions plus rares tel que le syndrome de Lyell

ou le DRESS (Drug Reaction with Eosinophilia and Systemic Symptoms). L'EMP a une présentation clinique polymorphe. Les lésions maculo-papuleuses de taille variable débutent généralement du tronc et des extrémités et confluent en plaques. C'est la toxidermie la plus fréquemment décrite et représenterait de 30 à 95% des réactions cutanées liées aux médicaments (Apaydin, 2000, Bigby 2001).

Les molécules les plus souvent inculpées sont les pénicillines (surtout les aminopénicillines telles que l'amoxicilline et l'ampicilline, les sulfonamides et les anticonvulsivants (Nigen et al. 2003, Gomes and Demoly, 2005).

Dans le paragraphe suivant, nous présenterons les éléments clefs de la reconnaissance du médicament : les IgE pour les HSA de type I, et le complexe CMH/peptide/TCR pour les HSA de type IV.

### III-5- Régulation de la réponse d'hypersensibilité allergique :

Une activation excessive du système immunitaire, qui conduira à des réactions allergiques ou d'auto-immunité est délétère. L'existence de mécanismes régulateurs permettant ainsi le maintien de la tolérance est donc essentielle afin de stopper une réponse inflammatoire engagée lors de la lutte contre un antigène, et de bloquer les processus d'auto-immunité. Ainsi, les pathologies inflammatoires telles que les HSA sont dues à l'activation de cellules effectrices avec un mauvais contrôle des effets anti-inflammatoires non spécifiques et spécifiques (Liyanaage et al. 2002).

La régulation qualitative de la réponse immune fait appel à des coopérations cellulaires et à la production notamment de cytokines et de facteurs humoraux.

Pour le versant humoral, la régulation du système fera appel aux IgG 4 produites par les lymphocytes et capables d'entrer en compétition avec les IgE spécifiques de l'antigène. D'autre part, les récepteurs Fcγ sont extrêmement importants dans cette régulation puisqu'ils interagissent avec les immunoglobulines et peuvent avoir un effet activateur ou un effet inhibiteur de la réponse immunitaire en fonction de la voie de signalisation qui suivra le contact.

Pour le versant cellulaire, plusieurs sous populations lymphocytaires ont été étudiées. La plupart sont des LT CD4+ mais il existe aussi quelques LT régulatrices CD8+. On parle de lymphocytes T régulateurs (ou suppresseurs), couramment appelés Treg. Ces cellules aident à prévenir l'activation des lymphocytes auto-immuns et sont très importantes pour le maintien de l'homéostasie. Leur rôle principal est de réprimer l'activité des cellules de l'immunité, soit auto-immune, soit en fin de réaction immunitaire. Les LT CD4+ présentent une grande capacité anti-inflammatoire du fait principalement de leur capacité à sécréter des cytokines immunosuppressives telles que l'IL10 et le TGF- $\beta$  (Transforming growth factor). On distingue plusieurs sous-populations de LT suppresseurs :

- Les lymphocytes Th3 sont des LTCD4+ sécrétant du TGF- $\beta$  ainsi que de l'IL-4 et de l'IL-10, et de ce fait favorisent la sécrétion d'IgA par les plasmocytes, tout en régulant à la baisse les Th1 et les Th2. (Weiner et al, 2001)

- Les lymphocytes Tr1 produisent également du TGF- $\beta$  mais aussi beaucoup d'interleukine 10 et suppriment la réponse immune grâce à des contacts cellulaires et à la sécrétion de cytokines. (Roncarolo et al.,2001).

- Les lymphocytes T CD4+CD25+ régulatrices. Ces cellules correspondent à environ 10% des LTCD4+. Elles ont un rôle majeur pour le contrôle des maladies inflammatoires chroniques (Dubois et al. 2003). Elles sont caractérisées par les marqueurs de surface CD4 et CD25 ainsi que la molécule FOXP3 (Forkhead box P3), facteur de transcription exprimé au niveau cytosolique. Ces cellules interviennent dans la régulation des maladies inflammatoires, auto-immunes (Suri-Payer, 1998), et dans les allergies de type IV de la classification de Gell et Coombs (Dubois, 2003). Elles joueraient un rôle dans le contrôle des HSA médicamenteuses

- Les lymphocytes CD8+ suppresseurs produisent du TGF-  $\beta$  et de l'IL10 et ont donc un effet négatif sur les LTh1, les LT cytotoxiques et les cellules présentatrices de l'antigène. L'activité suppressive de ces lymphocytes s'appuie aussi sur la sécrétion d'IFN-gamma (Cone et al. 2007).

### III-6- Facteurs de risques :

De nombreux facteurs influencent la survenue de réactions d'hypersensibilité allergique. Ils relèvent d'une part du caractère immunigène des médicaments mais aussi du patient lui-même.

Concernant les facteurs de risque liés au médicament, on sait que les phénomènes d'allergie seront plus fréquents pour des médicaments dont le poids moléculaire est supérieur à 1000 daltons, lors d'une administration intermittente et répétée, pour une dose antigénique forte (posologie quotidienne importante et/ou durée du traitement longue) et pour une administration par voie parentérale. La présence d'une pathologie concomitante peut aussi jouer un rôle.

Concernant les facteurs de risque liés à l'hôte, notons l'âge : les enfants étant en général moins touchés, le sexe : les hommes sont moins atteints et les facteurs génétiques avec l'atopie (mais qui n'est pas applicable aux allergies médicamenteuses).

### III-7- Diagnostic de l'hypersensibilité allergique médicamenteuse :

Le diagnostic formel des phénomènes allergiques est complexe et l'imputabilité d'un ou plusieurs médicaments n'est pas toujours facile. La confirmation d'une HSA (différente de l'HS non allergique) s'appuiera sur la mise en évidence des cellules effectrices spécifiques de l'antigène médicamenteux au moyen de tests cutanés et immunobiologiques.

La démarche diagnostique débutera toujours par un interrogatoire poussé avant de réaliser les tests cutanés et biologiques adaptés.

Dans un premier temps, l'interrogatoire clinique, qui retracera l'histoire du patient en rapport avec son traitement médicamenteux, devra être mené de façon très minutieuse afin de décrire au mieux la symptomatologie, la chronologie des événements cliniques, les médicaments consommés, l'exposition antérieure à la molécule suspectée, les antécédents cliniques et allergiques du patient...

Dans un second temps, les tests cutanés seront réalisés puis, éventuellement les tests biologiques. Un diagnostic précis est toujours basé sur des tests cutanés et quelquefois sur des tests de provocation.

- Tests cutanés : Utilisés dans le diagnostic de l'allergie aux médicaments depuis les années 60, ils restent aujourd'hui la meilleure méthode de détection des IgE spécifiques (pour les HSA type I) d'une molécule médicamenteuse. Dans le cas de réactions de type immédiat, seront réalisés des tests à lecture immédiate : les prick-tests qui consistent à piquer dans l'épiderme à travers une goutte d'allergène et à observer la réponse d'hypersensibilité immédiate qui se manifeste par un prurit, un érythème, une papule ou l'apparition d'un anneau rougeâtre. Il existe aussi les IDR (intra-dermoréactions) pour lesquelles une solution contenant les allergènes est injectée en intradermique. La réaction inflammatoire obtenue détermine si le sujet possède ou non les anticorps correspondants au produit testé. La lecture est réalisée dans les 15 à 20 minutes et le diagnostic prendra en compte le diamètre de la papule.

Dans le cas d'une réaction retardée, on réalise un patch-test où le médicament est appliqué sur la peau sous occlusion avec une double lecture à 48heures puis à 72 ou 96 heures après la pose du test.

Le diagnostic formel n'est pas toujours obtenu, et la réalisation de ces tests peut présenter des risques pour le patient (reproduction de la réaction allergique ou sensibilisation).

- Tests immunobiologiques :

Sachant que les effecteurs des réactions allergiques sont différents en fonction du type d'hypersensibilité, plusieurs tests ont été développés.

Pour les réactions immédiates, IgE médiées, le dosage des IgE spécifiques fait appel soit à leur mise en évidence directe par la méthode du Radio Allergo Sorbent Test (RAST) (dosage des IgE spécifiques chez le patient), soit à des tests cellulaires avec le Test d'Activation des Basophiles, qui mesure les marqueurs d'activation (histamine, leucotriènes, CD 203c, CD 63).

Pour les réactions non immédiates, des tests d'activation lymphocytaire peuvent être réalisés. Il est ainsi possible d'analyser les marqueurs membranaires d'activation lymphocytaire grâce à la cytométrie en flux, la mesure de la sécrétion de cytokines après restimulation *in vitro* (notamment l'IL5 et l'IFN- $\gamma$ . : Teraki et al. 2003 ; Rozieres et al. 2009), la mesure de la prolifération cellulaire (Test de Transformation Lymphocytaire) et l'analyse de la cytotoxicité.

Enfin, les tests de provocation constituent le meilleur moyen diagnostique positif de l'allergie aux médicaments. Le plus gros problème reste la sécurité du patient puisque ces tests consistent à reproduire une réaction allergique.

## **IV- Hypersensibilités allergiques aux bêta-lactamines**

Les réactions allergiques aux pénicillines ont été rapportées dès le début de leur commercialisation. Ces molécules, bien que présentant un excellent index thérapeutique, et très importantes dans la lutte contre les infections, restent les médicaments les plus souvent impliqués dans les réactions allergiques médicamenteuses et peuvent être à l'origine des quatre types d'HSA de la classification de Gell et Coombs. Toutefois, les réactions de type I et IV étant les plus fréquentes, nous ne développerons pas ici les types II et III. Nous nous concentrerons de plus sur les bêta-lactamines les plus prescrites que sont les pénicillines et les céphalosporines.

Du point de vue épidémiologique, on estime qu'environ 10% des patients déclenchent une réaction allergique après la première administration de bêta-lactamines. Parmi eux, 15 à 20% présenteront une réaction sévère à l'occasion d'une nouvelle administration.

La majorité des réactions présumées allergiques aux bêta-lactamines sont des réactions cutanées. Les plus fréquemment rapportées étant les rashes maculopapuleux (près de 75 % des réactions présumées allergiques aux  $\beta$ -lactames), viennent ensuite les réactions évoquant une HSA immédiate avec principalement l'urticaire et l'œdème de Quincke (Ponvert et Scheinmann 2003, Ponvert et al. 1999, Ponvert et al. 1996, Juchet et al. 1994, Romano et al. 1997). Le choc

anaphylactique survient dans 0,2% des cas avec une mortalité de 0,02% (Park et al. 2005, Birnbaum et al. 1997). On estime, toutefois, que le nombre de décès secondaires à une réaction anaphylactique aux bêta-lactamines représente 75% des chocs mortels d'origine médicamenteuse (Bousquet et al. 2007).

Les patients les plus touchés sont les adultes entre 20 et 49 ans (Birnbaum et al. 1997). Le terrain atopique ne paraît pas être un facteur de risque mais serait lié à des réactions plus graves. Bien que certaines ethnies aient une prédisposition aux accidents allergiques médicamenteux, ceux dus aux  $\beta$ -lactames ne semblent pas être liés aux phénotypes HLA (Juchet et al. 1994, Ponvert et al. 1996, Birnbaum et al. 1997, Guglielmi et al. 2005). D'après une étude réalisée à propos du contrôle génétique des réactions immunitaires secondaires à ces antibiotiques, l'origine ethnique n'affecte pas l'incidence de survenue de ces réactions (Birnbaum et al. 1997). Bien que ces réactions semblent être plus fréquentes chez l'adulte entre 20 et 49 ans, les âges extrêmes ne sont pas épargnés. Enfin, les réactions d'HSA aux bêta-lactamines ne semblent pas être plus fréquentes chez les patients ayant reçu antérieurement ces molécules. En effet, Rudolph et Price en 1973 n'ont pas noté de différence significative entre les patients qui n'ont jamais reçu de pénicillines (0,87%) et ceux qui en ont reçu antérieurement (0,6%). La présence de certaines maladies sous-jacentes (le syndrome d'immunodéficience acquise, la mononucléose infectieuse et la mucoviscidose) (Ponvert et al. 1999, Just et al. 2002) et la voie d'administration préconisée pour les  $\beta$ -lactames semblent être prédisposantes à la survenue des réactions d'HSA. En effet, ces pathologies sont plus fréquentes et sévères lorsque le médicament est administré par voie parentérale (Birnbaum et al. 1997).

## IV-1- Immunogénicité des bêta-lactamines :

Comme la plupart des médicaments, les molécules de type  $\beta$ -lactamine ont un poids moléculaire inférieur à 1000 Daltons et de ce fait ne sont pas immunogènes en tant que telles. Pour être reconnues par le système immunitaire, elles doivent auparavant se fixer de façon covalente à une protéine porteuse. Les bêta-lactamines sont donc des haptènes.

La liaison à une protéine se fait après ouverture du cycle  $\beta$ -lactame. Les déterminants antigéniques reconnus par les immunoglobulines ont été étudiés et identifiés pour la famille des pénicillines mais beaucoup moins concernant les céphalosporines.

#### IV-1-1-Déterminant majeur :

Lorsque la pénicilline pénètre dans l'organisme, elle subira deux types de modifications. On peut observer une ouverture spontanée du cycle  $\beta$ -lactame, exposant ainsi le groupement carbonyle susceptible de réagir chimiquement avec l'amine d'une lysine ou d'une sérine (principalement) de la protéine porteuse. Le groupement **pénicilloyl** issu de l'ouverture du cycle est appelé déterminant majeur puisqu'il constitue le principal produit de dégradation immunogène et que environ 95% de la pénicilline liée aux protéines tissulaires est sous cette forme. Dans le cas de la benzylpénicilline (ou pénicilline G), on parle du BPO pour benzylpénicilloyl. Du point de vue clinique, ce déterminant est responsable de réactions dans 70% des cas (Birnbaum et al., 1997).

#### IV-1-2-Déterminants mineurs :

La pénicilline peut également subir une isomérisation en acide pénicillénique qui se liera de façon covalente à des protéines conduisant ainsi à la formation, dans 5% des cas, de déterminants antigéniques dits mineurs tels que pénicillanyl et pénicillenate. Toutefois, ces déterminants peuvent être à l'origine d'accidents sévères (Demoly et al. 2003)

La figure 7 présente les principaux déterminants formés à partir de la pénicilline (Weiss, 1988).



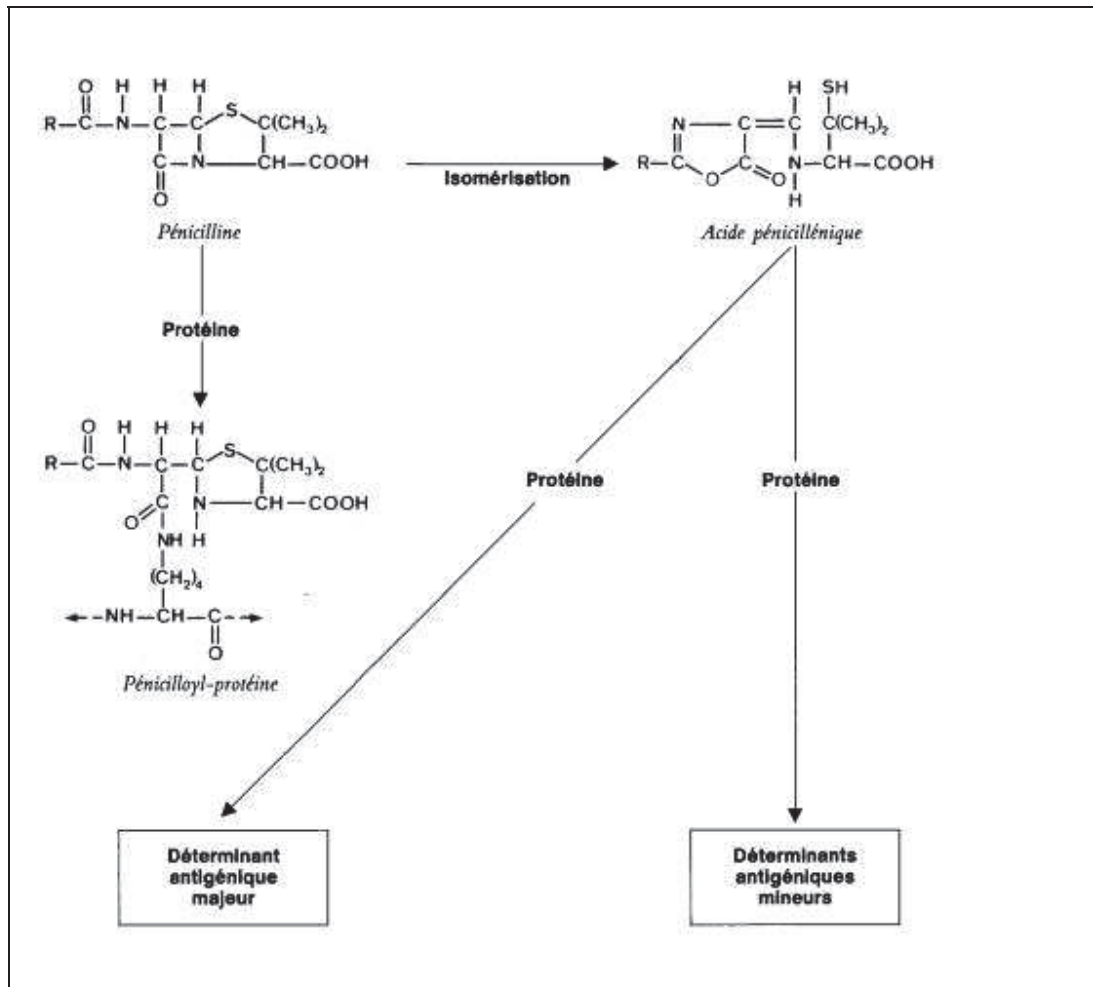


Figure 7 : Principales voies métaboliques de la pénicilline, *in vivo*, aux conditions physiologiques

Il faut noter que la reconnaissance par les immunoglobulines E spécifiques peut se faire au niveau d'autres épitopes que ceux mentionnés ci-dessus. En effet, des IgE spécifiquement dirigées contre la chaîne latérale des aminopénicillines ont été mises en évidence (Birnbbaum et al. 1997, Torres et al., 1998). Ainsi, plusieurs zones de reconnaissance par le IgE peuvent exister (Torres et al., 1998).

En ce qui concerne les céphalosporines, les déterminants antigéniques sont moins bien connus. Le groupement céphaloyl a été identifié mais est très instable (Romano, 1998).

Nous développerons dans les chapitres 5 et 6 les différentes zones de reconnaissance par le système immunitaire et pourrons ainsi présenter des hypothèses concernant des sensibilisations et réactions croisées entre les différentes molécules de type bêta-lactamine.

## IV-2-Réactions immédiates IgE médiées :

Comme nous l'avons présenté dans le paragraphe III-4-1, ce type de réaction (type I de la classification de Gell et Coombs) se caractérise par la production d'immunoglobulines E spécifiques de l'allergène, ici la bêta-lactamine. Le délai de survenue des symptômes va de quelques minutes à quelques heures. Rappelons que les IgE spécifiques de l'antigène sont fixées sur les récepteurs du Fc à haute affinité, et l'interaction spécifique avec l'antibiotique va entraîner un pontage de deux IgE et une dégranulation des mastocytes et basophiles qui libèrent les médiateurs préformés (histamine principalement) et néoformés (leucotriènes), entraînant ainsi la symptomatologie connue des allergies IgE médiées.

Ce travail de thèse s'est penché sur les interactions au niveau moléculaire entre les IgE et les bêta-lactamines. Ainsi, il est important de connaître les caractéristiques de ce type d'anticorps, Et plus particulièrement la partie hypervariable capable de reconnaître spécifiquement l'antibiotique.

La structure de base des anticorps (type IgG) est un complexe de 4 chaînes protéiques identiques 2 à 2 (2 chaînes lourdes H =Heavy et 2 chaînes légères L=Light), fixées par des ponts disulfure inter et intra-caténaires. Chaque chaîne d'Ig contient des domaines constants (Fc pour Fragment constant formé par les régions constantes des deux chaînes lourdes) et variables (Fab pour Fragment antigen binding formé par les régions variables et hypervariables des quatre chaînes polypeptidiques de la molécule).

Les **immunoglobulines de type E** ont un poids moléculaire de 190 kDaltons et sont les moins nombreuses des immunoglobulines circulantes. On estime que les IgE sériques correspondent à environ 50% des IgE totales, le reste étant lié aux mastocytes et basophiles via leurs récepteurs à haute affinité Fcε RI. Elles sont constituées de quatre chaînes protéiques (2 chaînes lourdes epsilon et 2 légères kappa ou lambda) organisées en domaines d'environ 110

acides aminés chacun. Les IgE ont environ 12% de sucres et présentent une structure similaire à celle des autres immunoglobulines mais avec un pont disulfure supplémentaire et un domaine constant en plus pour les chaînes lourdes. Elles sont synthétisées par les lymphocytes B après commutation isotypique induite par l'environnement cytokinique provenant des lymphocytes T CD4+. La commutation IgE passe par le réarrangement des gènes du lymphocyte B, jusque là producteur d'immunoglobulines de type M et qui devient désormais producteur d'IgE. Trois éléments sont requis pour induire cette commutation. D'une part les interleukines 4 et 13. D'autre part, le couple CD40/CD40L et enfin, le CD23/CD21. Les IgE ne fixent pas le complément mais exercent leurs effets biologiques par liaison de leur domaine constant epsilon 2 et 3 avec les récepteurs cellulaires (FcεR). Il existe deux types de récepteurs à IgE. Le récepteur à IgE de forte affinité, le FcεR1 qui possède une constante de dissociation avec l'IgE de  $10^{-9}$  à  $10^{-10}$ M et le récepteur à IgE de faible affinité, le FcεR2 qui est essentiellement présent sur les éosinophiles mais aussi sur la paroi des macrophages alvéolaires, les plaquettes, les lymphocytes. Notons que les interactions IgE/ FcεR2 sont faibles et c'est la sommation de liaisons proches qui peut induire une activation cellulaire. La figure 8 ci-après (Zanecki, 2007) présente la structure d'une IgE.

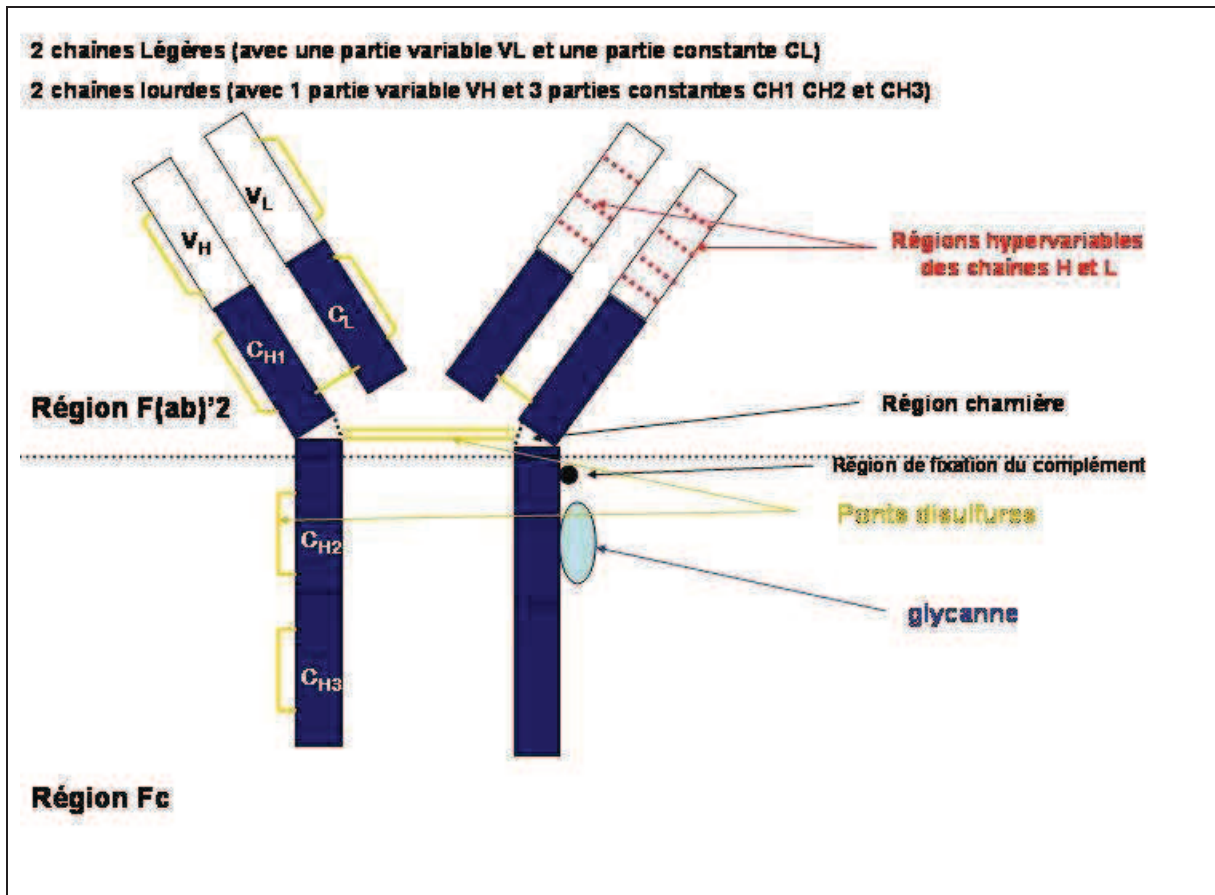


Figure 8 : Structure d'une immunoglobuline de type E

La reconnaissance spécifique de l'antigène se fera au niveau de régions hypervariables (HVR), encore appelées CDR pour Complementarity-Determining Region. Au sein du domaine variable, les régions hypervariables se présentent sous forme de trois boucles exposées à l'extrémité (on a donc HV1, HV2 et HV3). Ces régions sont entourées de régions « cadre » bien moins variables ce qui offre à l'immunoglobuline une variabilité localisée en s'appuyant sur une structure stable. L'appariement de la chaîne légère L et de la chaîne lourde H assemble des boucles hypervariables en deux domaines créant ainsi une surface composite qui forme à l'extrémité de chaque « bras » Fab le site d'anticorps se liant à l'antigène de manière ultra-spécifique.

## IV-2-Réactions retardées :

Elles sont dues à l'action incontrôlée de lymphocytes T spécifiques du médicament. Rappelons que les bêta-lactamines sont des haptènes qui doivent se lier de manière covalente à des protéines pour sensibiliser un individu. Le médicament hapténisé va être pris en charge par une cellule présentatrice de l'antigène qui après apprêtement présentera l'antigène toujours fixé au peptide (issu de la dégradation de la protéine porteuse) au sein de la niche à peptides, au TCR d'un lymphocyte T. Dans le cas d'un LTCD4, le CMH sera de classe II, et dans le cas de LT CD8, le CMH sera de classe I. Même si l'on considère que le médicament suit la voie du Pi concept (donc sans apprêtement), là encore, le CMH et le TCR seront indispensables à la phase de sensibilisation ou de révélation.

Les résultats de ce travail de thèse présenteront les interactions au niveau moléculaire impliquant le CMH, le peptide de la niche, le médicament et le TCR. Nous étudierons les différents scénarios possibles (hapténisation, Pi concept...).

Les molécules de CMH pour **Complexe Majeur d'Histocompatibilité** constituent le système de reconnaissance du soi. Le CMH correspond à un complexe de gènes différents entre deux individus et qui s'expriment notamment par des protéines de surface. Le rôle principal des antigènes du CMH de classe 1 et 2 consiste à assurer la « présentation des épitopes T » aux lymphocytes T. Ces molécules font partie de la superfamille des immunoglobulines. On distingue le CMH de classe I présent sur presque toutes les cellules de l'organisme et destiné à la présentation antigénique aux LTCD8. Les molécules de classe II quant à elles se retrouvent sur certaines cellules telles que les cellules dendritiques, les monocytes macrophages et les LB, et présenteront les antigènes aux LTCD4 étant ainsi à l'origine du déclenchement de la réponse immunitaire. Les molécules de CMH sont la clef de l'immunité cellulaire et de la communication entre les cellules. Pour cela, ces molécules sont exprimées en surface des CPA et servent à présenter l'antigène, au sein de la niche à peptides, aux cellules compétentes. En effet, les molécules de CMH ne sont stables que si elles sont associées à un peptide situé dans une poche appelée niche à peptides. Les résidus d'ancrage sont les acides aminés de la niche qui interagissent via des liaisons faibles de type hydrogène et ionique avec les acides aminés du peptide. De très nombreux peptides peuvent être ainsi présentés au sein de la niche. Nous

possédons au maximum six molécules de CMH de classe I et douze de classe II ce qui est suffisant pour reconnaître des dizaines de milliers de peptides différents. Ainsi, les molécules de CMH sont sélectives mais ne possèdent pas la spécificité de TCR.

La molécule de CMH I, exprimée de façon quasi ubiquitaire est impliquée dans la présentation de peptides endogènes aux LTCD8 et se compose d'une chaîne alpha et une chaîne bêta-microglobuline. La niche à peptide peut contenir des peptides de 8 à 10 acides aminés et les résidus d'ancrage sont situés aux deux extrémités.

La molécule de CMH de classe II est retrouvée en surface des cellules présentatrices de l'antigène telles que les cellules dendritiques, les LB, les macrophages et sera reconnue par les LTCD4. Elle se compose d'une chaîne alpha, une chaîne bêta et sa niche à peptides accueillera des peptides de 13 acides aminés au moins avec une interaction qui se fera sur toute la longueur de la poche.

On parlera de système HLA (human leucocytes antigen) chez l'Homme. Certains locus sont plus particulièrement étudiés, notamment HLA-A, HLA-B ou HLA-C (appartenant au MHC de classe I) et HLA-DP, HLA-DQ et HLA-DR (appartenant au MHC de classe II). La figure 9 (cours en ligne, Université de Liège) présente la structure des molécules de CMH 1 et 2.

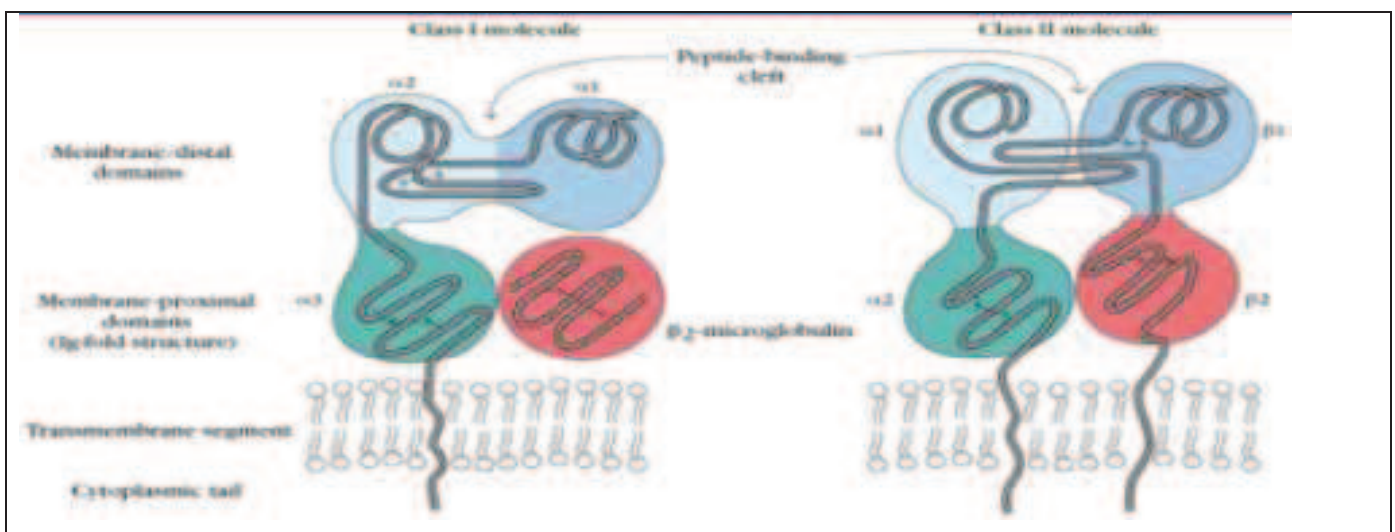


Figure 9 : Structure des molécules de CMH de type 1 et 2

Le récepteur des cellules T ou **TCR** (T Cell Receptor) correspond à un complexe moléculaire présent à la surface des LT. Ce récepteur membranaire constitue un paratope et permet une reconnaissance spécifique des complexes CMH/peptide. C'est un hétérodimère composé de 2 chaînes glycoprotéiques de membrane  $\alpha$  et  $\beta$  le plus souvent (ou  $\gamma$  et  $\delta$ ) reliées de façon covalente par un pont disulfure. L'interaction du TCR avec son antigène cible entraîne l'activation du lymphocyte via une série d'évènements impliquant enzymes, co-récepteurs et molécules spécialisées. La reconnaissance de l'antigène n'est pas directe puisqu'il est indispensable que l'antigène soit présenté au TCR par l'intermédiaire de molécules du CMH. Notons que les antigènes reconnus par le TCR sont obligatoirement linéaires.

Chaque chaîne du TCR fait partie de la superfamille des immunoglobulines et comporte une région N-terminale variable (V) composée de 3 régions hypervariables (CDR1, CDR2, CDR3 comparables à celles de la région V des chaînes des Ig), une région constante (C), une région charnière (H) où se fait la liaison covalente entre les 2 chaînes (pont S-S), une région transmembranaire qui possède des acides aminés chargés positivement et enfin une extrémité cytoplasmique très courte. Le CDR3 est la principale région impliquée dans la reconnaissance de l'antigène car placé au centre du paratope. Les CDR1 et CDR2 interagissent surtout avec la molécule de CMH. Le paratope présente une forme assez constante correspondant à une cavité définie par les six CDR avec une diversité concentrée sur le CDR3. La région transmembranaire du TCR est composée d'acides aminés basiques, c'est-à-dire chargés positivement. Cette charge électrostatique permet au TCR de s'associer avec d'autres protéines (comme le complexe CD3), qui sont des molécules « accessoires » avec une région transmembranaire chargée négativement, et qui sont nécessaires à la transduction du signal à partir du TCR. Sachant que l'affinité de l'interaction du TCR avec le complexe peptide-CMH est faible, des corécepteurs sont nécessaires, et la transduction du signal est alors augmentée par la liaison simultanée des complexes CMH/peptide avec un co-récepteur spécifique. Les cellules portant un TCR capable de se lier à une molécule du CMH de classe II portent également la molécule CD4 ; celles qui lient le CMH de classe I portent la molécule CD8. Ces co-récepteurs renforcent et prolongent l'interaction entre le lymphocyte T et la cellule présentatrice d'antigène, et participent aussi à la transduction du signal.

Le schéma 3 ci-après représente la structure du TCR

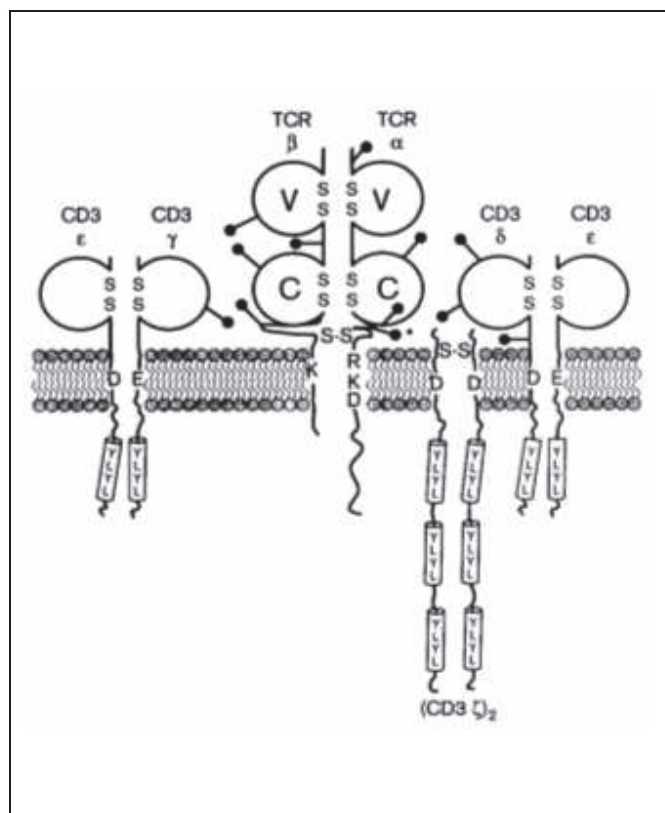


Schéma 3 : Structure du TCR (Malissen et al., 1996)



# Chapitre 2 : La modélisation moléculaire

La conception rationnelle de produits actifs, ou modélisation moléculaire, est une méthodologie couramment utilisée dans l'industrie pharmaceutique. Depuis un peu plus d'une trentaine d'années, elle s'est imposée progressivement comme un outil de choix pour la découverte et la conception orientée de nouvelles molécules actives. Auparavant, n'étaient pratiqués que des tests biologiques systématiques appelés « screening » sur un grand nombre de molécules et, souvent, seule la chance permettait de mettre en évidence une piste intéressante.

Tous les pays industrialisés sont confrontés aujourd'hui à une croissance forte de leurs dépenses de santé, et tout particulièrement de leurs dépenses pharmaceutiques. Ces dernières augmentent sous l'impulsion des différents facteurs que sont le vieillissement de la population, une plus forte proportion de consommateurs de médicaments à tous les âges (notamment sous la pression de la publicité), et des prix très élevés pour les nouveaux médicaments. Ces tarifs découlent du fait que les firmes pharmaceutiques ont obtenu la liberté des prix dans tous les pays industrialisés, et/ou que les autorités chargées de la fixation des prix dans ces pays acceptent des tarifs proches de ceux exigés par les firmes.

Le coût de recherche et développement le plus souvent cité est de « 802 millions de dollars » par médicament voir même un milliard de dollars US suivant plusieurs sources. Et même si il est probablement surestimé, il n'en reste pas moins très important et mérite que des approches scientifiques novatrices, telle que la modélisation moléculaire, soient développées.

A ce jour, la nécessité de développer de nouvelles drogues, adaptées aux nombreuses pathologies existantes et à venir, pousse l'ensemble de la communauté scientifique et industrielle vers une même voie : concevoir un médicament capable de soigner un patient tout en étant sans danger pour l'organisme. Un médicament se doit d'être d'une grande efficacité et d'une grande spécificité.

La conception d'une nouvelle molécule médicamenteuse constitue une aventure exaltante et enrichissante, pouvant permettre de grands progrès scientifiques, mais aussi extrêmement longue, coûteuse, et risquée. Avant sa mise sur le marché, le médicament

passera par les phases de recherche et développement, puis sera testé en terme d'efficacité, spécificité et d'effets secondaires. Ainsi, il faudra plus de 10 ans pour que la molécule soit commercialisée. Ceci entraîne des coûts extrêmement importants et de ce fait, une forte pression financière s'exerce au cours du développement d'une molécule thérapeutique. Les dépenses des phases de R&D sont augmentées par l'ensemble des tests biologiques, *in vitro*, *in vivo* chez l'animal puis chez les patients, qu'il est nécessaire d'effectuer. L'amélioration de la drogue représente également un investissement supplémentaire. A cela, il faut ajouter le risque d'effets secondaires potentiellement dangereux pour le patient. Ainsi, les industries pharmaceutiques sont elles soumises à de fortes pressions économiques et humaines.

La modélisation moléculaire, du fait qu'elle n'implique que des analyses *in silico*, offre de nombreux avantages à toutes les étapes de conception d'une molécule d'intérêt thérapeutique. En effet, lors des phases de R&D, elle permettra soit la conception *de novo* du futur médicament, soit l'étude de l'action d'une substance connue pour ses effets vis-à-vis d'une cible avec une amélioration possible de l'interaction drogue/cible. Elle permettra également la création de médicaments ultra-spécifiques diminuant ainsi les effets secondaires. D'autre part, lors des phases de développement et d'essais, la modélisation progressera en parallèle avec les biologistes et permettra de réduire les tests *in vitro* et *in vivo* réalisés. Il est également possible de modifier un médicament déjà existant en vue d'en améliorer l'efficacité et la spécificité. Les études de modélisation offrent également l'avantage de comprendre de manière fondamentale le mécanisme d'action d'une drogue, ou certaines réactions secondaires comme c'est le cas pour les hypersensibilités allergiques induites par les bêta-lactamines.

Etant donné que la modélisation n'est que théorique, une étroite et forte interaction avec l'ensemble de la communauté scientifique, réunie pour la conception du médicament est absolument nécessaire. Les équipes de modélisateurs proposent une version théorique, validée ou invalidée par les biologistes, les médecins...et un jeu de « ping pong » se met en place entre modélisateurs et autres scientifiques afin de concevoir une molécule candidate efficace, spécifique et sans danger pour le patient.

# **I- Evolution des différentes méthodes utilisées pour découvrir de nouveaux principes actifs :**

Concentrons-nous sur le secteur de la chimie médicinale. Cette problématique est extrêmement ancienne, car le désir de guérir les malades est apparu à l'aube de l'humanité. Il est vrai qu'à ses débuts, elle a hésité et s'est basée exclusivement sur l'apprentissage par l'échec et l'empirisme. Durant l'ère égyptienne et l'ère romaine, les premiers remèdes, basés sur une formulation complexe à base de plantes, ont permis de traiter quelques patients. Cette évolution s'est poursuivie pendant des siècles, toujours basée sur ces deux mêmes concepts de base. Les connaissances scientifiques progressent fortement à partir du 18<sup>ème</sup> siècle, permettent de comprendre les problèmes qui sont posés, et d'y répondre par une démarche plus rationnelle. La recherche de nouveaux complexes actifs d'origine végétale, ainsi que leurs modifications chimiques, a évolué. Les progrès de la chimie ont permis un peu plus tard de créer les premières molécules de synthèse. Pourtant, la structure des molécules ainsi que leurs modifications structurales restent sous l'influence de l'empirisme.

De nos jours, plusieurs méthodes permettent de découvrir ou de redécouvrir des principes actifs. L'ensemble de ces méthodes s'organise selon un schéma général exposé ci-après :

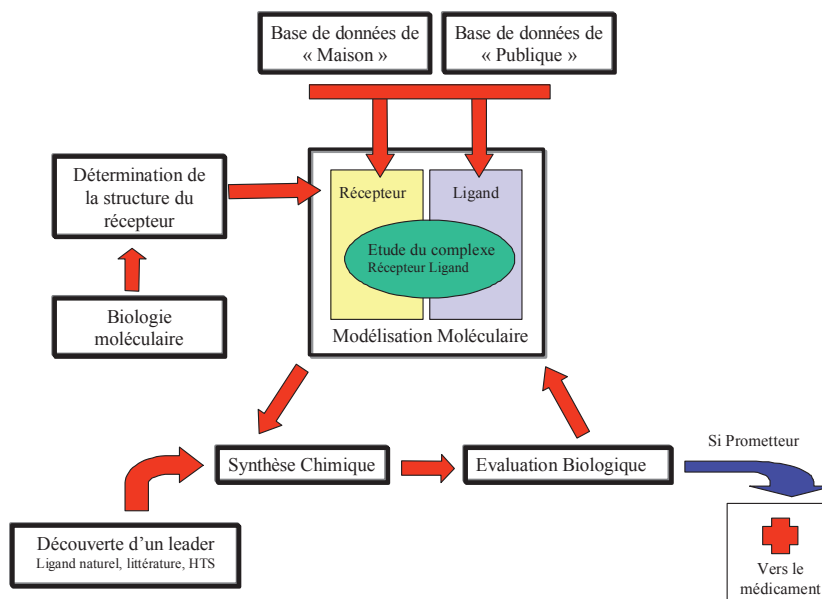


Schéma 4 : Schéma général de la recherche d'une molécule active

Plusieurs stratégies de recherche sont envisageables pour découvrir de nouvelles molécules :

1. Dans un premier cas, est recherchée une molécule, appelée molécule « leader », présentant l'activité désirée. Pour cela, les méthodes modernes de « **High Throughput Screening** » (HTS), qui permettent de tester des dizaines de milliers de composés par jour, ont révolutionné le domaine. Ces collections de molécules ont plusieurs origines possibles :
  - Ce peut être un produit naturel qui interagit avec le récepteur. De nouvelles molécules peuvent être trouvées en récoltant des espèces végétales et animales dans des régions dont la biodiversité est très grande. Elles formeront des bibliothèques de molécules.
  - On peut également redécouvrir des molécules de remèdes ancestraux afin de constituer de nouvelles bibliothèques.
  - Il existe un large éventail de sources pour trouver un produit « leader ». Cela peut aller de la firme d'un concurrent jusqu'au domaine public. En effet, la description d'un très grand nombre de composés est aujourd'hui accessible.
2. Dans un deuxième cas, il n'y a pas de molécule « leader ». La mise au point des méthodes de HTS a provoqué la création de nouvelles stratégies de synthèse de molécules : la

chimie combinatoire et la synthèse parallèle. En effet, si l'on synthétise molécule par molécule, un service de chimie ne produira que quelques centaines de molécules par an.

- **La chimie combinatoire** propose une approche complètement différente. Deux méthodes se distinguent, suivant le type de molécules désirées : la conception d'un biopolymère débutera avec des monomères qui permettront de synthétiser plusieurs polymères possédant la longueur voulue et présentant toutes les combinaisons possibles. Si, par contre, nous voulons une molécule disposant d'un motif central commun et effectuer des variations structurelles autour de celui-ci, il faudra faire réagir plusieurs groupements afin de « décorer » ce motif central. Chaque combinaison de groupements aboutira à une molécule différente.
  - **La synthèse en parallèle** est une voie différente de la chimie combinatoire. Elle permet d'automatiser des réactions chimiques grâce à des robots. Ceux-ci disposent de plusieurs réacteurs. Cette méthode permet de synthétiser automatiquement des composés similaires. Elle sera utilisée soit pour synthétiser une série de molécules, soit pour optimiser une réaction chimique en synthétisant la même molécule sous des conditions opératoires différentes. Chaque molécule est isolée, dans un puits ou sur un support de type cellulose dont le fond est enduit par le récepteur étudié lui-même couplé à un chromatophore. Une réaction colorimétrique témoignera de l'interaction drogue / récepteur. L'affinité de la molécule sera lue par un lecteur optique.
3. Enfin, le service de **modélisation moléculaire** peut soit concevoir une molécule à partir d'un récepteur connu soit, dans une deuxième phase, optimiser une molécule. Plusieurs types d'études peuvent être réalisés :
- Des études portant seulement sur la structure des ligands : elles permettent essentiellement de tirer des relations structure - activité qui serviront soit à prédire l'activité de nouvelles molécules virtuelles, soit à rechercher la molécule pouvant interagir de façon optimale avec son récepteur.
  - Des études visant le complexe récepteur / ligand. Si la structure du ligand peut être calculée, celle du récepteur ne peut par contre pas provenir uniquement du calcul. Soit la structure est dans le domaine public, soit elle est assez proche d'une structure déjà publiée permettant ainsi une modélisation par homologie. Lorsque la structure est totalement inconnue, nous ferons appel à la cristallographie de rayons X ou à la RMN-

2D. Il est également possible de créer un récepteur virtuel à partir des informations extraites des différentes molécules interagissant avec le récepteur. Ce type d'approche est utilisé dans les méthodes de 3D QSAR ou de *de-novo* design. Ainsi, lorsque la structure du récepteur sera résolue, une étude ligand / récepteur sera possible et permettra de déterminer le meilleur ligand.

Il faut garder en tête que la recherche de la molécule optimale est basée sur un cycle de circulation de l'information, visant à améliorer son activité par des modifications structurales à chaque nouveau cycle. Plusieurs cycles sont nécessaires afin que soit mise au point la molécule active optimale qui sera peut-être incluse dans une formulation médicamenteuse.

Rappelons l'importance des bases de données et des librairies chimiques, capitales dans la découverte de nouvelles molécules. Actuellement, de grandes sociétés se sont lancées dans la collection de molécules, afin de traiter les problèmes d'aujourd'hui et de réduire le temps de recherche d'une molécule pour les futurs problèmes. Elle prennent ainsi une longueur d'avance sur la concurrence et pourront conquérir demain des parts de marché. L'avenir se crée, se gère et s'anticipe : les grandes entreprises pharmaceutiques mondiales l'ont déjà compris, même si cela se fait parfois au mépris des règles déontologiques.

## **II- Du modèle aux méthodes :**

La modélisation moléculaire permet de simuler la forme d'une molécule et son interaction avec d'autres éléments, à partir d'un modèle du système, et en ayant recours à différentes méthodes de calcul. Un modèle correspond à une représentation schématique de la réalité. Réduisons notre étude au cas de la chimie médicinale. Ce sont les interactions entre des ligands et des récepteurs ou entre deux biopolymères qui sont le plus souvent étudiées.

### **II-1- Les origines expérimentales du modèle :**

L'étude des interactions d'un ligand vis à vis de son récepteur n'est possible que si l'on dispose d'un modèle moléculaire fidèle du récepteur. Si aujourd'hui on peut, avec plus ou moins de confiance, simuler le repliement d'une protéine, il faut d'abord que le modèle soit défini expérimentalement. Il existe deux grandes familles de méthodes expérimentales de

résolution de la structure d'un biopolymère : la Résonance Magnétique Nucléaire bi-dimensionnelle (RMN-2D), et la cristallographie de rayons X.

### II-1-1- La Résonance Magnétique Nucléaire 2D :

La **RMN** est une méthode datant de plusieurs dizaines d'années. Elle permet de déterminer la structure d'une molécule. La RMN-2D sert à résoudre des structures de protéines ou d'acides nucléiques. Une série d'impulsions de type NOESY, COSY... permet de déterminer le voisinage spatial d'un atome en particulier. L'évolution des technologies des appareils de RMN-2D a augmenté leurs résolutions et permet de déterminer des relations de voisinage au sein de molécules de plus en plus grosses et complexes. Ces relations sont traduites en contraintes qui vont servir à bâtir le modèle convenable, tiré d'un modèle théorique non replié. L'un des avantages de la RMN-2D est de pouvoir étudier une structure en milieu physiologique. Afin d'obtenir la structure correcte de la molécule, il est nécessaire de satisfaire les différentes contraintes de distances qui ont été déterminées par RMN-2D. Pour cela, des simulations de dynamique moléculaire sont lancées. Il faut noter qu'une structure, même si elle satisfait toutes les contraintes RMN-2D, peut être encore à moitié dénaturée. Pour remédier à ce problème, la durée de simulation de dynamique moléculaire, les conditions de la simulation, ainsi que les méthodes appliquées sont d'une importance critique pour le résultat final. La RMN-2D permet non seulement de déterminer la structure, mais également de mieux comprendre le repliement des protéines.

### II-1-2- La cristallographie de rayons X:

La deuxième grande famille de méthodes correspond à la **cristallographie de rayons X**. Cette méthode consiste à faire cristalliser la protéine en agissant sur le milieu. Après sélection du cristal, celui-ci est soumis à l'irradiation de rayons X. Les rayons X diffractants permettent de déterminer une carte de densité électronique du cristal, à partir de laquelle la position des différents atomes de la structure sera déduite. La carte de densité électronique sera alors traduite en différentes contraintes utilisées pour affiner la structure grâce à une simulation de dynamique moléculaire.

Les méthodes cristallographiques connaissent des limites. Tout d'abord, la résolution de l'appareil, mais les progrès scientifiques et techniques ont amélioré ce point négatif. D'autre part, une des plus importantes limitations vient de la difficulté à obtenir des cristaux

de bonne qualité et, lorsque l'on y parvient, le milieu de cristallisation reste très éloigné du milieu physiologique.

## II-2- Les modèles théoriques :

L'étude et la simulation des propriétés d'une molécule requièrent la construction d'un modèle structural de celle-ci. Les simulations pourront décrire d'autant plus finement les propriétés que le modèle sera exact et que les méthodes de simulations nécessaires relieront bien le modèle aux propriétés simulées. A partir de systèmes biologiques, un modèle théorique peut être construit en utilisant la méthode la plus adéquate. C'est le cas des systèmes d'acides nucléiques pour lesquels il est possible de concevoir la structure tertiaire d'un ensemble avec une assez bonne fiabilité.

Le cas des protéines est différent car le polypeptide est très flexible et dispose d'un grand nombre de degrés de liberté. Pour ce type de systèmes, nous recherchons dans les bases de données structurales une protéine présentant des similitudes avec la protéine dont on veut construire le modèle. Cette technique est appelée **modélisation par homologie**. Il s'agit sans doute de la méthode la plus répandue actuellement, notamment depuis l'augmentation de la disponibilité de structures résolues expérimentalement dans les banques de données. La modélisation par homologie permet de générer la structure d'une protéine à partir d'une protéine homologue de structure connue. Elle est basée sur l'hypothèse que deux séquences homologues auront des structures tridimensionnelles proches. Dans les différents niveaux d'organisation d'une protéine, plus on se rapproche de la structure tridimensionnelle, plus l'information sera conservée. En d'autres termes, la structure tertiaire d'une protéine sera beaucoup plus conservée que sa séquence. Ainsi, la recherche d'homologues de la protéine à modéliser, possédant des structures connues, va permettre de construire par homologie la structure de la protéine cible. Cette modélisation fonctionne bien pour des identités de séquences supérieures à 30% (Chothia et al., 1986), puisqu'au delà de cette limite, 80% des carbones alpha se superposent structurellement. En dessous, il faut utiliser d'autres critères que l'identité pour inférer l'information structurale à la protéine cible (par exemple la conservation des structures secondaires, Geourjon et al., 2001). Les étapes principales seront: la recherche d'homologues dans les banques de données, l'alignement de la protéine cible avec la ou les séquences homologues, l'extraction des données structurales de l'empreinte (contraintes de distances, angles dièdres, etc) et l'application de ces données sur la protéine cible et enfin, la validation du modèle. L'inférence de l'information structurale de la « protéine empreinte » à la



protéine à modéliser détermine le groupe auquel appartient la méthode. On distinguera la modélisation par assemblage de fragments rigides (rigid body assembly), par identification de segments (segment matching), par satisfaction de contraintes spatiales (spatial constraints satisfaction) et plus récemment par évolution artificielle (artificial evolution). Les méthodes par assemblage de fragments rigides ont été les premières développées. Citons SwissModel (Schwede et al., 2004), 3D-Jigsaw (Bates et al., 2001) ou Builder (Koehl et al., 1994). Les méthodes par identification de segments utilisent des sous ensembles de positions atomiques dérivées de l'alignement comme guide pour trouver des segments correspondants dans une banque représentative de toutes les structures connues. La méthode la plus ancienne est SegMod/ENDCAD (Levitt, 1992). Les méthodes par satisfaction de contraintes spatiales appliquent les contraintes extraites de l'empreinte sur la protéine à modéliser puis les violations de contraintes sont minimisées. Les méthodes les plus utilisées sont Modeller (Fiser et al., 2003) et Geno3D (Combet et al., 2002), accessibles par le web. Enfin, les méthodes par évolution artificielle utilisent la notion de mutation. Ce sont les méthodes les plus récentes. L'alignement entre l'empreinte et la protéine cible est considéré comme une succession de mutations, insertions et délétions. Chaque opération conduisant de l'empreinte à la protéine cible représente un coût énergétique positif ou négatif au niveau de la structure. La construction du modèle débute par l'opération ayant le plus faible coût énergétique et ainsi de suite. Le programme le plus courant utilisant cette méthode de construction est Nest (Petrey et al., 2003) du logiciel Jackal.

Les modèles issus de la modélisation par homologie nécessitent souvent une étape de minimisation d'énergie afin d'obtenir une conformation plus stable énergétiquement. Des bibliothèques de rotamères (différentes conformations possibles pour les chaînes latérales des résidus) peuvent être utilisées afin d'améliorer la qualité du modèle (Bower et al, 1997).

## II-3- Les méthodes de simulation :

On distingue deux classes de méthodes : la première est basée sur la **mécanique quantique**, et la seconde sur la **mécanique moléculaire**. Il est possible de les combiner et l'on parle alors de **méthode mixte QM / MM**.

## II-3-1- Les méthodes quantiques :

Cette classe de méthodes base les calculs sur l'équation de Schrödinger. Celle-ci ne peut être résolue exactement, sauf pour l'ion  $H^{2+}$ , ce qui est extrêmement limitant. Pour y remédier, plusieurs approximations sont effectuées: la première correspond à l'approximation de Born-Oppenheimer qui fixe la position des noyaux. Malgré cela, la résolution de l'équation est toujours impossible, et d'autres approximations sont nécessaires.

A ce stade on distingue deux méthodes de traitement différentes:

1- La première méthode, qui est aussi la plus ancienne, repose sur le calcul de propriétés à partir de l'orbitale moléculaire polyélectronique  $\psi$ . Cette orbitale ne peut être construite de façon exacte, il est alors nécessaire d'utiliser plusieurs niveaux d'approximation pour y parvenir. La première approximation appelée « approximation orbitale » permet de calculer l'orbitale moléculaire polyélectronique comme étant le produit anti-symétrique d'orbitales moléculaires monoélectroniques ( $\varphi_i$ ). Les différentes orbitales moléculaires  $\varphi_i$  ne peuvent plus être construites strictement, c'est pour cela que l'on a recours à l'approximation Linear Combination of Atomic Orbital (LCAO) qui permet de construire  $\varphi_i$  à partir d'une combinaison linéaire d'orbitales atomiques monoélectroniques. A leur tour, les différentes orbitales atomiques monoélectroniques seront approximées afin de faciliter les calculs. Cette dernière approximation va permettre de plus ou moins bien décrire les orbitales atomiques monoélectroniques, et l'on parlera alors de bases.

- Les **méthodes *ab-initio*** se divisent en deux sous familles : les méthodes Hartree - Fock (HF, RHF, UHF, ROHF) (Hartree, 1928 ; Fock, 1930), et les méthodes post Hartree-Fock, (MPn, CAS, ...) (Moller, 1934). La principale différence entre ces deux méthodes est que les interactions électroniques sont négligées dans les méthodes HF et réintroduites dans les méthodes post HF. Ces méthodes ne peuvent être appliquées qu'à des systèmes de quelques dizaines d'atomes pour les méthodes HF et d'une dizaine d'atomes seulement pour les méthodes post HF.
- Les **méthodes semi-empiriques** résultent de la simplification extrême des méthodes HF. Certaines intégrales difficiles à calculer sont remplacées par des termes empiriques ajustés afin de faire correspondre au mieux les résultats calculés avec les grandeurs expérimentales (potentiel d'ionisation, électroaffinité) ou avec les résultats obtenus par des méthodes plus avancées. Généralement, ces méthodes s'appliquent

seulement aux électrons des couches de valence. Les électrons de cœur quant à eux sont traités par le potentiel du noyau qu'ils induisent.

2- La seconde classe de méthodes, apparue plus récemment, traite les électrons comme un nuage de gaz d'électrons qui englobe les noyaux de la molécule. La densité de ce nuage de gaz est représentée par une fonction unique, d'où le nom de **DFT** (Density Functional Theory) (Parr et al. 1989, Kyachko et al. 1990, Becke et al. 1989). Pour des résultats de qualité équivalente, cette méthode est beaucoup plus rapide et permet de traiter quelques dizaines d'atomes dans un temps raisonnable.

Ces méthodes de calcul permettent de fournir des renseignements sur la répartition électronique, les niveaux d'énergies et de décrire plus ou moins bien la liaison chimique. Elles permettent en outre d'identifier les états de transition et donc d'étudier les réactions chimiques. La limitation liée à la taille du système ne permet pas d'analyser des objets tels que des biopolymères, même avec les méthodes semi-empiriques (Dewar et al. 1985, Stewart et al. 1989a, Stewart et al. 1989b).

### II-3-2- Les méthodes de mécanique moléculaire :

C'est la deuxième grande catégorie des méthodes ici présentées. L'énergie de déformation du système est calculée en faisant la somme d'expressions mathématiques qui rendent compte d'un type particulier de déformation. L'ensemble des paramètres des liaisons chimiques, des expressions mathématiques des déformations, ainsi que les paramètres des atomes, est appelé **champ de forces**.

Il faut savoir qu'un champ de forces capable de simuler tous les types de molécules n'existe pas. Chacun est paramétré pour un ou plusieurs types de molécules. De nombreux champs de forces sont disponibles, les principaux étant Amber (Case et al. 1997, Pearlman et al. 1990), Charmm (Brooks et al. 1998, MacKerell et al. 1995 MacKerell et al. 1997), MMFF et CFF (Peng et al. 1997, Hwang et al. 1998, Maple et al. 1998). Amber et Charmm sont paramétrés pour des biopolymères comme les protéines et les acides nucléiques. D'autres champs de forces sont plus généralistes, c'est le cas du champ de forces de Merck (MMFF) qui a été paramétré sur de petites molécules.

La mécanique moléculaire permet de simuler la conformation des molécules ainsi que leurs mouvements. A l'inverse de la description d'une liaison chimique par mécanique

quantique, la liaison chimique est codée dans le système et ne peut être rompue; il est donc impossible d'étudier la réactivité d'une molécule par la mécanique moléculaire.

Deux classes de champs de forces sont distinguées. Les champs de forces de type I ont moins de termes croisés que les champs de type II, plus récents. Néanmoins, les champs de forces de classe I sont toujours très utilisés et donnent d'excellents résultats. Analysons plus en détail le champ de forces Amber (Cornell et al., 1995), qui est paramétré pour les simulations de protéines et d'acides nucléiques. L'énergie de déformation du système est calculée en faisant la somme de termes qui représentent chacun un type d'interaction. Nous disposons dans ce cas de six termes qui apparaissent sur la figure 10.

$$\begin{aligned}
 E_{\text{pot}} = & \sum_b K_2 (b - b_0)^2 + \sum_{\theta} H_{\theta} (\theta - \theta_0)^2 + \sum_{\phi} \frac{V_n}{2} [1 + \cos(n\phi - \phi_0)] \\
 & + \sum \varepsilon [(r^*/r)^{12} - 2(r^*/r)^6] + \sum q_i q_j / \varepsilon_{ij} r_{ij} + \sum \left[ \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right] \\
 & \qquad \qquad \qquad \text{(1)} \qquad \qquad \qquad \text{(2)} \qquad \qquad \qquad \text{(3)} \\
 & \qquad \qquad \qquad \text{(4)} \qquad \qquad \qquad \text{(5)} \qquad \qquad \qquad \text{(6)}
 \end{aligned}$$

Figure 10 : Les différents termes de champ de forces Amber.

Les trois premiers termes permettent de calculer la contribution énergétique de la déformation de la structure chimique. Le premier terme évalue l'énergie d'élongation d'une liaison chimique. Cette contribution suit la fonction dite de « Morse », mais au voisinage de la longueur d'équilibre, elle peut être approximée par une fonction parabolique. Le second terme permet d'évaluer l'énergie de déformation angulaire entre deux liaisons chimiques. La contribution est évaluée suivant le même principe que pour le premier terme. Le troisième terme calcule l'énergie de déformation apportée par les torsions des liaisons chimiques en utilisant des fonctions périodiques. Il faut noter que ces termes permettent également d'évaluer la contribution de la déformation de la géométrie des différentes liaisons de l'atome, et de maintenir ainsi la chiralité des centres d'asymétrie. Les trois derniers termes donnent la contribution non covalente de l'énergie du système. Le terme quatre représente la contribution des interactions de type Van der Waals. Le terme cinq correspond à l'apport électrostatique vis-à-vis de l'énergie interne du système. Cet apport est défini par la loi classique de l'interaction coulombienne. Afin de limiter la taille des calculs, une distance limite de l'interaction est paramétrée. Cette distance limite (cut-off) ne permet pas de décrire correctement les interactions à longue distance : pour y remédier, des algorithmes de sommation de type Ewald comme « Particle Mesh Ewald » (York et al. 1993, York et al.

1994) ont été rajoutés dans les simulations. Les deux avant derniers termes sont critiques pour obtenir une bonne description des interactions non covalentes. Enfin, le terme six permet de calculer la participation des liaisons hydrogène du système. La contribution de ce terme est limitée à  $0.5 \text{ kcal.mol}^{-1}$ , l'essentiel de la contribution énergétique d'une liaison hydrogène provenant du terme électrostatique.

En mécanique moléculaire, il est possible de placer le système à étudier dans des conditions périodiques. Un volume est défini autour du système, et sera dupliqué dans toutes les directions de l'espace. Le volume peut se présenter sous la forme d'un parallélépipède ou d'un parallélépipède tronqué. Cette méthode décrira donc une matrice infinie à partir d'un ensemble fini. Elle peut également être appliquée à la simulation d'un solvant autour de l'objet, en tenant compte des conditions de pression.

La mécanique moléculaire permet de simuler les mouvements d'une molécule sur un temps très court. On parle de dynamique moléculaire (paragraphe III-1-1 de ce chapitre). Il existe deux grandes familles de simulations de dynamique moléculaire : Newtonnienne et de Monte-Carlo.

- La **simulation newtonnienne** est basée sur la loi de Newton (équation 1). Les équations de déplacement sont résolues pour le système complet, en respectant l'énergie potentielle décrite dans le champ de forces. Les forces appliquées sur l'atome  $i$  (en fonction de  $V$  qui est l'énergie potentielle issue du champ de forces en accord avec ses coordonnées) sont données par l'équation 2.

$\mathbf{F}_i(t) = m_i \mathbf{a}_i(t)$	$-\frac{\delta V}{\delta \mathbf{r}_i} = m_i \frac{\delta^2 \mathbf{r}_i}{\delta t_i^2}$
---	--

Equations 1 et 2 : Equation de Newton et équation de mouvement pour l'atome  $i$

Ces équations peuvent être résolues par des méthodes d'intégration numérique. Ce type de méthodes permet de calculer les différentes coordonnées ainsi que la vitesse des différents atomes à l'instant  $t + \Delta t$  à partir de l'instant  $t$ . Le choix du pas d'intégration ( $\Delta t$ ) est critique pour la qualité des calculs qui sont également fonction de la méthode d'intégration choisie. On compte plusieurs méthodes telles que : Verlet – Leafrog qui est la plus utilisée (Verlet et al. 1967), Verlet Velocity, Adams-Bashforth-Moulton 4

(ABM4) ou Runge - Kutta 4. Le temps d'intégration est de 1 femtoseconde (fs) pour les deux premières méthodes, tandis que 0.1 fs constitue un pas d'intégration correct pour la dernière méthode.

Afin de simuler l'agitation thermique, de l'énergie est rajoutée au système. Cette énergie supplémentaire va s'équilibrer entre un terme d'énergie cinétique et un terme d'énergie potentielle. L'énergie totale étant gardée constante. La vitesse des atomes est directement reliée à la température du système par le mouvement brownien des particules. Cette température est contrôlée par un algorithme qui agit sur la vitesse des atomes dans le système. Certains termes doivent être gardés constants afin d'intégrer les équations de mouvement. Le choix de ces termes constants crée autant de types de dynamique moléculaire. On distingue :

- énergie et volume constants (NVE)
- température et volume constants (NVT)
- température et pression constantes (NPT)
- pression et enthalpie constantes (NPH)

Les simulations les plus utilisées sont celles qui correspondent soit à une simulation isochore (NVT), soit à une simulation isobare (NPT).

- Les **simulations de Monte-Carlo** sont basées sur le hasard. En effet, les conformations sont générées au hasard, puis l'énergie interne est évaluée par le champ de forces. Le conformère formé est accepté lorsque son énergie satisfait la distribution d'énergie gaussienne. La méthode de test la plus utilisée est celle de Métropolis. Ce type de dynamique moléculaire est utilisé pour générer des conformations car il n'y a aucune relation de mouvement entre deux conformères.

De nos jours, de nouveaux algorithmes tel que « shake » (Van Gunsteren et al. 1977) permettent d'avoir des pas d'intégration de 2 fs, permettant de rallonger le temps de simulation. Quelques équipes, particulièrement bien dotées en super-calculateurs, ont réussi à simuler quelques microsecondes. La durée moyenne des simulations est actuellement de quelques nanosecondes. Elle a fortement augmenté ces dernières années grâce à une nouvelle classe de super-calculateurs constitués par des agrégats de PC utilisant un « operating

system » parallèle de type Unix. Ces systèmes présentent un rapport qualité-prix très bas, et ont su s'imposer comme le standard des calculs de haute performance à prix raisonnable.

### II-3-3- Les méthodes mixtes :

Cette dernière classe de méthodes permet de modéliser les réactions chimiques dans des systèmes trop importants pour être traités par des méthodes de mécanique quantique pure. Elle combine les deux méthodes de calcul pour un même système. Celui-ci est divisé en deux régions: le cœur et la périphérie. Le cœur comprend généralement le site d'interaction du ligand et il est traité par des méthodes quantiques semi-empiriques. La périphérie, qui correspond à la région la plus grande, est quant à elle, traitée par la mécanique moléculaire. L'énergie totale du système peut être séparée en trois termes distincts comme en témoigne la figure 11 qui présente la zone bleue traitée par mécanique moléculaire, et la zone jaune par mécanique quantique.

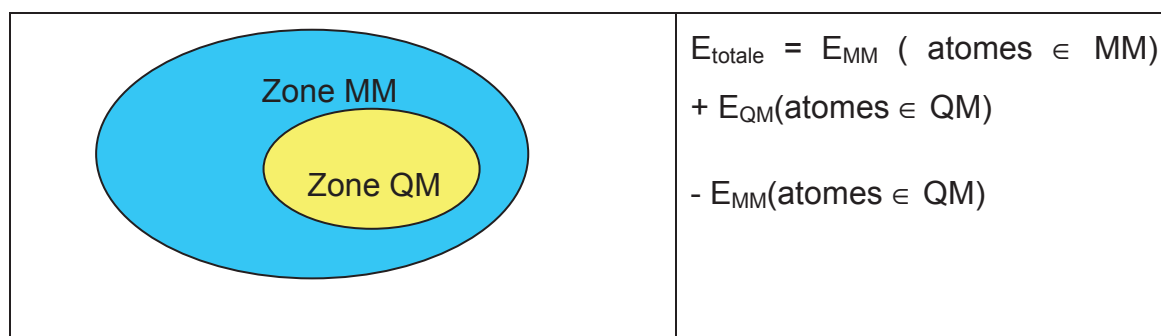


Figure 11 : Termes définissant l'énergie totale du système dans un système mixte QM/MM

Les termes peuvent être calculés par le programme de mécanique quantique pour le premier, et par le programme de mécanique moléculaire pour le second. Le troisième terme est une fonction de potentiel de couplage entre les deux régions. Les différentes liaisons chimiques reliant les deux systèmes peuvent être soit :

- rompues, et un atome d'hydrogène sera inséré au milieu (méthode du « Link atom »).
- gelées via leurs orbitales p (Laboratoire de Chimie Théorique de Nancy)

Plusieurs programmes tels que Gaussian, NWchem, ou encore ADF permettent de créer des systèmes décrits par la méthode de QM/MM. La méthode QM/MM sera particulièrement utile pour simuler des réactions chimiques sur des biopolymères telles que des protéines.

# III- Apport de la modélisation moléculaire dans la découverte de nouveaux composés :

La majeure partie des études qui touchent la « science du vivant » a pour objectif de déterminer ou d'optimiser un ligand vis à vis d'un récepteur. Deux grandes voies peuvent être envisagées en fonction de la connaissance ou non du récepteur.

## III-1-Approche sans récepteur :

Nous l'avons vu dans le paragraphe II, il existe deux grandes méthodes pour résoudre la structure du récepteur. En réalité, les études de cristallographie ou de RMN-2D sont longues et coûteuses. De plus, le récepteur d'une molécule n'est pas toujours connu. Généralement dans ce cas, nous disposons d'un ensemble de molécules qui présentent toutes une propriété commune, et la recherche s'oriente vers la découverte d'un nouvel analogue plus actif. Il existe de nombreux outils permettant au modélisateur de réaliser ce type d'étude.

### III-1-1-Analyse conformationnelle et flexibilité moléculaire :

La première tâche consiste à bien connaître les différentes molécules dont nous disposons. Pour cela, il faut déterminer leur flexibilité ainsi que leurs conformations les plus probables. Ces informations sont extrêmement importantes pour la suite de l'étude. Afin de connaître la flexibilité d'une molécule, deux grandes familles de méthodes sont disponibles.

- La première méthode est basée sur la **dynamique moléculaire**. Un modèle de la molécule est créé puis son énergie interne est minimisée. A partir de la conformation ainsi obtenue, une simulation de dynamique moléculaire en phase aqueuse décrite explicitement ou non sera effectuée. La simulation de dynamique moléculaire va échantillonner une conformation à des intervalles de temps fixes. Les différents conformères seront analysés et les structures les plus proches seront rassemblées dans différentes familles. Cette méthode permet d'explorer l'espace conformationnel de la molécule accessible grâce à l'addition d'énergie. Il est également possible de réaliser des cycles de fort chauffage et de refroidissement (qui sont appelés « aneling ») afin de permettre à la molécule de passer des barrières conformationnelles d'énergies élevées. Dans ce cas, la structure archivée sera celle de plus basse énergie, c'est-à-dire à la température la plus basse. Plusieurs



dizaines de cycles sont ainsi effectuées afin d'explorer un espace conformationnel le plus vaste possible.

Une autre approche consiste à réaliser non plus une dynamique moléculaire utilisant les lois de Newton, mais une dynamique moléculaire basée sur le hasard (méthode de Monte-Carlo). Ce type de simulation génère un grand nombre de conformères couvrant un espace conformationnel plus étendu que dans un échantillonnage réalisé à partir d'une simulation de dynamique moléculaire Newtonienne.

- La seconde méthode (Torsional Search) s'appuie sur une recherche systématique, en faisant varier pas à pas l'angle de torsion de chaque liaison simple non cyclique du système. A chaque torsion, l'ensemble de la molécule (sauf la liaison considérée) sera minimisé en énergie afin de rationaliser la structure. Si cette méthode permet d'explorer l'espace conformationnel, elle trouve rapidement ses limites dans le nombre de liaisons chimiques. En effet, le nombre de conformations possibles croît de façon exponentielle avec le nombre de liaisons.

### III-1-2-Relations structure - activité quantitative (QSAR) :

On trouve dans la littérature du XVII<sup>ème</sup> siècle des allusions à la justification d'un remède en fonction des relations propriété - activité. Plus sérieusement, les études de relations structure - activité ont été menées dès que les scientifiques ont disposé de la structure des molécules qu'ils synthétisaient. Les **QSAR** représentent une amélioration significative. Elles ont pour but de découvrir, grâce à des méthodes statistiques, une équation permettant de décrire l'activité de la molécule en se basant sur sa structure chimique.

Le QSAR se décompose en deux étapes bien distinctes. La première revêt un aspect plus chimique car elle comprend la collecte, la modélisation et la description des molécules étudiées. La seconde étape correspond à la recherche de la relation capable de relier les descripteurs à l'activité de la molécule. Les premières études furent réalisées par Hansch qui utilise trois descripteurs, et qui établit une relation basée sur une combinaison linéaire de ces trois descripteurs (Hansch, 1969).

Examinons plus en détail la première étape. La collection de molécules est une phase très difficile, car il nous faut disposer d'un nombre suffisant de molécules pour réaliser une analyse statistique. De plus, les molécules doivent présenter une grande diversité du point de

vue de leurs activités, allant de molécules très actives jusqu'à des molécules ne présentant aucune activité. De même, les structures doivent être les plus diverses tout en restant des ligands potentiels du récepteur. Des erreurs peuvent être induites par une collection de molécules se décomposant en plusieurs sous-familles assez proches. Dans ce cas, il y a de fortes redondances d'information au sein de chaque sous-famille. Il sera alors plus judicieux de subdiviser l'étude et de développer une relation pour chaque sous-famille.

Une fois définie la collection de molécules, les différents modèles sont soumis à des programmes qui servent à décrire la molécule sous forme d'une valeur numérique unique. Le chiffre obtenu correspond à un calcul basé sur la structure chimique. Si Hansch a travaillé avec trois descripteurs, les programmes de QSAR disposent aujourd'hui d'une à plusieurs centaines de descripteurs différents. Nous ne développerons pas ici la méthode de calcul de chaque descripteur mais notons qu'ils sont répertoriés en plusieurs catégories et pourront être:

- topologiques
- conformationnels
- spatiaux
- électroniques
- basés sur la mécanique quantique
- thermodynamiques

Avec la multiplication de la puissance de calcul des stations de travail, de nouvelles méthodes de QSAR sont apparues : les méthodes 3D QSAR.

Pour les 3D QSAR, il ne faut plus voir un descripteur comme un outil permettant la description d'une molécule par un chiffre unique, mais plutôt comme une sonde qui renvoie un résultat en fonction des coordonnées du modèle. Dans ce cas, il peut y avoir des milliers de descripteurs du même type. Leur répartition sera spécifique à chaque méthode. Deux grandes classes se distinguent: la méthode CoMFA pour « Comparative Molecular Field Analysis » (Bohacek et al. 1992, Kroemer et al. 1995) et RSA pour « Receptor Shape Analysis » (Hahn 1995, Hahn et al. 1995). Une étape préliminaire reste nécessaire : il faut superposer les différentes molécules, soit par une superposition rigide, soit par une superposition flexible.

La méthode **CoMFA** permet de placer des descripteurs évaluant l'énergie d'interaction (en terme stérique et électrostatique) entre une molécule et un fragment chimique jouant le rôle de sonde. La distribution des sondes se fait selon une grille tridimensionnelle qui entoure la collection de molécules superposées. La grille peut être de forme rectangulaire ou sphérique. Notons que la méthode **CoMSIA** est basée sur le même principe mais l'énergie évaluée prend compte également des termes hydrophobes et relatifs à la liaison hydrogène.

La méthode **RSA** présente des points communs avec la méthode CoMFA. A partir d'un jeu de molécules les plus actives, une surface d'accès au solvant unique est créée. Par la suite, sera appliquée, pour chaque polyèdre de surface, la valeur des potentiels électrostatiques et stériques. La surface ainsi générée devrait correspondre à la surface interne du récepteur.

Dans toutes les méthodes QSAR, les descripteurs sont rassemblés par molécules. Les descripteurs sont disposés en colonnes tandis que les molécules sont en lignes, ce qui aboutit à une matrice. L'activité des molécules est intégrée et servira à la calibration du système prédictif. A chaque étape, l'objectif consiste à trouver la relation corrélant les descripteurs à l'activité des molécules. Lorsque la collection est suffisamment grande, elle est scindée en deux jeux. Le premier, qui est aussi le plus gros, va servir à découvrir la relation : on parle du jeu de calibration. Le second va servir à tester la relation précédente : le système va prédire les différentes activités qui seront alors confrontées aux activités réelles. Si le jeu est trop petit, d'autres méthodes peuvent être employées telle que la validation croisée. Cette méthode permet de prédire l'activité de  $i$  molécules de l'ensemble total (constitué de  $n$  molécules) en utilisant comme groupe de calibration les  $n-i$  molécules. Les calculs sont ensuite répétés jusqu'à prédire les  $n$  molécules. Le nombre  $i$  pourra varier de 1 à  $n/5$  soit 20% des molécules.

Les méthodes statistiques se sont étoffées au cours du temps, allant de la méthode la plus simple, celle de régression linéaire, jusqu'aux algorithmes génétiques ou aux réseaux neuromimétiques.

A titre d'illustration, et parmi de nombreux exemples, les méthodes QSAR ont permis de mettre au point une nouvelle famille d'hypertenseurs. L'étude QSAR portait sur l'optimisation d'une famille de molécules conçues à partir d'un leader peptidique (Duncia et al., 1990).

Le QSAR amène à découvrir une relation structure-activité, mais il permet également de prédire l'activité d'une molécule qui n'a pas encore été synthétisée. De plus, l'analyse de la nature des descripteurs (dans le cas classique) ou de la position des principaux descripteurs (dans le cas du 3D QSAR) apporte des renseignements très importants pour l'optimisation d'une molécule.

Cette méthode est particulièrement utilisée dans l'industrie pharmaceutique mais aussi dans les industries chimiques classiques. Citons l'exemple des relations structure-activité qui ont été choisies pour optimiser le choix du surfactant à ajouter à un ciment.

### III-1-3- Les pharmacophores :

Cette méthode est plus récente. Un **pharmacophore** se définit comme un ensemble de fragments ou de fonctions chimiques dans une configuration spatiale particulière, et qui sont nécessaires à l'activité de la molécule. Il est déterminé à partir d'une série de molécules actives. Plusieurs types de fonctions chimiques sont distinguées :

- accepteur de liaisons hydrogène
- donneur de liaisons hydrogène
- charge positive
- charge négative
- groupe hydrophobe aliphatique
- groupe hydrophobe aromatique
- groupe ionisable

La méthode va donc déterminer un ensemble de fragments ou de fonctions chimiques représentatives de l'activité des modèles. Un pharmacophore peut être utilisé soit pour réaliser un modèle de l'activité, soit pour comprendre le mode d'interaction de la molécule vis à vis de son récepteur.

L'utilisation principale consiste à rechercher, au sein d'une base de données structurale, les molécules pouvant correspondre au pharmacophore étudié.

Cette méthode a permis entre autre de mettre au point une nouvelle famille d'anti-inflammatoires. L'étude portait sur des molécules anti-inflammatoires existantes et a aboutit à un modèle pharmacophore basé sur 4 fonctions chimiques et un groupement hydrophobe (Glen et al. 1995)

La méthode a été améliorée en générant des zones d'exclusion. Ceci permet d'éliminer toutes les molécules possédant un ou plusieurs atomes appartenant à une zone d'exclusion. Ces zones sont utilisées en grand nombre et créent la surface d'accès au solvant de la cavité du récepteur. Cette surface est générée à partir des molécules actives, mais en supposant que les molécules interagissent correctement avec le récepteur, la surface sera alors assez proche de la surface réelle du récepteur. Ainsi, la méthode qui utilise les zones d'exclusion se rapproche de la méthode 3D QSAR « receptor ».

### III-1-4- La méthode de De novo design sans récepteur :

Le **De novo design** est une technique très complète puisqu'elle fonctionne soit avec la structure du récepteur, soit avec une collection de molécules actives (c'est-à-dire sans récepteur). Détaillons ici l'approche sans récepteur.

Le mode sans récepteur s'appuie sur une collection de molécules actives qui seront superposées. Par la suite, un traitement est effectué sur leur superposition. Il consiste en l'analyse de l'ensemble des molécules afin de dégager les zones acceptrices et donneuses de liaisons hydrogène, ainsi que les régions hydrophobes. Ces dernières peuvent être visualisées par un maillage de traits. Les régions acceptrices de liaisons hydrogène sont pour leur part illustrées sous forme de liaisons chimiques HN dont la densité varie selon les paramètres fixés. Les zones donneuses de liaisons hydrogène sont codées par des liaisons HO.

Après ce premier traitement, les différentes molécules qui ont servi à bâtir le modèle des interactions du site récepteur ne seront plus utilisées. Le programme va chercher dans des bases de fragments structuraux, soit commerciales (ACD, Cambridge...) soit maison, un fragment capable d'interagir avec les différentes liaisons hydrogène et les régions hydrophobes qui ont été définies au niveau du site d'interaction. Le programme va ainsi sélectionner les meilleurs fragments et les placer dans la cavité. L'ensemble des solutions proposées sera rangé par affinité croissante. Le choix du meilleur fragment est fait par le modélisateur. Lorsque la cavité permet de positionner d'autres fragments, le processus est répété jusqu'à ce que la cavité soit remplie. Dans une seconde phase, il faut préciser au

logiciel que l'on ne cherche plus de nouveaux fragments, mais un lien entre deux fragments. Différents liens vont être proposés. Une fois le choix fait, la nouvelle molécule est créée. Elle sera évaluée, et une estimation de l'affinité sera calculée.

Bien qu'elle présente de nombreux avantages, cette méthode est décriée. Elle est très rapide car les calculs nécessaires ne sont pas très lourds. La principale critique touche le choix du fragment laissé au modélisateur. De plus, ce choix bloque certaines autres possibilités. Le parcours optimal de construction de la molécule est difficile et de nombreux essais sont nécessaires.

Cependant, cette méthode est aujourd'hui employée par de nombreuses entreprises pharmaceutiques, et a abouti à la mise au point de nombreuses anti-protéases dirigées contre le VIH-1.

## III-2- Approche avec récepteur

### III-2-1- La méthode d'arrimage (Docking)

Cette méthode (Bohm et al., 1996 – Bohm et al., 1992) qui permet de simuler l'approche et l'interaction entre un ligand et son récepteur, comme un bateau vis-à-vis du port, d'où le nom de docking. A l'origine, le ligand était rigide. Les logiciels proposent actuellement une géométrie flexible du ligand. Il existe plusieurs façons de prendre en compte cette flexibilité. La principale est utilisée par le logiciel SURFLEX-DOCK qui découpe le ligand en différents fragments. Le fragment central est positionné dans le site, puis la molécule est reconstruite en faisant varier les angles de rotation à chaque connexion entre les fragments. La position de chaque fragment ajouté est donc optimisée par rotations, afin d'avoir le plus d'interactions avec le site. Une fois que la molécule est totalement reconstituée, les interactions entre la molécule et le récepteur seront évaluées une fonction de notation (scoring). Les fonctions de scoring, tel que CScore (chapitre 4, paragraphe II-3-1), permettent d'avoir un résultat assez pertinent sur l'affinité d'une molécule pour son site.

Cette méthode permet de valider ou d'optimiser le ligand vis-à-vis de son récepteur, et ce avant même de le synthétiser. Elle est utilisée dans toutes les études de modélisation moléculaire.

Si l'étude des interactions entre une molécule et son récepteur est possible, le calcul ne permet pas d'accéder aux données thermodynamiques. De plus, le docking ne prend pas en compte l'énergie de déplacement du solvant nécessaire à la molécule pour pénétrer dans le site récepteur.

### III-2-2- Méthode du De novo design avec récepteur :

La méthode de **De novo design** (Bohm et al., 1996 – Bohm et al., 1992) permet de concevoir un nouveau ligand, soit à partir d'une superposition de molécules (nous l'avons développé dans la partie III-1-2-), soit avec la structure du récepteur. Le programme va analyser la cavité du récepteur et détailler les zones hydrophobes ainsi que les régions acceptrices et donneuses de liaisons hydrogène. La représentation graphique des différentes zones est identique à celle du mode sans récepteur. Le programme va ensuite rechercher dans sa base de fragments celui pouvant satisfaire le plus d'interactions dans une région de la cavité. Chaque fragment est ainsi évalué, et une note d'interaction est renvoyée. Il faut préciser que la fonction de notation prend en compte l'aspect directionnel des sites donneurs et récepteurs de liaisons hydrogène. Le résultat du calcul se présente sous la forme d'un tableau de fragments rangés par note décroissante. Pour chaque fragment, la meilleure position dans la cavité est retenue. Cette opération est répétée jusqu'à remplir entièrement la cavité. Un mode de création de liens est alors disponible et permet de proposer une unité de jonction entre les deux fragments choisis. Une fois réunis les différents fragments, le nouveau ligand obtenu est noté afin de déterminer le ou les meilleurs candidats à synthétiser.

Les reproches adressés à cette méthode ont déjà été exposés dans la section III-1-4. Une critique supplémentaire pourrait être faite : le récepteur reste fixe. Or, nous savons que les chaînes latérales bordant le site actif du récepteur sont mobiles et peuvent changer de position en fonction du ligand.

L'utilisation de cette méthode a permis de mettre au point de nombreux inhibiteurs de la protéase du VIH-1. Dans le tutoriel du logiciel Ludi (de novo design), les cinq meilleurs ligands proposés sont devenus des inhibiteurs commercialisés par Roche, Hoest, Glaxo-Wellcome... Ceci démontre l'importance pour les industries pharmaceutiques de créer leurs propres bases de fragments.

### III-2-3- La méthode de « Multiple Copy Simultaneous Search (MCSS) » :

Le **MCSS** (Miranker et al., 1991, MacKerell et al., 1998) est une méthode plus récente que la méthode de novo. Elle permet également de concevoir des ligands pour un récepteur, sans recourir à des molécules « leader ». Elle consiste à prendre un fragment et à le répliquer dans la cavité en plusieurs centaines d'exemplaires. Les complexes « ligand-récepteur » sont minimisés, mais la structure du récepteur est gardée figée et il n'y a pas d'interaction entre les différentes copies du ligand. De plus, toutes les copies sont minimisées simultanément et, lorsque deux copies sont très proches dans l'espace, l'une des deux sera éliminée. Cette méthode permet d'explorer, pour un ligand donné, toutes les positions possibles au sein de la cavité et de retenir les meilleures. Le calcul doit être réalisé plusieurs fois afin de retenir les meilleurs candidats. Après avoir validé le choix des différents fragments et leur emplacement, le modélisateur conçoit le ou les liens nécessaires au raccordement des différents éléments constituant le ligand.

### III-2-4- Calcul d'énergie libre :

Ce type de méthode est peu utilisé car il nécessite de longues simulations sur des systèmes complexes. Deux grandes familles existent : la Free Energy Perturbation pour perturbation d'énergie libre (Case et al. 1997), et la CMC/MD pour Chemical Monte-Carlo / Molecular Dynamics (Kollmann 1993).

La méthode **FEP** permet de calculer la variation d'énergie libre au cours de la variation d'un paramètre. Ainsi, il est possible de faire varier une contrainte géométrique telle qu'une distance, un angle ou une torsion, ou la nature d'un atome. Le paramètre fluctue entre la valeur 0 et 1. Des atomes « fantômes » peuvent également être définis: ils existent et disposent d'une connectivité, mais ils n'ont aucune interaction de Van der Waals ni électrostatique avec les autres atomes du système. Dans ce dernier cas, la topologie ainsi que la nature des atomes sont codées pour l'état 0 et pour l'état 1 dans le même fichier.

La variation de paramètres géométriques peut servir à étudier l'interaction entre deux molécules. L'un des meilleurs exemples est celui de C.Chipot (Chipot et al. 1996), qui a analysé l'interaction entre deux plateaux de bases d'acides nucléiques. Il a fait varier la distance entre les deux plateaux et a pu calculer l'énergie d'interaction.



La variation de la nature des atomes est, elle, plus utilisée, car elle permet d'accéder à la variation d'énergie libre ( $\Delta G$ ) d'interaction entre un ligand et son récepteur. Plusieurs équipes (Pearlman et al. 1989, Singh et al.1999) ont réalisé une série de simulations afin de déterminer le  $\Delta G$  d'affinité d'une molécule vis à vis de son récepteur. Pour cela, il est nécessaire de générer des apparitions et des disparitions du ligand. L'apparition d'une molécule revient à simuler la variation de la nature des atomes en partant des atomes fantômes vers les atomes réels ; la disparition étant l'inverse. Pour calculer le  $\Delta G$  entre le ligand et le récepteur, il faut prendre en compte les molécules de solvant qui remplacent le ligand absent. Dans ce cas, il est nécessaire de réaliser les simulations suivantes d'apparitions et de disparitions :

- solvant de la cavité dans le récepteur
- solvant de la cavité dans le solvant
- ligand dans la cavité
- ligand dans le solvant

Pour chaque cas, il faut simuler l'apparition et la disparition afin d'évaluer l'erreur réalisée lors de la simulation. La sommation des quatre simulations citées permet de déduire le  $\Delta G$  d'interaction du ligand en phase aqueuse vis-à-vis de son récepteur. Une étude a montré que ce type de calcul, effectué sur une série de molécules, avait une très bonne corrélation avec les analyses expérimentales (Becke et al. 1989).

La méthode du **CMC/MD** est beaucoup plus récente. Elle nécessite moins de puissance de calcul car elle ne fait pas appel à la variation d'un paramètre, mais évalue le  $\Delta\Delta G$  entre deux états. Pour cela, deux régions sont créées : la région dite « Monte-Carlo » et la région qui l'englobe. Le programme va ensuite, en utilisant une dynamique de Monte – Carlo, évaluer l'énergie libre d'interaction entre les deux régions. Cette technique a été utilisée par l'équipe du Pr Kollman (Kollman 1993) afin d'obtenir les énergies libres d'affinité entre des analogues du TIBO (inhibiteur) et de la reverse transcriptase du VIH-1. L'étude conclut à une bonne corrélation entre les expérimentations et les simulations.

### III-3- Modélisation des structures 3D à haut débit : le logiciel

#### Modeome :

Ce logiciel a été développé au sein de notre institut par Nicolas Garnier sous la direction du Dr. Emmanuel Bettler et nous reprenons ici le principe général de Modeome avec lequel nous avons créé de nombreux modèles utilisés dans ce travail de thèse.

**Modeome3D** est un système automatique, à grande échelle, de génération de modèles tridimensionnels par homologie de séquences (paragraphe II-2). Il est composé d'un module de génération des modèles, d'un module de gestion et d'une interface web. Nous ne présenterons ici que le principe de génération des modèles.

Modeome3D prend en entrée un fichier regroupant l'ensemble des séquences protéiques d'un organisme au format « fasta » et une banque de séquences de protéines de structures connues (en général la PDB) pour effectuer les recherches de similarités. En sortie, une base de données relationnelle regroupant les modèles 3D et un ensemble d'informations qui y sont rattachées est créée. L'ensemble des scripts du système Modeome3D a été développé par N. Garnier en intégrant des programmes pré-existants tels que **Geno3D** (développé à l'IBCP et employé pour générer les modèles), Sov, Fit, CE (Combinatorial Extension) ou DSSP (Dictionnaire de Structures Secondaires de Protéines) pour aboutir à la construction des modèles et au chargement de la banque de données.

La modélisation comporte deux étapes clefs : 1- La **sélection de l'empreinte** à utiliser et 2- L'**alignement** de l'empreinte avec la protéine cible. Ces deux étapes conditionnent la qualité des modèles générés. La sélection des empreintes est absolument essentielle car plus l'empreinte sera proche de la protéine cible, meilleure sera la qualité du modèle. Ainsi, la première étape consiste en une recherche de similarité dans la banque de données de séquences possédant des structures connues pour chacune des protéines à modéliser. Cette recherche se déroule en deux temps : une recherche de similarité globale grâce à Fasta pour rechercher les empreintes les plus évidentes, partageant plus de 50% d'identité avec la séquence cible. Lorsque aucune empreinte n'est trouvée, une recherche locale avec PsiBlast en sept itérations par défaut est réalisée. Les empreintes potentielles ainsi trouvées sont alors filtrées en fonction de différents critères ajustables. Une ou plusieurs empreintes pour une protéine sont alors obtenues. Les alignements de l'empreinte et de la protéine cible sont

extraits du fichier de sortie Fasta ou PsiBlast et donnés en entrée à Geno3D. Il procède alors en quatre étapes :

- Il réalise une recherche d'homologues dans les séquences de la PDB par PsiBlast. La (ou les) séquence(s) homologue(s) ayant au moins 30% d'identité convenable est (sont) alignée(s) avec la protéine à modéliser.
- Chaque alignement, par paire, de l'empreinte et de la protéine cible est donné en entrée à Geno3D, qui va extraire les contraintes de distances pour les résidus identiques (au moins 20) et construire trois modèles tridimensionnels.
- Une étape de dynamique et de minimisation d'énergie permet de relâcher les contraintes et d'optimiser la structure.
- La validation des modèles se fait par superposition structurale de l'empreinte avec le modèle en utilisant le programme CE.

Le modèle finalement retenu sera celui dont l'énergie de minimisation est la plus faible.

Enfin, les structures secondaires sont déterminées à partir de la structure tridimensionnelle par DSSP et ajoutées au fichier PDB de sortie.

Le temps de construction des modèles est relativement long, variant de quelques minutes à plusieurs heures en fonction de la taille des protéines. Seuls les modèles de plus de 20 résidus sont construits. Il faut préciser que Modeome 3D est capable de générer également des oligomères permettant ainsi de modéliser la forme biologique multimérique active lorsque celle-ci existe.

### III-4-Le logiciel SuMo (Surfing the Molecules) :

Cette nouvelle approche, développée au sein de notre institut (Jambon *et al.* 2003), fait l'objet du chapitre trois. Bien qu'initialement SuMo était très orienté « bioinformatique », les études que nous avons mené nous ont prouvé son potentiel en tant qu'outil de modélisation moléculaire et il peut tout à fait s'inscrire dans une stratégie de conception rationnelle de médicaments.

# Chapitre 3 : Le logiciel SuMo – Optimisation et nouveaux paramètres

## I- Principe du logiciel SuMo

Le logiciel SuMo, « Surfing the Molecules », a été développé au sein du laboratoire de Bioinformatique et RMN Structurales, à l'IBCP en 2000. Il a fait l'objet d'une thèse rédigée par Martin Jambon sous la direction de Christophe Geourjon.

SuMo est un logiciel de comparaison de structures tridimensionnelles. S'intéressant uniquement à la surface des protéines, il peut être assimilé à un outil d'analyse des « peaux » protéiques c'est-à-dire d'étude de l'interface vis-à-vis du solvant ou d'une autre protéine. La « peau » est caractérisée par la nature des acides aminés qui la compose mais également par les repliements qu'elle présente. L'objectif principal de SuMo est de pouvoir retrouver des régions communes entre plusieurs protéines. Il est possible de comparer une protéine entière, ou seulement une zone d'intérêt, tel qu'un site d'interaction, à un ensemble plus ou moins important de protéines. Bien entendu, les calculs concernant des protéines entières comparées à plusieurs milliers de protéines, seront beaucoup plus lourds que si l'on ne veut analyser qu'un site précis et restreint en terme de taille.

Ce chapitre détaillera les différents aspects techniques de SuMo après un bref rappel du principe.

Nous créons des bases de données, de taille plus ou moins importante (allant de quelques protéines à plusieurs milliers), contenant les modèles tridimensionnels issus de la PDB ou de Modeome. Chacun de ces modèles verra sa surface décrite par ce que l'on appelle les « objets SuMo », dont la répartition dans l'espace dépendra de la conformation de l'interface étudiée, et qui coderont les acides aminés présents. A titre d'exemple, un acide aminé de type lysine sera codé par trois objets : un centré sur l'atome d'azote de la fonction amine terminale de la chaîne

latérale, un placé sur le carbonyle du squelette peptidique et enfin le dernier sur l'azote de la liaison peptidique. Ainsi, tous les acides aminés de l'interface sont décrits par le programme et stockés dans les bases de données. Par la suite, l'utilisateur va définir une requête, qu'on appellera « query », correspondant à ce qu'il souhaite comparer. Cela peut être une protéine complète contre toutes les protéines de la base, ou certaines parties d'une protéine contre certaines parties des protéines d'une base. Les requêtes sont très modulables. Le logiciel va ensuite effectuer les comparaisons demandées en recherchant les objets SuMo communs en terme de nature et de position spatiale entre la référence, c'est-à-dire la requête ou « query », et les modèles de la base. Les objets SuMo sont donc des descripteurs chimiques de la surface des protéines qui prennent en compte la conformation de la cible.

Les possibilités offertes par ce logiciel sont extrêmement variées puisque, dès lors que l'on possède un modèle 3D d'une protéine, sa surface sera décrite par les objets SuMo et pourra alors être comparée à la base de données. En application de ce travail de thèse, outre les phases purement liées au sujet des hypersensibilités allergiques, nous avons mené plusieurs études qui démontrent la diversité des analyses que ce logiciel peut effectuer. Cependant, la question sera toujours : « Sur cette protéine, quelle(s) région(s) sera (seront) la (les) même(s) que sur une autre ou un ensemble d'autres protéines ? ».

Nous allons brièvement présenter les différentes études développées et les chapitres 4,5 et 6 détailleront les résultats se rapportant à cette thèse.

Une de nos collaborations industrielles concerne le domaine de l'immunologie et plus précisément un ensemble d'IgE qui reconnaît spécifiquement une protéine ainsi que quatre peptides issus de la dégradation de celle-ci. Le partenaire industriel ignore quelle partie de la protéine est reconnue par une IgE spécifique, capable d'interagir également avec certaines séquences peptidiques issues de la dégradation de cette protéine. Ces peptides étant très différents de la protéine entière, une approche classique par alignement de séquences n'a donné aucun résultat. Ainsi, nous nous sommes appuyés sur le logiciel SuMo afin de chercher les zones communes entre la protéine et les peptides. Dans un premier temps, nous avons construit les modèles des peptides à partir de structures déposées à la PDB puis nous les avons soumis à une simulation de dynamique moléculaire sur 10 nanosecondes en phase de solvant non décrit.

L'utilisation du solvant non décrit, qui tient compte de l'énergie de solvation grâce à la fonction de Poisson-Boltzman, permet d'autoriser le maximum de flexibilité pour des petits objets tels que les peptides. Durant la simulation, 500 conformères ont été échantillonnés pour chacun des quatre peptides. Nous avons donc créé une base de 2000 peptides et nous avons demandé à SuMo de retrouver les points communs entre la protéine de départ qui nous sert ici de référence et les 2000 entrées de la base. Nos calculs aboutissent à une zone unique de superposition des objets SuMo, ce qui montre que tous les peptides reconnaissent cette même zone, qui est de ce fait impliquée dans l'interaction spécifique avec les IgE.

Une autre étude concerne les allergènes, et plus précisément ceux de l'arachide. En collaboration avec le laboratoire d'immunologie de Lyon Sud, nous avons cherché à expliquer les réactivités croisées cliniquement observées entre différents allergènes mais également à retrouver les épitopes d'un allergène donné. Concernant l'arachide, nous avons étudié la protéine Ara h1 d'*Arachis hypogaea*. Nous avons créé une base de 6020 séquences de protéines (issues de multiples sources d'allergènes : aliments, acariens, pollens...) déposées dans la banque de séquences « Uniprot ». Nous avons par la suite généré des modèles tridimensionnels de ces protéines au moyen de notre logiciel de modélisation automatique par homologie (Modeome). 2404 modèles de protéines allergènes ont été construits. Les surfaces sont alors comparées grâce à SuMo afin d'identifier des zones communes entre la protéine de référence (ici Ara h1) et celles de la base de données (2404 structures). Les résultats obtenus montrent des motifs communs notamment entre Ara h1 et des allergènes du lupin, de la lentille et de fruits à coques (noix, noisetier, noix de cajou, pistache...) et sont en accord avec certaines observations cliniques d'allergies croisées. Nous avons également appliqué ce procédé afin d'expliquer le syndrome communément appelé « syndrome bouleau – pomme » impliquant les protéines BetV1 et Mald1. Le logiciel a comparé les surfaces globales de ces deux protéines et a retrouvé les épitopes décrits dans la littérature. Ce principe d'analyse bioinformatique peut être appliqué à tous les allergènes protéiques et permet de prédire les réactions allergiques croisées déclarées et potentielles. Parmi les nombreuses applications, il sera possible de limiter le nombre d'allergènes testés chez les patients (puisque si plusieurs allergènes présentent une forte similitude et surtout des épitopes communs alors un seul test sera suffisant), d'identifier des motifs essentiels reconnus par les structures de l'immunité permettant, entre autres, de construire des protéines hypoallergisantes et d'améliorer la spécificité des tests diagnostiques.

La deuxième collaboration industrielle nous a amené à utiliser le logiciel SuMo afin de réduire les effets secondaires potentiels de molécules visant spécifiquement la protéine kinase C humaine. Nous avons construit pour cela une base de modèles réunissant l'ensemble des protéines kinases humaines. Pour cela, dans un premier temps, nous avons extrait de la base de séquences Uniprot l'ensemble des séquences des protéines kinases humaines et les avons soumises au logiciel Modeome. Nous avons pu construire environ 3000 modèles en partant des 6000 séquences. Le site catalytique de la protéine kinase C alpha sert de référence et il a été comparé à l'ensemble des modèles de la base. Les protéines dont le site présente le plus d'objets communs avec la référence sont les mieux classées. Ce sont ces protéines qui ont le plus de chance de fixer la molécule dont la conception a été optimisée pour la PKC alpha (notre référence). L'approche par le logiciel SuMo s'intègre donc très bien dans l'optimisation de la conception de molécules.

Avant de rentrer dans le détail du logiciel, rappelons que cette partie du travail de thèse n'a pas pour objectif, de créer un «nouveau SuMo». Nous nous sommes ici intéressés à l'amélioration des qualités du logiciel initial, par des modifications effectuées sur certains éléments, aussi, nous ne présenterons pas le développement et l'aspect informatique du logiciel.

## II- Concept d'origine

Le programme décrit la surface de la protéine grâce à des objets qui codent pour les fonctions chimiques. Il faut voir la surface de la protéine comme un ensemble d'acides aminés obtenu en calculant l'accessibilité au solvant pour chacun d'eux. Un seuil est défini et permet d'identifier les acides aminés de surface (ceux qui possèdent donc une valeur supérieure à la valeur seuil d'accessibilité).

A partir de la liste des acides aminés de surface, le programme, en utilisant les coordonnées tridimensionnelles de chacun des atomes, va placer les objets SuMo en appliquant les règles de ce que l'on appelle un dictionnaire. Développé par M. Jambon, il est basé sur plusieurs concepts détaillés ci-après.

## II-1- Description des liaisons hydrogène

Le dictionnaire est d'abord basé sur la description des liaisons hydrogène avant celle des fonctions chimiques. Cette description s'appuie sur deux types d'objets : les « delta plus » qui décrivent les donneurs de liaisons hydrogène et les « delta minus » qui décrivent les accepteurs de liaisons hydrogène. Ces deux types d'objets servent à définir la polarisation de la liaison chimique lorsqu'elle est de type hydrogène. Ils sont placés au niveau des atomes porteurs d'un proton labile et sur ceux capables de recevoir ce proton. Un second concept est alors appliqué pour décrire les liaisons hydrogène : le « chemical shift » ou CS. Il correspond à un décalage de l'objet par projection en direction de la liaison hydrogène. Le décalage ne concerne que l'objet « delta plus » et correspond à 2,3 Angströms. Ainsi, sur le squelette d'une protéine, l'azote donneur de liaison hydrogène va présenter un objet « delta plus » qui sera décalé de 2,3 Å dans la direction de la liaison hydrogène. Contrairement à ce que l'on pourrait penser, l'objet « delta minus » qui code pour l'atome accepteur de la liaison hydrogène est situé directement sur l'atome, sans projection. Cette différence de traitement semble assez curieuse et engendre probablement une dissonance dans le traitement de l'information chimique. Nous n'avons cependant pas développé un dictionnaire qui corrigerait uniquement cette différence de traitement et qui permettrait d'évaluer exactement cette particularité.

Chacun des acides aminés dispose ainsi de plusieurs objets pour définir les liaisons hydrogène. Leur nombre est variable en fonction des résidus mais on remarque également que cette description n'est pas rigoureuse car, pour certains acides aminés, plusieurs objets manquent, ce qui entraîne une sous-description en terme de liaisons hydrogène. C'est le cas notamment pour l'acide aspartique (ASP ou D) où il manque des objets par rapport à l'acide glutamique (GLU ou E). Le tableau 1 présente les objets « delta plus » et « delta minus » pour chaque acide aminé. La première colonne donne le nom du résidu, les deuxième et troisième colonnes présentent le nombre d'objets « delta plus » et « delta minus » respectivement. La position des objets sur les chaînes latérales ou le squelette est également indiquée.



<b>Nom de l'acide aminé</b>	<b>Nombre d'objet(s) « delta plus » sur squelette ou chaîne latérale</b>	<b>Nombre d'objet(s) « delta minus » sur squelette ou chaîne latérale</b>
Acide Aspartique (ASP)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Acide glutamique (GLU)	1 squelette / 2 chaîne latérale	1 squelette / 2 chaîne latérale
Alanine (ALA)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Arginine (ARG)	1 squelette / 5 chaîne latérale	1 squelette / 0 chaîne latérale
Asparagine (ASN)	1 squelette / 2 chaîne latérale	1 squelette / 1 chaîne latérale
Cystéine (CYS)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Glutamine (GLN)	1 squelette / 0 chaîne latérale	1 squelette / 2 chaîne latérale
Glycine (GLY)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Histidine (HIS)	1 squelette / 2 chaîne latérale	1 squelette / 1 chaîne latérale
Isoleucine (ILE)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Leucine (LEU)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Lysine (LYS)	1 squelette / 1 chaîne latérale	1 squelette / 0 chaîne latérale
Méthionine (MET)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Phénylalanine (PHE)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Proline (PRO)	0 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale
Sérine (SER)	1 squelette / 1 chaîne latérale	1 squelette / 1 chaîne latérale
Thréonine (THR)	1 squelette / 1 chaîne latérale	1 squelette / 1 chaîne latérale
Tryptophane (TRP)	1 squelette / 1 chaîne latérale	1 squelette / 0 chaîne latérale
Tyrosine (TYR)	1 squelette / 1 chaîne latérale	1 squelette / 0 chaîne latérale
Valine (VAL)	1 squelette / 0 chaîne latérale	1 squelette / 0 chaîne latérale

Tableau 1 : Description des liaisons hydrogène par les objets « delta plus » et « minus » pour chaque acide aminé

Comme le montre le tableau ci-dessus, le nombre d'objets décrivant la liaison hydrogène est très variable mais souvent important par résidu. Ainsi, ces deux objets sont les plus représentés, ce qui crée un biais dans la recherche d'objets communs lors de la requête. En effet,

la liaison hydrogène est sur-définie par rapport aux autres fonctions chimiques. Nous analyserons cet aspect dans la section des critiques du dictionnaire.

## II- 2- Description des autres fonctions chimiques et types géométriques :

Les autres fonctions chimiques sont codées par différents objets répartis sur l'acide aminé. On dénombre ainsi 13 différents objets qui sont parfois spécifiques à un résidu comme, par exemple, le « proline point » qui est spécifique à l'acide aminé proline, ou encore l'objet imidazole chiral que l'on ne trouve que sur l'histidine. D'autres objets ne sont pas spécifiques et décriront donc plusieurs acides aminés. Le tableau 2 liste les 13 différents objets, les acides aminés qui disposent de ces objets et leur type géométrique.

Nom de l'objet	Acide aminé disposant de cet objet	Type de l'objet
Acyl plan	ASP ; GLU	4
Amide chiral	ASN ; GLN	5
Guanidinium biplan	ARG	3
Glycine polar	GLY	2
Hydrophobe point	ALA ; ILE ; LEU ; PHE ; PRO ; TRP ; TYR ; VAL	1
Hydroxyl polar	SER ; THR ; TYR	2
Imidazole chiral	HIS	5
Négatif	ASP ; GLU	1
Positif	ARG ; LYS	1
Proline point	PRO	1
Thioéther	MET	1
Thiol polar	CYS	2
Stacker aromatique	HIS ; PHE ; TRP ; TYR	3

Tableau 2 : Objets SuMo d'origine : types géométriques et acides aminés porteurs

Les objets peuvent être de cinq types différents que nous ne détaillerons pas ici. Ces types géométriques s'appuient sur un aspect vectoriel qui augmente les critères nécessaires pour la superposition des objets SuMo. A titre d'exemple, les objets « stacker aromatique » sont codés avec un vecteur perpendiculaire au plan du cycle aromatique, donc, pour qu'un autre objet « stacker aromatique » puisse se superposer, il faut nécessairement que les deux plans des cycles aromatiques soient parallèles. Ces contraintes géométriques se retrouvent pour la plupart des objets et nous discuterons ce point qui nous a semblé limitant dans la section des critiques.

### II-3- Poids des objets :

Chaque objet sera également défini par un poids, qui sera inclus dans la fonction de score dédié à classer les protéines les plus proches de celle de référence. Les poids ont une importance particulièrement critique. Ils ont été modifiés par la société Medit SA à de nombreuses reprises dans le but d'optimiser les résultats de requête en fonction de la question. Nous avons pensé qu'une telle méthode influençait les résultats ; aussi, la section concernant le nouveau dictionnaire d'objets présentera des poids strictement définis pour chaque objet et ne variant pas en fonction de l'étude envisagée.

Le tableau ci-dessous présente les objets par acides aminés, leurs poids et leurs positions.

Nom de l'acide aminé	Nom de l'objet	Poids de l'objet
Acide aspartique	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Acyl plan (C gamma)	0,75
	Négatif (entre O delta 1 et 2)	0,75
Acide glutamique	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Delta moins (O epsilon 1)	0,6
	Delta moins (O epsilon 2)	0,6
	Acyl plan (C delta)	0,75
	Négatif (entre O epsilon 1 et 2)	0,75

Alanine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (C bêta)	0,75
Arginine	Delta plus (N backbone)	0,6
	Delta plus (N epsilon)	0,6
	Delta plus (N h1)	0,6
	Delta plus (N h1)	0,6
	Delta plus (N h2)	0,6
	Delta plus (N h2)	0,6
	Delta moins (O backbone)	0,6
	Positif (entre Nh1 et Nh2)	0,75
Guanidinium biplan (C zêta)	0,85	
Asparagine	Delta plus (N backbone)	0,6
	Delta plus (N delta)	0,6
	Delta plus (N delta)	0,6
	Delta moins (O backbone)	0,6
	Delta moins (O delta 1)	0,6
	Amide chiral (C gamma)	0,75
Cystéine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Thiol polar (S gamma)	0,65
Glutamine	Delta plus (N backbone)	0,6
	Delta plus (N epsilon 2)	0,6
	Delta plus (N epsilon 2)	0,6
	Delta moins (O backbone)	0,6
	Delta moins (O epsilon 1)	0,6
	Amide chiral (C delta)	0,75
Glycine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Glycine Polar (N backbone)	0,75

Histidine	Delta plus (N backbone)	0,6
	Delta plus (N delta 1)	0,6
	Delta plus (N epsilon 2)	0,6
	Delta moins (O backbone)	0,6
	Positif (N epsilon 2)	0,75
	Imidazole chiral (centre du cycle)	0,9
	Stacker aromatique (centre du cycle)	0,9
Isoleucine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (entre C bêta/ C gamma 1)	0,75
	Hydrophobe point (entre C bêta/ C gamma 2)	0,75
Leucine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (C bêta)	0,75
	Hydrophobe point (C gamma)	0,75
Lysine	Delta plus (N backbone)	0,6
	Delta plus (N zeta)	0,6
	Delta moins (O backbone)	0,6
	Positif (N zeta)	0,75
Méthionine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Thioether (S delta)	0,65
Phénylalanine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (entre C bêta et C gamma)	0,75
	Hydrophobe point (centre du cycle)	0,75
	Stacker aromatique (centre du cycle)	0,9
Proline	Delta moins (O backbone)	0,6
	Hydrophobe point (dans le cycle)	0,75
	Proline point (entre C alpha, C delta et N)	0,75

Sérine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Hydroxyl polar (O gamma)	0,65
Thréonine	Delta plus (N backbone)	0,6
	Delta plus (O gamma 1)	0,6
	Delta moins (O backbone)	0,6
	Delta moins (O gamma 1)	0,6
	Hydroxyl polar (O gamma 1)	0,65
Tryptophane	Delta plus (N backbone)	0,6
	Delta plus (N epsilon 1)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (C gamma)	0,75
	Hydrophobe point (centre du cycle à 6 carbones)	0,75
	Stacker aromatique (centre du cycle à 5 carbones)	0,9
	Stacker aromatique (centre du cycle à 6 carbones)	0,9
Tyrosine	Delta plus (N backbone)	0,6
	Delta plus (O h)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (entre C bêta /C gamma)	0,75
	Hydrophobe point (centre du cycle à 6 carbones)	0,75
	Hydroxyl polar (O h)	0,65
	Stacker aromatique (centre du cycle à 6 carbones)	0,9
Valine	Delta plus (N backbone)	0,6
	Delta moins (O backbone)	0,6
	Hydrophobe point (C bêta)	0,75

Tableau 3 : Description des acides aminés par les anciens objets : position et poids des objets

## II-4- Critiques du dictionnaire d'origine :

Les grandes qualités descriptives et comparatives du logiciel d'origine ne sont en aucun cas remises en question. Cependant, il nous a semblé tout à fait intéressant et nécessaire d'apporter plus de cohérence concernant le traitement du positionnement des objets sur les atomes, la description trop importante des liaisons hydrogène, les contraintes liées aux vecteurs directionnels assignés à de nombreux objets, la spécificité d'un objet par acide aminé, et enfin la variation des poids en fonction de l'étude menée.

Les solutions que nous proposons pour répondre aux différents problèmes soulevés sont donc les suivants, et notre nouveau dictionnaire d'objets a été créé afin de minimiser les aspects qui nous ont semblés négatifs au fonctionnement optimal de SuMo :

1- Suppression des objets « delta plus » et « minus » : comme nous le montrerons à l'aide des deux exemples dans le paragraphe 4 de ce chapitre, ces objets, étant placés sur ou à proximité de tous les hétéroatomes capables de créer une liaison hydrogène, sont extrêmement nombreux et génèrent un bruit de fond qui altère les facultés du logiciel à classer les meilleurs résultats en tête de liste, ou même à retrouver deux sites très proches.

2- Suppression du « chemical shift » : d'une part, il ne nous a pas semblé nécessaire de décaler certains objets par rapport à l'atome porteur d'une fonction à décrire, et d'autre part, le CS n'était pas appliqué de façon uniforme et cohérente.

3- Suppression des vecteurs : comme nous l'avons déjà dit, les contraintes de superposition liées à l'aspect directionnel des objets ont été jugées inutiles, et les tests comparatifs entre l'ancien et le nouveau dictionnaire (paragraphe 4) le montrent.

4- Suppression des objets spécifiques à un acide aminé. Nous avons remarqué que, non seulement, il existait des objets créés pour un seul acide aminé, (c'est le cas pour les objets Proline et Glycine), mais d'autres objets qui ne portent pas le nom d'un acide aminé sont toutefois retrouvés uniquement sur un seul type de résidu. Ainsi, l'objet thiol ne se trouve que sur la cystéine, alors que l'on aurait pu penser qu'il s'adapterait très bien à la méthionine qui, elle,

possède un objet thioéther, qui lui est propre. De même, l'objet imidazole ne décrit que l'histidine et l'objet guanidinium que l'arginine.

Nous sommes partis du principe qu'il fallait décrire tous les acides aminés de façon plus harmonieuse et logique en évitant d'attribuer un objet à un unique résidu.

5- Réajustement du nombre d'objets pour certains acides aminés trop faiblement décrits, ou au contraire trop abondamment définis. A titre d'exemple, la lysine ne possède aucun objet sur toute la longueur de sa chaîne latérale (ce qui sous-estime l'hydrophobicité et la topologie), tandis que l'arginine possède sept objets juste pour décrire l'extrémité de sa chaîne latérale.

En créant de nouveaux objets, nous avons souhaité une description plus homogène des acides aminés les uns par rapport aux autres.

6- Impossibilité de faire varier les poids en fonction de l'étude afin de ne pas influencer les résultats. Nous verrons dans le paragraphe suivant que les poids que nous accordons aux objets sont les mêmes pour tous, sauf pour les objets « hydrophobes » et « carbone alpha ». En effet, il est nécessaire de ne pas sur-décrire l'hydrophobicité des protéines. D'autre part, concernant les objets « carbone alpha », puisqu'ils sont communs à tous les acides aminés et que nous ne voulions pas apporter un bruit de fond, ni orienter la superposition sur le squelette de la protéine, nous avons diminuer leur poids.

## **III- Nouvelle version de SuMo**

### **III-1- Concept: un pharmacophore appliqué aux protéines**

Notre nouvelle approche s'appuie sur un fonctionnement beaucoup plus simple et offre une grande souplesse en terme d'applications. Nous sommes partis du concept du pharmacophore, qui rappelons-le, correspond à l'ensemble des fonctions chimiques réparties dans l'espace et essentielles à l'activité d'une molécule, et nous l'avons appliqué aux protéines. Ainsi, les objets SuMo servent à définir le « pharmacophore » de la protéine, et nous avons dans ce sens procédé à plusieurs grands changements :



Tout d'abord, nous avons décidé d'utiliser des objets codant pour une fonction chimique donnée. Ainsi, tous les acides aminés ayant une fonction chimique commune disposeront du même objet. Ceci permet d'éviter d'avoir des objets spécifiques à un acide aminé unique et de ce fait, tous les résidus sont traités de manière équitable.

L'autre originalité consiste à ne pas décrire les liaisons hydrogène. En effet, nous avons constaté que leur nombre très important influence le score global. Cumulées, elles représentent un poids bien trop fort, ce qui risque de « noyer » le système en générant un bruit de fond inutile. De plus, la description des liaisons hydrogène est implicitement donnée par le type de fonction chimique et les objets « delta plus » ou « minus » entraînent alors une redondance d'information.

Ensuite, nous avons supprimé les vecteurs assignés aux objets. Ainsi, notre approche est tout à fait différente et se veut plus souple dans la description des surfaces des protéines. Nous avons supprimé l'aspect « directionnel » des objets car celui-ci correspond à une vision beaucoup trop statique et figée de la protéine. En effet, lorsque nous observons une simulation de dynamique moléculaire, même après stabilisation de la protéine, nous constatons des variations importantes au niveau des chaînes des acides aminés de surface, ce qui donne à la protéine un aspect « dynamique ». Cette flexibilité n'est pas prise en compte dans la structure unique d'une protéine définie expérimentalement et qui nous sert de modèle pour placer les objets. Les structures déposées à la PDB sont d'ailleurs, soit issues de minimisations sous contraintes, soit de structures moyennes. Pour remédier à ce problème, nous avons choisi de n'utiliser qu'un type géométrique d'objet : le type « point » qui ne possède aucune fonction géométrique directionnelle. Ainsi, nous estimons que si spatialement deux fonctions chimiques identiques sont extrêmement proches, c'est-à-dire placées à une distance inférieure à 1 Å, alors les objets décrivant ces deux fonctions sont déclarés communs.

D'autre part, nous avons réduit à dix (contre treize) le nombre d'objets SuMo. Nous souhaitons apporter plus de cohérence dans la description des fonctions chimiques et nous les présentons en détail dans le paragraphe suivant.

Enfin, nous avons supprimé le « chemical shift » puisque d'une part, il n'y avait pas d'uniformité dans l'application de ce « décalage des objets », et que d'autre part, le fait de placer les objets directement sur les fonctions chimiques comme nous le proposons ici, permet de

respecter parfaitement la topologie de la surface décrite. Ceci va, une fois de plus, dans le sens d'une plus grande flexibilité et cohérence dans le traitement de l'information.

## III-2- Nouveau dictionnaire :

### III-2-1- Flexibilité des protéines :

La perte de l'aspect directionnel (suppression des vecteurs) permet d'augmenter la flexibilité du système. Cette modification est importante, par rapport à l'approche d'origine. En effet, auparavant, pour que deux objets identiques se superposent, il leur fallait une direction vectorielle commune. Or, cette contrainte nous paraît négative dans la pertinence de la description. L'analyse des chaînes latérales des acides aminés lors d'une simulation de dynamique moléculaire, montre une flexibilité très importante. Elles sont d'autant plus flexibles que l'acide aminé est placé en surface puisque ainsi la chaîne latérale n'est pas contrainte par les autres acides aminés environnants. L'analyse que nous avons réalisée, pour des résidus de surface, des angles de torsion et de la distance entre le centre de la fonction chimique de la chaîne latérale et le carbone alpha, montre effectivement de fortes variations.

Ces conclusions sont issues d'une simulation de dynamique moléculaire sur la protéine kinase C isoforme alpha effectuée sur une durée de 10 ns en phase aqueuse. Le champ de forces Amber FF03 a été utilisé. La température (300 Kelvin) et la pression (1 Atmosphère) sont constantes. La protéine a été préalablement insérée dans un ensemble périodique cubique rempli avec des molécules d'eau de type TIP3P. Toutes les 20 ps, une conformation de la protéine est échantillonnée. L'accessibilité de l'acide aminé a été vérifiée par un calcul de DSSP. Pour chaque acide aminé en surface de chacun des conformères ainsi obtenus, nous avons pris deux mesures : 1- La distance entre le carbone alpha et le centre de la fonction chimique terminant la chaîne latérale, 2- La variation de l'angle de torsion entre le centre de la fonction et le carbone alpha. Ainsi, nous évaluons parfaitement les changements de position des chaînes latérales. Pour chaque type d'acide aminé, plusieurs résidus en différents endroits de la surface ont été sélectionnés de telle façon à pouvoir moyenniser les distances. Le tableau 4 présente l'ensemble des mesures obtenues, avec dans la colonne 1, le nom de l'acide aminé (aa), puis le nombre d'acides aminés analysés, en colonne 3 la distance moyenne entre le centre de la fonction chimique et le carbone

alpha du résidu, enfin la dernière colonne donne l'angle de torsion moyenné entre le centre de la fonction chimique et le carbone alpha de l'acide aminé. Pour la glycine et la proline, bien entendu, nous ne pouvons donner ces mesures (Non Calculé dans le tableau).

Nom de l'aa	Nombre d'aa	Distance moyenne entre le centre de fonction chimique et le C $\alpha$ de l'aa	Variation de l'angle de torsion moyen entre le centre de la fonction chimique et le C $\alpha$ de l'aa
Alanine	3	1,53 $\pm$ 0,0571	$\pm$ 6,698
Arginine	3	6,27 $\pm$ 0,00583	$\pm$ 0,935
Asparagine	4	3,48 $\pm$ 0,04157	$\pm$ 7,741
Aspartate	2	2,58 $\pm$ 0,00493	$\pm$ 15,625
Cystéine	5	2,83 $\pm$ 0,08188	$\pm$ 15,838
Glutamique	3	4,15 $\pm$ 0,06925	$\pm$ 15,380
Glutamine	3	3,88 $\pm$ 0,04144	$\pm$ 17,499
Glycine	NC	NC	NC
Histidine	3	4,59 $\pm$ 0,08852	$\pm$ 5,339
Isoleucine	4	3,88 $\pm$ 0,06349	$\pm$ 13,404
Leucine	3	3,93 $\pm$ 0,03052	$\pm$ 5,961
Lysine	5	6,31 $\pm$ 0,01722	$\pm$ 11,402
Méthionine	3	3,77 $\pm$ 0,01108	$\pm$ 1,789
Phénylalanine	4	5,18 $\pm$ 0,07468	$\pm$ 2,586
Proline	NC	NC	NC
Sérine	4	2,46 $\pm$ 0,01183	$\pm$ 16,306
Thréonine	3	2,33 $\pm$ 0,03210	$\pm$ 18,112
Tryptophane	2	6,42 $\pm$ 0,08223	$\pm$ 13,553
Tyrosine	3	6,51 $\pm$ 0,09855	$\pm$ 12,419
Valine	3	2,53 $\pm$ 0,07605	$\pm$ 14,977

Tableau 4 : Analyse de la flexibilité de la protéine kinase C par dynamique moléculaire

Ces résultats confirment la mobilité des chaînes latérales ce qui justifie notre choix de supprimer l'aspect directionnel des objets SuMo. Effectivement, l'orientation des chaînes

latérales va varier au cours du temps et elles « vibrent » en décrivant un large espace conformationnel qu'il nous a semblé essentiel de prendre en considération. La direction des fonctions chimiques dans une conformation unique, comme nous l'a trouvons dans la PDB, n'est finalement pas très représentative de l'ensemble des mouvements observés normalement, et il paraît préjudiciable de négliger cette information. Nous pensons que si deux objets identiques sont présents dans la même zone géographique, ils doivent être reconnus comme étant communs même s'ils n'ont pas la même orientation. Rappelons que dans la version précédente, deux objets qui ne convergeaient pas en terme directionnel n'étaient pas considérés comme communs. La souplesse introduite par la suppression des vecteurs permet de tenir compte des mouvements observés lors des simulations de dynamique moléculaire sans avoir à rajouter un algorithme simulant ces variations. Enfin, cette approche permet de simplifier la superposition des objets.

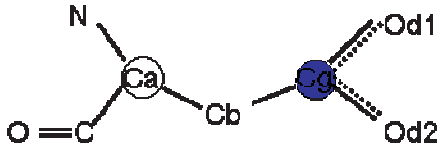
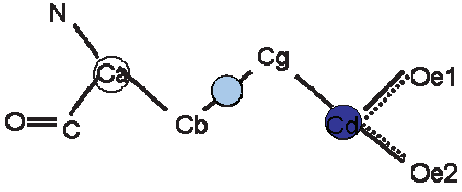
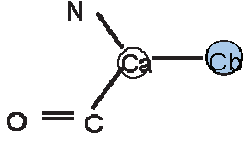
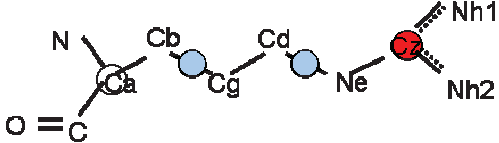
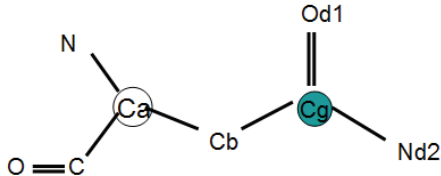
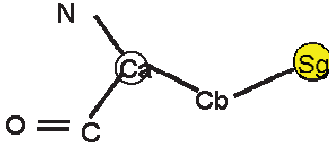
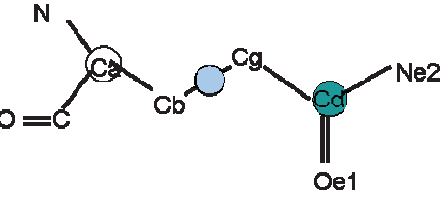
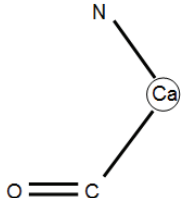
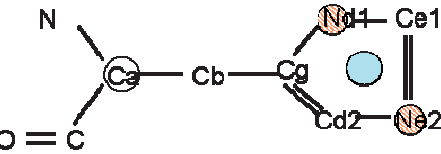
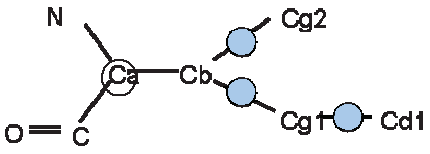
### III-2-2- Les nouveaux objets :

Nous avons appliqué l'idée du pharmacophore aux chaînes latérales des acides aminés, en créant un objet pour chaque type de fonction chimique caractéristique. Nous obtenons les objets suivants :

- Acyl
- Amide
- Amine aliphatique
- Amine aromatique
- Guanidinium
- Hydrophobe aliphatique
- Hydrophobe aromatique
- Hydroxyl
- Thiol

Ces neuf objets codent pour différentes fonctions chimiques portées par les chaînes latérales de surface. Il faut remarquer que certaines fonctions chimiques sont codées par plusieurs objets simultanément. C'est notamment le cas pour les cycles aromatiques, qui disposent de l'objet « hydrophobe aromatique » pour décrire le cycle hyperconjugué et de l'objet « amine

aromatique» si des azotes sont présents, comme pour l'histidine. Afin de prendre à compte l'hydrophobicité des chaînes latérales, nous avons créé l'objet « hydrophobe aliphatique » qui sera placé entre deux atomes de carbone et qui représente le groupement CH<sub>2</sub>-CH<sub>2</sub>. Les schémas ci-après représentent les objets. La légende est présentée à la suite du tableau.

	
Acide aspartique (Asp)	Acide glutamique (Glu)
	
Alanine (Ala)	Arginine (Arg)
	
Asparagine (Asn)	Cystéine (Cys)
	
Glutamine (Gln)	Glycine (Gly)
	

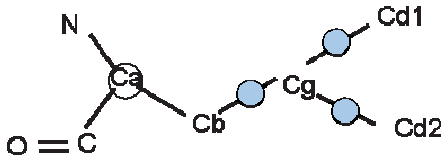
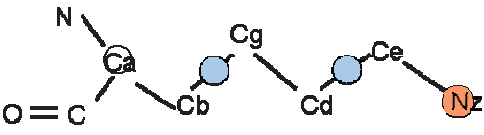
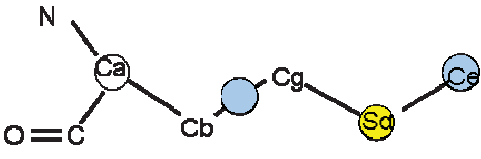
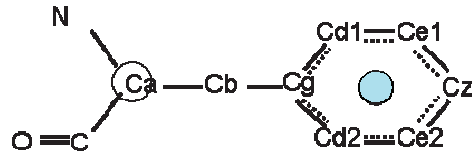
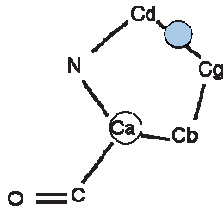
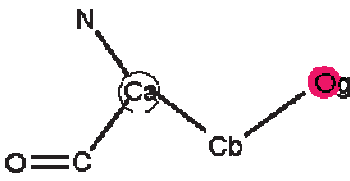
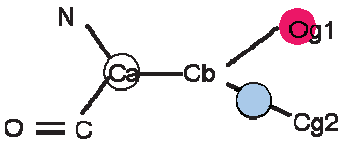
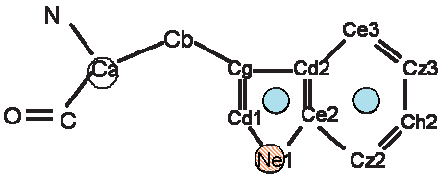
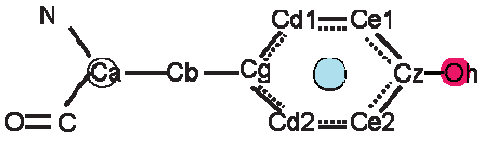
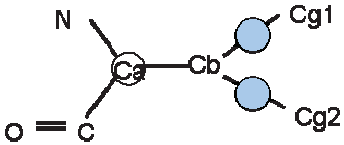










Histidine (His)	Isoleucine (Ile)
	
Leucine (Leu)	Lysine (Lys)
	
Méthionine (Met)	Phénylalanine (Phe)
	
Proline (Pro)	Sérine (Ser)
	
Thréonine (Thr)	Tryptophane (Trp)
	
Tyrosine (Tyr)	Valine (Val)

Tableau 5 : Description des acides aminés par les nouveaux objets

Ci-dessous, les codes couleurs pour les dix objets :

- Acyl 
- Amide 
- Amine aliphatique 
- Amine aromatique 
- Carbone alpha 
- Guanidinium 
- Hydrophobe aliphatique 
- Hydrophobe aromatique 
- Hydroxyle 
- Thiol 

Notre dixième objet est centré sur le carbone alpha (on l'appelle donc: « objet carbone alpha », représenté par un rond gris clair) de chaque acide aminé et il a pour but de rendre compte du repliement de la protéine. Nous savons que la structure d'une protéine est plus conservée que la fonction qu'elle exerce, d'où l'idée de créer cet objet qui augmentera le poids accordé au repliement et donc majorera le score des protéines possédant un repliement similaire.

Toutefois, s'agissant des travaux sur les phénomènes d'hypersensibilité allergique, l'objet « carbone alpha » a dû être supprimé. En effet, dans ce cas, nous avons constaté que la reconnaissance par les structures de l'immunité ne se faisait que sur un très faible nombre d'objets SuMo. De ce fait, si nous avions conservé l'objet carbone alpha, toutes les structures possédant un repliement identique à la référence auraient vu leurs objets « carbone alpha » se superposer. Cela aurait alors augmenté le score final de ces structures qui seraient passées en début de liste même si elles ne possédaient aucun objet impliqué dans la reconnaissance par les IgE ou les TCR.

### III-2-3- Poids des objets :

Nous avons choisi d'assigner le même poids à nos objets pour l'ensemble des 23 acides aminés. Seuls les objets « hydrophobe aliphatique » et « carbone alpha » possèdent un poids inférieur car ils codent pour des fonctions chimiques très nombreuses et nous ne voulions pas créer de redondance d'information concernant l'hydrophobicité ou le repliement. Les poids respectifs des dix objets sont présentés dans le tableau 6 suivant :

<b>Nom de l'objet</b>	<b>Poids de l'objet</b>
Acyl	1,00
Amide	1,00
Amine aliphatique	1,00
Amine aromatique	1,00
Thiol	1,00
Hydroxyle	1,00
Guanidinium	1,00
Hydrophobe aliphatique	0,60
Hydrophobe aromatique	1,00
Carbone alpha	0,50

Tableau 6 : Poids des nouveaux objets

Le choix que nous avons fait, d'accorder à tous les objets le même poids, nous a paru essentiel afin de décrire le plus fidèlement la réalité chimique, de limiter l'influence liée aux poids individuels sur le score global et donc de respecter notre démarche de cohérence, marquée



également par la suppression du « chemical shift », des vecteurs et des objets spécifiques à un acide aminé donné.

### III-3- Conclusions et remarques :

Il ne nous a pas semblé logique d'accorder plus d'importance à une fonction amine qu'une amide ou un cycle aromatique. Toutefois, deux objets présentent un poids inférieur. L'objet hydrophobe aliphatique a un poids de 0,60 car c'est un objet qui se retrouve plusieurs fois au sein d'un acide aminé hydrophobe, et nous avons constaté que si nous laissons le poids à la valeur 1, nous faussions le résultat et surestimions l'importance du caractère hydrophobe. Nous avons fixé la valeur à 0,60 en nous basant sur le constat que les acides aminés hydrophobes possèdent en moyenne 2 à 3 objets de ce type. D'autre part, nous avons remarqué, à travers les études que nous avons menées afin de valider nos objets, que l'objet « carbone alpha » devait nécessairement avoir un poids inférieur. En effet, si l'on conserve la valeur 1 pour décrire le carbone alpha, la superposition est trop orientée sur le squelette de la protéine et l'importance des chaînes latérales est minimisée. Le paragraphe qui suit présente les deux études-tests qui nous ont servi à mettre à l'épreuve nos nouveaux objets et paramètres.

Cependant, comme nous l'avons déjà mentionné, il nous a été nécessaire d'éliminer cet objet « carbone alpha » dans les situations où la description d'une cible ne se fait que par quelques objets. Par exemple, si la requête correspond à un site de très faible taille, soit environ trois ou quatre objets (autres que le « carbone alpha »), alors le fait de conserver l'objet carbone alpha dévie la superposition et annule la prise en considération des chaînes latérales.

Il faut noter que la modification du dictionnaire d'objets a permis de réduire fortement le temps de calcul des « jobs SuMo ». En effet, d'une part, notre dictionnaire présente moins d'objets (dix contre treize), d'autre part, la superposition est plus simple à effectuer puisque nous avons supprimé les vecteurs et le « chemical shift ».

Enfin, la suppression des objets « delta plus » et « minus » a largement allégé le système en terme de temps de calcul. Nous n'avons pas effectué de benchmark entre les deux dictionnaires, mais nous évaluons le gain de temps de l'ordre d'un facteur deux.

## IV- Tests comparatifs entre SuMo d'origine et SuMo actuel :

Nous avons réalisé plusieurs tests visant d'une part à évaluer la pertinence du nouveau dictionnaire développé, et d'autre part à comparer les capacités de notre version SuMo par rapport au logiciel d'origine.

### IV-1-Test des lectines :

Le premier test a été effectué sur le même ensemble que le Dr. M. Jambon (Jambon *et al.*, 2003) afin de comparer les deux dictionnaires. La base de données se compose de 90 lectines vraies et de 18 lectines fausses, soit 108 structures. Les lectines sont des glycoprotéines qui se lient spécifiquement et de façon réversible à un oligosaccharide au niveau d'un site connu (présent sur des carbohydrates, des glycoprotéines ou des glycolipides), de façon comparable aux enzymes et aux immunoglobulines. Elles interviennent dans divers processus biologiques, lors de la reconnaissance entre les cellules, et notamment lors de l'activation de l'immunité innée. Les 18 « fausses lectines » sont identifiées dans la PDB par le terme « lectin like ». Elles possèdent bien un site de fixation pour les oligosaccharides, qui présente par ailleurs une forte similarité de repliement avec le site des vraies lectines, mais elles sont incapables de fixer les sucres.

Une comparaison de séquences de type BLAST (Basic Local Alignment Search Tool) ne permet pas de discerner les vraies lectines des fausses alors que l'approche SuMo le peut. Le jeu de test a été développé par M. Jambon lors de la conception de SuMo. Par la suite, le logiciel ainsi que son client (MED-SuMo) a été repris par la société Medit SA. Aussi, pour plus de pertinence et une bonne comparaison des capacités du logiciel, nous avons testé les trois versions :

- Serveur Web : Serveur développé au laboratoire avec le dictionnaire d'objets d'origine. La requête est directement soumise au serveur sans passer par le client MED-SuMo de Medit.

- Anciens objets : Serveur développé au laboratoire avec les objets d'origine retravaillés par Medit SA. La requête est soumise au serveur et passe par le client MED-SuMo.
- Nouveaux objets : Serveur développé au laboratoire avec les nouveaux objets. La requête est soumise au serveur en passant par le client MED-SuMo.

La requête est identique dans les trois situations. La structure de référence, qui est comparée au jeu total de vraies et fausses lectines, correspond à la lectine d'*Arachis hypogaea* : la cacahuète. C'est une vraie lectine définie en 1996 par Banerjee *et al.* au moyen de la cristallographie de rayons X avec une résolution de 2,25 Å et déposée à la PDB sous le code 2PEL. Cette protéine est capable de fixer un ligand : le bêta-lactose qui est co-cristallisé avec la lectine au niveau d'un site de fixation nommé LAT qui nous sert de site de référence. Ainsi, nous avons sélectionné les objets SuMo entourant le bêta-lactose, en respectant une distance maximale de 6 Å (ceci correspond à la « query ») et nous avons demandé aux trois versions de comparer la requête aux 108 lectines de la base. Les résultats sont compilés dans le tableau 7 et ne présentent que les 18 fausses lectines, puisque, rappelons-le, l'objectif de cette étude est de discerner les vraies lectines, des fausses. Autrement dit, nous attendons du logiciel qu'il ne retrouve que des vraies lectines puisque la requête porte sur une vraie lectine. Notons que NP signifie Non Présente, ce qui veut dire que la protéine n'a pas été retrouvée par le logiciel. Les cases grises mettent en évidence les fausses lectines. Le score SuMo permet de classer les protéines selon plusieurs paramètres (Jambon et al. 2003) dont le nombre d'objets SuMo communs avec les objets qui définissent la « query ». Ainsi, plus une protéine sera proche de la référence, plus sa place dans le classement sera haute.

	Nouveaux objets		Anciens objets		Serveur Web	
Code PDB	Score SuMo	Nombre d'objets communs	Score SuMo	Nombre d'objets communs	Score SuMo	Nombre d'objets communs
1IOA	NP	NP	NP	NP	NP	NP
1HSS	NP	NP	3,49	7	NP	NP
3AIT	NP	NP	2,45	4	2,10	4
1AVB	NP	NP	3,00	6	NP	NP
4AIT	NP	NP	3,65	6	NP	NP
1CLV	NP	NP	2,45	4	NP	NP
1HOE	NP	NP	2,45	4	NP	NP
1HTX	NP	NP	2,15	3	NP	NP
1CN1	NP	NP	2,44	5	2,10	4
1APN	NP	NP	2,44	5	2,10	3
1DQ2	NP	NP	NP	NP	2,10	4
1ENQ	NP	NP	NP	NP	2,10	3
1OK0	NP	NP	2,45	4	NP	NP
1QFD	NP	NP	NP	NP	NP	NP
2AIT	NP	NP	2,45	4	NP	NP
1CES	NP	NP	2,44	5	NP	NP
2UU8	4,44	7	2,29	5	NP	NP
1ENS	NP	NP	2,40	4	NP	NP

Tableau 7 : Résultats comparatifs des trois tests SuMo sur le jeu des lectines

Nous attendons de SuMo qu'il soit capable de différencier les sites de fixation de lectines vraies de ceux des fausses lectines. La structure de référence correspondant à celle d'une vraie lectine, le résultat de la requête devrait éliminer de la liste des résultats toutes les fausses lectines, pour ne garder que les vraies et les classer en fonction de la similarité de leur site de fixation avec le site LAT de la lectine 2PEL. Comme le montre le tableau 7, il existe une très grande différence entre les résultats obtenus avec les nouveaux objets et ceux avec les anciens objets. En effet, seule une fausse lectine est retrouvée dans le classement qui utilise les nouveaux objets (ce qui fait une réponse erronée sur le jeu de 18 fausses lectines, soit environ 5% d'erreur), alors qu'avec

les anciens objets 12 fausses lectines sont retrouvées et classées, ce qui correspond à douze réponses inexacts, soit environ 66% d'erreur. Concernant l'entrée 2UU8, avec les anciens objets, son score SuMo étant inférieur à 30% du score maximal, nous considérons le résultat comme non significatif, et donc nous ne tenons pas compte de cette entrée incorrecte. S'agissant des résultats issus du serveur web, qui rappelons-le utilise la version originale et historique de SuMo, on voit que cinq fausses lectines sont retrouvées, soit un pourcentage d'erreur d'environ 28%.

Ainsi, ces résultats montrent très clairement qu'avec les nouveaux objets, le logiciel SuMo est nettement plus performant puisqu'une seule fausse lectine (correspondant à l'entrée 2UU8) est classée parmi les lectines vraies. Remarquons que cette fausse lectine présente 7 objets communs avec le site de référence, ce qui justifie son score relativement élevé (4,44), signifiant ainsi que cette fausse lectine possède un site de fixation aux sucres très proche de celui de référence. Notons qu'aucune autre fausse lectine n'apparaît dans le classement. Ce résultat montre indéniablement que la modification du dictionnaire d'objets, marquée par la suppression du « chemical shift », des vecteurs directionnels, ainsi que la réduction du nombre d'objets (particulièrement l'élimination des « delta plus » et « minus ») permet de réduire le bruit de fond et d'augmenter la sélectivité et les capacités de discernement du logiciel.

## IV-2-Test des protéases :

Toujours avec pour objectif de mettre à l'épreuve les nouveaux paramètres de notre logiciel, nous avons imaginé un test plus poussé. Notre ambition était de distinguer les enzymes de type protéase par rapport à d'autres enzymes de type isomérase. Ce choix n'est pas fait par hasard puisqu'il nous fallait, pour ce test, deux séries de protéines qui ne présentaient pas les mêmes fonctions biologiques afin de voir si SuMo pouvait différencier les unes des autres. Les protéases, encore appelées peptidases, sont des enzymes capables de couper la liaison peptique, et ce phénomène nécessitant la présence d'une molécule d'eau, les protéases sont aussi des hydrolases. Quant aux isomérases, ce sont des enzymes qui catalysent les changements au sein des molécules, et qui n'ont aucune fonction d'hydrolase. Ainsi, nous avons créé un jeu de 1031 protéines extraites de la PDB et qui se compose de la façon suivante:

- 970 isomérasés, qui ne présentent pas une fonction secondaire de type protéase
- 61 protéases, parmi lesquelles 16 sérines protéases. Le reste étant divisé en différents sous-groupes tel que celui des métalloendopeptidases, des cystéine endopeptidases...

La référence correspond à une sérine protéase déposée à la PDB sous le code 1AFQ. Il s'agit de la gamma chymotrypsine, découverte par diffraction de rayons X avec une résolution de 1,8 Å (Kashima, 1998). La requête correspond aux objets placés au niveau du site actif. La protéine 1AFQ a été cristallisée avec un substrat synthétique, qui aurait pu constituer notre « query », mais nous avons préféré définir manuellement les objets à comparer afin de ne sélectionner que le site actif et son environnement immédiat. Ainsi, notre query correspond à trois sphères de sélection d'un rayon de 10 Å chacune, centrées sur les acides aminés suivants : W215, F302 et S218.

La question est double :

1- SuMo est-il capable de différencier les protéases des isomérasés? Autrement dit, et sachant que la référence est une protéase, va-t-on obtenir les protéases en tête de classement?

2- SuMo est-il capable de distinguer les sérines protéases des autres protéases, autrement dit, sachant que la requête porte sur une sérine protéase, va-t-on obtenir les sérines protéases dans les résultats à plus forts scores ?

Tout comme pour le test sur les lectines, nous avons effectué la même requête sur le serveur web, sur le cluster qui utilise la dernière version du serveur de la société Medit (anciens objets), et enfin avec notre serveur qui utilise les nouveaux objets. Le résumé des résultats se trouve dans le tableau 8 qui ne présente que les premières entrées. Il faut noter que pour la version web, sur les 1031 modèles, 77 entrées dont 5 protéases (codes PDB: 2H4W, 1PSN, 2VA6, 2EA2, 3BI1) ont dû être retirées. En effet, le serveur n'a pas pu les traiter car elles entraînent un « crash ». Ceci a pu être corrigé dans la version tournant sur le cluster, ainsi, les résultats pour les anciens et nouveaux objets sont exploitables, mais ceux provenant de la version web sont incomplets et ne sont présentés dans le tableau qu'à titre indicatif. Dans la colonne « Fonction », nous avons indiqué le type d'enzyme avec pour les protéases: SP pour Sérine

Protéase, CEP pour Cystéine EndoPeptidase, MCP pour MétalloCarboxyPeptidase, MEP pour MétalloEndoPeptidase, TEP pour Thréonine EndoPeptidase, CCP pour Cystéine type CarboxyPeptidase, AEP pour AminoEndoPeptidase et ISO pour toutes les isomérases.

Nouveaux objets				Anciens objets				Serveur Web			
Code PDB	Score SuMo	Objets communs	Fonction	Code PDB	sumo score	Objets communs	Fonction	Code PDB	volume	Objets communs	
1H4W	52,97	131	SP	1H4W	16,76	39	SP	H4W	5.1	7	SP
1A5H	47,16	109	SP	1FAX	10,98	25	SP	1D3P	4.5	6	SP
1GI7	46,68	106	SP	1A5H	8,87	22	SP	1BML	3.6	5	SP
1FAX	45,51	106	SP	1GI7	4,9	10	SP	1A5H	3.6	5	SP
1FLE	45,23	101	SP	1Q6H	4,34	8	ISO	1K41	3	4	ISO
1EZX	40,13	90	SP	1XYH	4,12	9	ISO	1BCR	2.7	4	SP
1LTO	39,51	93	SP	3REQ	4,08	7	ISO	1SO0	2.7	4	ISO
1BML	38,76	92	SP	1IAU	4,07	8	SP	2DKA	2.7	4	ISO
1D3P	37,65	87	SP	1EZX	4	9	SP	1SNZ	2.7	4	ISO
1FQ3	36,37	95	SP	1PJH	3,99	8	ISO	1TCO	2.7	4	ISO
1IAU	33,97	75	SP	2GZM	3,95	7	ISO	1FQ3	2.7	4	SP
1DDJ	28,89	63	SP	1B6C	3,94	7	ISO	1HZJ	2.1	3	ISO
1NSS	11,45	33	ISO	1N23	3,9	7	ISO	2D3T	2.1	3	ISO
1JC4	11,05	28	ISO	1GXD	3,9	8	MEP	2UDP	2.1	3	ISO
1B6C	10,66	34	ISO	1P5Q	3,9	8	ISO	1WDK	2.1	3	ISO
1UPI	10,59	30	ISO	2IAM	3,8	6	ISO	1TIM	2.1	3	ISO
1VGA	10,33	34	ISO	2OJU	3,78	8	ISO	2GSA	2.1	3	ISO
2HXG	10,31	23	ISO	1D3P	3,75	9	SP	1UMP	2.1	3	ISO
1Z8K	9,99	22	ISO	1EQ2	3,75	8	ISO	1KVS	2.1	3	ISO
1ZVC	9,97	23	ISO	1Q6I	3,75	8	ISO	1U79	2.1	3	ISO
1R2T	9,89	20	ISO	2CIR	3,69	7	ISO	1I3M	2.1	3	ISO
2F6Q	9,8	31	ISO	2CIS	3,69	7	ISO	1KVU	2.1	3	ISO
1RCQ	9,78	35	ISO	1TTJ	3,69	7	ISO	3GSB	2.1	3	ISO
1LZO	9,75	23	ISO	2BI8	3,67	7	ISO	1X8M	2.1	3	ISO
1O5X	9,75	22	ISO	1HOT	3,65	6	ISO	1P5Q	2.1	3	ISO
1WOB	9,75	22	ISO	1FS5	3,65	6	ISO	1I3N	2.1	3	ISO
1M7O	9,75	22	ISO	1TCO	3,64	7	ISO	2OOH	2.1	3	ISO
1M7P	9,74	22	ISO	1GR0	3,64	7	ISO	2O5E	2.1	3	ISO

1IV8	9,66	24	ISO	1ZLI	3,64	7	MCP	1YWK	2.1	3	ISO
2BTM	9,61	25	ISO	1JOF	3,64	8	ISO	2OOW	2.1	3	ISO
1AU0	9,6	20	CEP	1UWY	3,63	6	MCP	1YPI	2.1	3	ISO
1P5G	9,57	26	ISO	2P5Y	3,62	7	ISO	1I3L	2.1	3	ISO
1HG3	9,56	23	ISO	1SEU	3,6	6	ISO	1FAX	2.1	3	SP
1MEM	9,51	18	CEP	1WLT	3,6	6	ISO	1IVY	2.1	3	ISO
1NSU	9,49	39	ISO	2IFY	3,6	8	ISO	1LTO	2.1	3	SP
1KFU	9,37	20	CEP	1TTJ	3,6	7	ISO	1UDA	2.1	3	ISO
1IIP	9,37	29	ISO	1BML	3,6	8	SP	2DMM	2.1	3	ISO
2BRJ	9,33	20	ISO	2GYI	3,56	7	ISO	1WDL	2.1	3	ISO
1AW2	9,28	19	ISO	1PCJ	3,55	7	ISO	1J5S	2.1	3	ISO
1SUX	9,27	19	ISO	1CY8	3,5	6	ISO	1A9Z	2.1	3	ISO
1JX1	9,26	22	ISO	2C5E	3,49	7	ISO	2OOZ	2.1	3	ISO
1R2R	9,25	19	ISO	2JFO	3,49	7	ISO	1I3K	2.1	3	ISO
2VEN	9,24	19	ISO	2H4L	3,49	7	ISO	1D8W	2.1	3	ISO
1TPH	9,21	18	ISO	1YA7	3,48	7	TEP	1UDB	2.1	3	ISO
2GKE	9,2	28	ISO	1YAU	3,48	7	TEP	1EF7	2.1	3	CCP
1SW0	9,2	18	ISO	1D6M	3,45	6	ISO	1GI7	2.1	3	SP
2F8M	9,18	31	ISO	2VA6	3,45	6	AEP	1IST	2.1	3	ISO
1BTM	9,16	25	ISO	1A31	3,45	8	ISO	1D6M	2.1	3	ISO
1XLB	9,15	25	ISO	1S5N	3,43	7	ISO	1ZJA	2.1	3	ISO
2V2D	9,14	19	ISO	2PJM	3,43	7	ISO	4GSA	2.1	3	ISO
15 P / 61 = 24,6%				14 P / 61 = 22,9%				10 P / 56 = 17,85%			
12 SP / 16 = 75%				8 SP / 16 = 50%				9 SP / 16 = 56,25%			

Tableau 8 : Résultats comparatifs des trois tests SuMo sur le jeu des protéases et isomérases

Les résultats obtenus grâce aux nouveaux objets sont significativement meilleurs que lorsque l'on utilise les anciens objets. Cette amélioration prend toute son importance dans les capacités du logiciel à différencier les sérines protéases des autres protéines comparées. En effet, et sachant que la « query » correspond à une sérine protéase, le logiciel devrait classer en premier les seize sérine protéases de la base. Devraient ensuite venir les 45 autres protéases puis enfin 970 isomérases. Avec les anciens objets, 14 protéases parmi les 61 sont retrouvées (soit 22.9% de bonnes réponses), et seules 8 sérines protéases sur les 16 sont classées. Avec les nouveaux objets,



15 protéases sur 61 sont retrouvées, soit 24.6% de réponses correctes (ce qui n'est pas très différent de l'ancienne version), par contre, 12 sérines protéases sur 16 sont classées et elles le sont avec les plus forts scores comme le montrent les douze premières entrées qui sont toutes des sérine protéases. Ainsi, avec les nouveaux objets, SuMo a pu retrouver 75% des sérine protéases alors que les anciens ne permettaient d'en retrouver que 50%, par ailleurs moins bien classées (quatre SP sont bien en premières positions, mais les autres sont dispersées parmi les isomérases). Non seulement les nouveaux paramètres de SuMo classent en premier 12 sérines protéases parmi les 16 de la base, mais rappelons-le parmi plus de 1000 protéines.

A travers ce nouvel exemple, nous constatons que le bruit de fond lié au « chemical shift » et la redondance d'information attachée aux objets descripteurs de liaisons hydrogène est effacé par les nouveaux paramètres que nous avons suggéré. D'autre part, sachant que nos résultats sont meilleurs puisqu'ils permettent de retrouver plus de protéases, et notamment de placer en premières positions la majorité des sérine protéases, nous concluons que les contraintes liées aux vecteurs de l'ancien dictionnaire n'apportent rien puisqu'en les supprimant nous ne perdons pas d'information, au contraire, nous sommes plus performants. Cependant, nous constatons les limites du nouveau dictionnaire à travers la très faible amélioration concernant le nombre de protéases totales récupérées parmi les isomérases.

## **V- Critique globale et conclusions**

L'approche tout à fait novatrice et unique du logiciel SuMo a démontré, à travers de nombreuses études personnelles, ou collaboratives, dont certaines sont en cours de valorisation à travers une publication (Chemelle et al., à publier), ses capacités de comparaison de structures protéiques. Les applications sont extrêmement nombreuses et touchent divers domaines scientifiques, que ce soit sur un plan très fondamental, ou plus appliqué. Par exemple, sur le sujet présenté dans ce travail de thèse, et nous développerons ce point dans les chapitres 4,5 et 6, nous avons cherché les zones d'interaction entre l'antigène (la bêta-lactamine) et les structures de reconnaissance spécifiques que sont les anticorps IgE et les TCR. L'aspect fondamental concerne la compréhension des mécanismes moléculaires de reconnaissance de l'antibiotique, tandis que l'aspect appliqué a permis d'analyser quelles protéines sériques pouvaient jouer un rôle de

protéine d'hapténisation dans le but ,entre autre, d'améliorer les tests *in vitro* de diagnostic. Les exemples-tests développés dans le paragraphe précédent montrent clairement les capacités de SuMo à rechercher au sein d'une base de données plus ou moins grande, des points communs entre les modèles de cette base. Rappelons que les requêtes que l'on peut soumettre au logiciel sont extrêmement variées, allant d'un site très précis d'une protéine donnée comparé à des zones délimitées sur d'autres protéines (ou un seule autre), à une comparaison à large échelle de toutes les protéines les unes par rapport aux autres. Il faut aussi noter la flexibilité de ce système : pour chaque étude, il est possible de créer une base de données spécifique, allant de quelques modèles à plusieurs milliers, et de demander au logiciel de retrouver les communautés chimiques de surface sur ces protéines.

Encore une fois, les qualités de SuMo d'origine ne sont absolument pas remises en question, et cette partie du travail de thèse s'appuie très fortement sur le potentiel initial du logiciel. Toutefois, le développement du nouveau dictionnaire permet une amélioration significative des capacités de SuMo comme les deux tests le montrent. La suppression du « chemical shift », des vecteurs directionnels, des objets « delta minus » et « plus » sur-décrivant les liaisons hydrogène, et des objets spécifiques à un acide aminé (tel que l'objet proline), permet d'augmenter les capacités de comparaison du logiciel et de mieux discerner les sites les plus proches en terme de fonctions chimiques de surface. Enfin, le nouveau dictionnaire est plus cohérent dans la définition des surfaces protéiques par les nouveaux objets, les applications possibles sont donc plus nombreuses et surtout, elles ne nécessitent pas de redéfinir au cas par cas les poids attribués aux objets en fonction des protéines étudiées. Ainsi SuMo est à la fois plus performant, plus pertinent et plus polyvalent.

# Chapitre 4 : Classification et réactivité moléculaire des bêta-lactamines

## I- Classification des bêta-lactamines - Relation Structure-Activité:

I-I- Création d'un dendrogramme pour aider au choix d'une molécule de remplacement chez un patient hypersensible :

Les bêta-lactamines, dont on distingue cinq groupes en fonction du cycle accolé au noyau commun bêta-lactame (pénicillines, céphalosporines, carbapénèmes, monobactames et inhibiteurs de bêta-lactamases) sont les antibiotiques les plus fréquemment accusés dans les phénomènes d'hypersensibilités allergiques médicamenteuses.

A partir d'une série de 42 antibiotiques disponibles pour les hôpitaux européens et ayant un motif chimique central: le noyau bêta-lactame, nous avons réalisé plusieurs études par modélisation moléculaire. Ces molécules sont représentatives des cinq groupes chimiques de  $\beta$ -lactamines (pénicillines, céphalosporines...). Les structures chimiques se trouvent en annexe, dans le tableau 18.

L'ensemble des molécules a été modélisé à l'aide du logiciel Sybyl (Tripos Inc.) et du champ de forces MMFF94. Chaque modèle a été relaxé en énergie afin de trouver sa conformation énergétique minimale. A partir de cet ensemble de modèles, nous avons cherché à construire un dendrogramme dans le but de discriminer les molécules non plus en fonction de leurs familles chimiques, mais selon leurs champs moléculaires. Afin de prendre en compte l'aspect tridimensionnel, nous avons utilisé les méthodes CoMFA et CoMSIA.

Avant de pouvoir réaliser le dendrogramme, les différents modèles moléculaires ont été alignés selon deux approches:

- Alignement rigide sur motif commun: l'alignement des molécules est alors réalisé à partir du noyau bêta-lactame. Les chaînes latérales nécessitent un réalignement manuel, sur une position commune, en respectant l'énergie interne de chaque modèle. Cette approche manuelle a permis d'obtenir un premier alignement, de grande qualité. L'alignement final apparaît sur la figure 12.

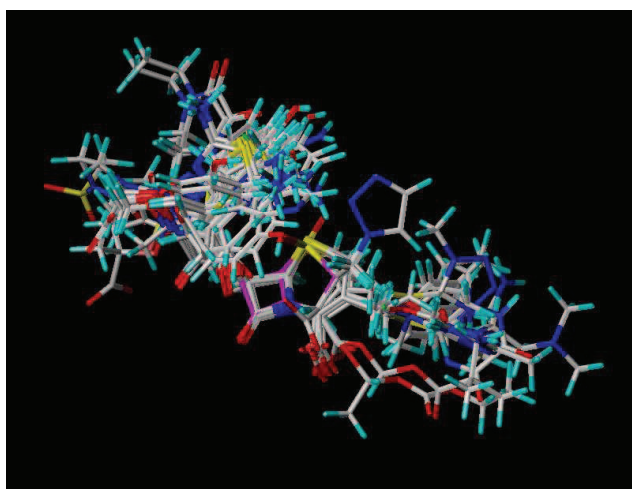


Figure 12 : Alignement des 42 antibiotiques sur le noyau  $\beta$ -lactame

- Alignement flexible sur un pharmacophore: dans un premier temps, nous avons construit un pharmacophore, que l'on peut, dans notre étude, qualifier d'allergophore, et ce à partir des 5 molécules les plus allergisantes de la série (amoxicilline, ampicilline, pénicilline G, pipéracilline et bacampicilline). Dans un second temps, les 37 autres molécules ont été alignées sur le pharmacophore de façon flexible par le programme. L'allergophore a été calculé sur l'ensemble de chaque molécule, mais le noyau central s'est révélé prépondérant. Le résultat final est peu différent de celui obtenu avec la première méthode, mais les chaînes latérales étant moins bien alignées, nous avons préféré continuer l'étude avec la première approche. L'alignement final par la méthode du pharmacophore est présenté sur la figure 13.

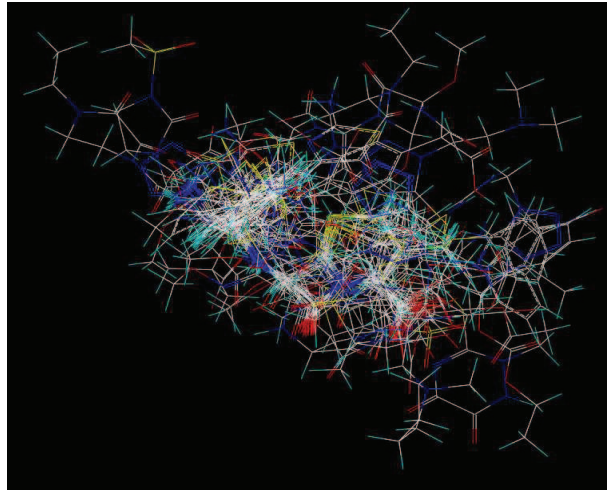


Figure 13 : Alignement des 42 antibiotiques sur le pharmacophore

Sur la base du premier alignement obtenu, nous avons ajouté une grille de potentiels afin de calculer les champs moléculaires de chacune des 42 molécules. Par la suite, ces champs sont traités par PLS (Partial Least Square) afin de déterminer un dendrogramme de différenciation. En partant de l'alignement rigide, nous avons choisi de former 6 familles, dont 2 petites de 2 membres qui ont été éliminées car trop peu représentatives. Le dendrogramme est présenté sur la figure 14 qui suit. Nous ne présenterons pas ici le dendrogramme obtenu à partir de l'alignement flexible car il était moins proche de la réalité clinique. Le dendrogramme se décompose en quatre sous-familles distinctes :

La première sous-famille ne regroupe que des céphalosporines de troisième générations (sauf une de deuxième) qui sont très rarement impliquées dans des allergies immédiates.

La deuxième sous-famille qui contient notamment l'amoxicilline, est plus hétérogène en terme de taux d'allergénicité et de sous-groupes chimiques. Notons toutefois que l'on n'a pas de céphalosporine de 3<sup>ème</sup> génération.

La troisième sous-famille relativement hétérogène regroupe des pénicillines, des monobactames, des inhibiteurs de  $\beta$ -lactamases, mais pas de céphalosporine.

La dernière sous-famille plus homogène présente surtout des céphalosporines de deuxième et troisième génération.

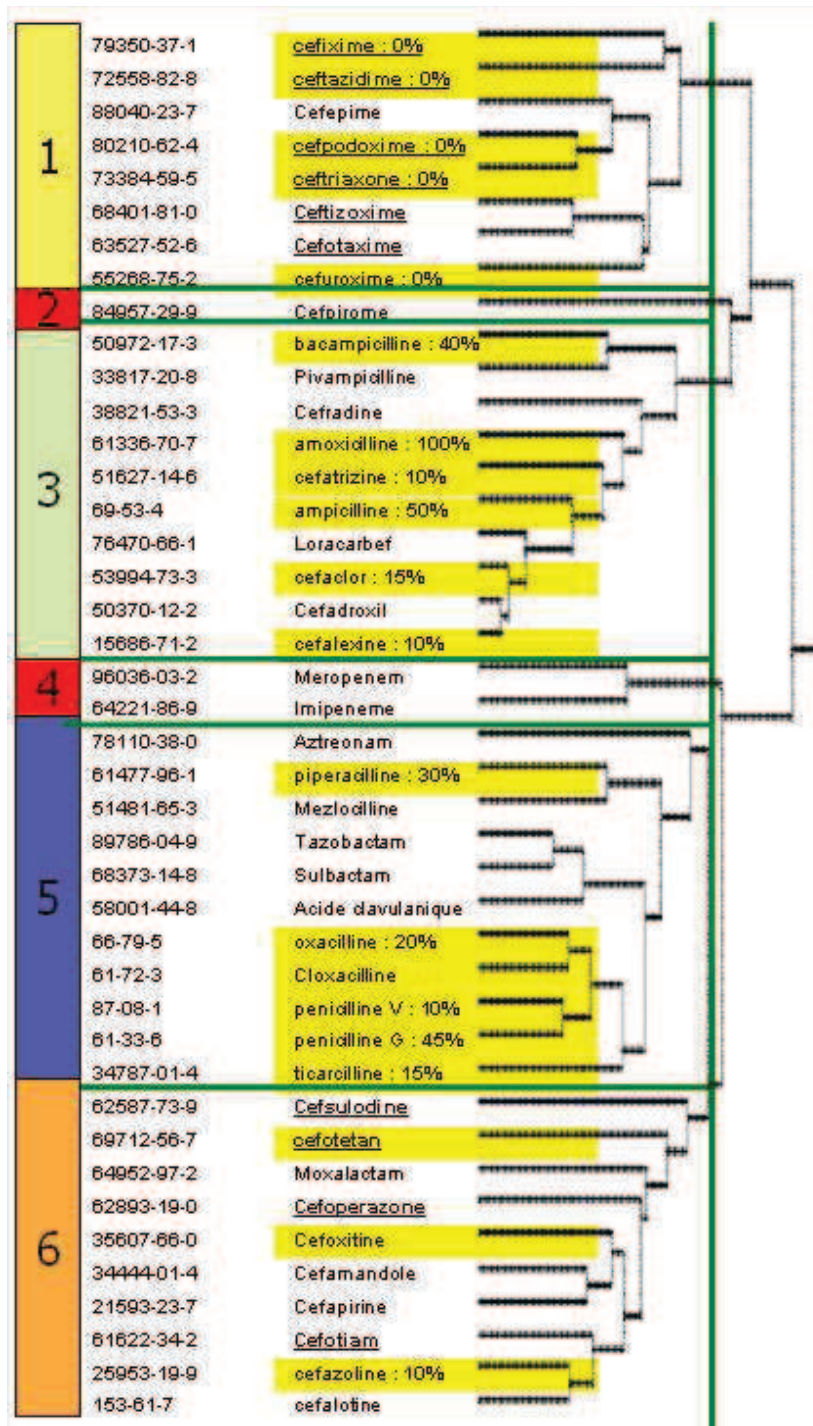


Figure 14 : Dendrogramme obtenu par alignement rigide de 42 molécules antibiotiques

Il est particulièrement utile pour aider les praticiens allergologues dans le choix d'un nouvel antibiotique lorsque le patient présente une hypersensibilité à la molécule initialement prescrite. De plus, cette classification permet d'étudier les phénomènes d'HSA croisées entre différentes molécules de la famille des bêta-lactamines.

Ce travail, valorisé par un article en cours de publication, a donné lieu à une analyse plus poussée et une validation qui seront développées dans le chapitre 5.

## I-2- Création de relations structure-activité quantitatives :

L'étude 3D-QSAR s'appuie sur les données cliniques issus de patients testés durant les « semaines HyperSensibilité Immédiate » à l'hôpital Lyon Sud. Rappelons que le dendrogramme, lui, ne tient pas compte des données patients. Nous avons sélectionné des molécules dans chacune des quatre familles du dendrogramme. Nous obtenons un jeu de 17 molécules testées par la méthode de Prick-Test. L'activité des molécules correspond au pourcentage d'allergénicité. La « batterie » de médicaments testée est fonction, d'une part de la possibilité de tests sur un patient (disponibilité des produits, et batterie de taille raisonnable), et d'autre part, permet de représenter chaque familles du dendrogramme. Notons que les patients sont sélectionnés sur le fait qu'ils ont déjà fait une réaction d'HSA Immédiate à une pénicilline : l'amoxicilline. Ainsi, les taux d'allergénicité correspondent à des réactions allergiques croisées entre l'amoxicilline et une autre bêta-lactamine. Le dendrogramme de la figure 14 montre ces pourcentages.

Nous avons réalisé l'étude QSAR en utilisant les méthodes CoMFA et CoMSIA. Pour chacune des deux méthodes nous avons validé la prédictivité de notre modèle grâce à une validation croisée (leave one out). Nous avons obtenu un coefficient de corrélation  $r^2$  de 0,752 par la méthode CoMSIA et un  $r^2 = 0,779$  par la méthode CoMFA. Les résultats montrent dès lors que notre étude va prédire le taux d'allergénicité des molécules antibiotiques avec une bonne fiabilité. Nous pouvons également analyser les zones qui sont les plus importantes pour l'activité des molécules, et qui auront tendance à les augmenter, ou à les diminuer. Ces zones sont présentées sur la figure 15. Les volumes bleus et verts sont plutôt « favorables » aux manifestations d'hypersensibilité, alors que les volumes rouges et jaunes sont « protecteurs ». On observe que la zone majeure induisant un phénomène allergique est centrée autour de l'azote quaternaire de l'une des chaînes latérales alors que l'autre chaîne a plutôt un rôle protecteur. L'implication de l'azote quaternaire rappelle le fait que les ions ammonium, trouvés dans de nombreux médicaments, produits cosmétiques, désinfectants..., constituent un motif d'allergénicité dominant trouvé dans l'environnement. De plus, ces résultats semblent confirmer le fait

que les antibiotiques de type céphalosporine de 3<sup>ème</sup> génération n'induiraient que peu de réactions allergiques, car ils ne possèdent pas d'azote quaternaire et disposent d'une chaîne latérale « protectrice ».

Les résultats obtenus *in silico* sont en accord avec les résultats expérimentaux qui nous ont été donnés par l'équipe du Pr. Jean François Nicolas lors de « Semaines Hypersensibilité Immédiate » sur une quarantaine de patients.

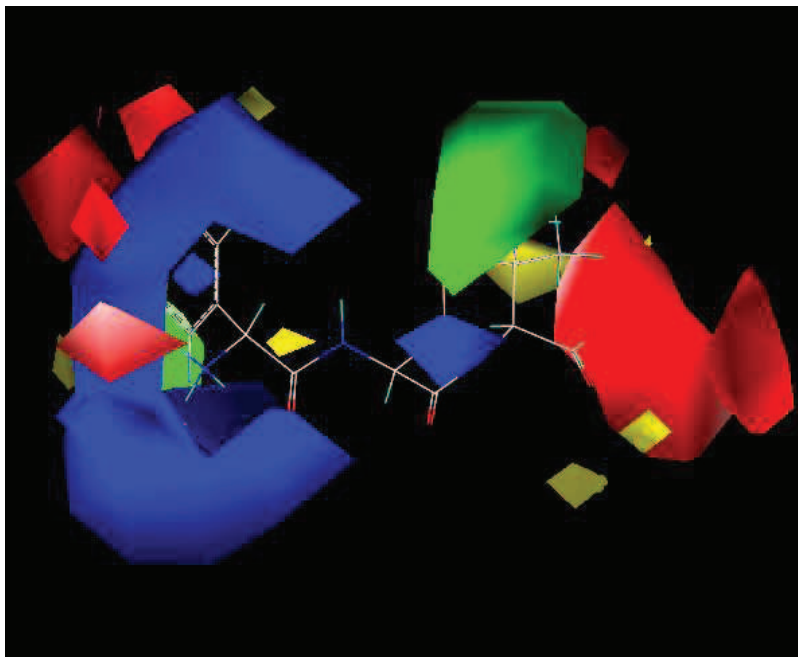


Figure 15 : Surfaces d'isoprépondérances des descripteurs sur une bêta-lactamine

## II- Réactivité moléculaire :

Le dendrogramme obtenu montre que certaines familles de molécules présentent un taux très homogène d'allergénicité croisée (famille 1 : taux nuls), alors que d'autres regroupant des molécules extrêmement proches au niveau de leurs champs moléculaires ont des taux très différents d'allergénicité. Cette variabilité nous a poussé à émettre l'hypothèse que le noyau commun des pénicillines devait posséder une réactivité différente du noyau commun des céphalosporines.

Les études immunochimiques antérieures (Landsteiner, 1962) ont prouvé l'absence d'antigénicité pour les petites molécules chimiques, c'est-à-dire de masse moléculaire de



500 à 1000 Daltons. Les antibiotiques, à l'instar de la majorité des médicaments, des cosmétiques, des colorants... sont des xénobiotiques qui nécessitent la liaison chimique, covalente et stable à une protéine endogène afin de devenir immunogène. Ainsi, les  $\beta$ -lactamines devront, pour induire une réponse immune spécifique, qu'elle soit humorale ou cellulaire, se comporter comme des haptènes et se coupler à une protéine porteuse sérique ou tissulaire (le plus souvent autologue, comme l'albumine plasmatique) (Lafaye et al. 1988). L'allergène prend alors le nom d'haptène. La liaison à la protéine est irréversible et très souvent covalente. Il faut noter que seuls quelques médicaments ont un poids moléculaire suffisant pour se comporter comme des antigènes complets, capables d'induire la sensibilisation d'un patient. Il s'agit généralement de protéines comme l'insuline, la corticotrophine et autres hormones (Bousquet et al. 1993).

Nous avons choisi d'étudier deux familles de  $\beta$ -lactamines les plus importantes: les pénicillines avec l'amoxicilline qui est la plus prescrite et les céphalosporines avec la ceftriaxone (troisième génération) qui est délivrée en cas d'HSA à l'amoxicilline. Ce choix prend appui sur le fait que les HSA à l'amoxicilline semblent être beaucoup plus fréquentes que les HSA à la ceftriaxone.

En nous basant sur l'ensemble des données de la littérature, nous avons donc cherché à approfondir l'hypothèse selon laquelle la bêta-lactamine se fixe de façon covalente sur une protéine porteuse endogène.

Pour cela, nous avons utilisé la structure co-résolue de la Protéine Fixant la Pénicilline avec un antibiotique de type  $\beta$ -lactamine: la cefuroxime (G. Gordon et al. 2000) et avons pu analyser le site d'interaction de la PFP avec le médicament. Nous avons par la suite cherché un ou plusieurs site(s) similaire(s) sur d'autres protéines en utilisant, en outre, notre approche SuMo.

Sachant que l'albumine humaine sérique est présente en forte concentration, et qu'elle est connue pour ses capacités à transporter des drogues, nous avons, au cours d'une première approche, étudié le ou les sites sur cette protéine qui fixeraient une ou plusieurs molécules d'antibiotique. Par la suite, nous étudierons l'ensemble des protéines susceptibles d'être hapténisées par le médicament (voir paragraphe III de ce chapitre). Ainsi, dans un premier temps, grâce aux techniques extrêmement précises de mécanique

quantique, nous avons étudié la réaction de fixation d'une partie du médicament sur une fonction lysine ou sérine. Dans un deuxième temps, nous avons recherché le(s) site(s) d'interaction au sein de l'albumine au moyen d'une étude de docking. Dans un troisième temps, nous avons modélisé la fixation de la bêta-lactamine complète sur une lysine de l'albumine entière, en utilisant la méthode mixte QM/MM décrite dans le chapitre 2.

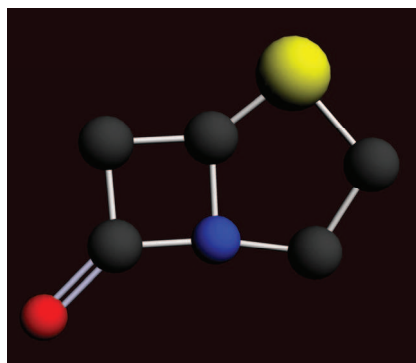
L'albumine humaine est une protéine plasmatique d'environ 65 kDa composée d'une seule chaîne polypeptidique de 585 acides aminés contenant 17 paires de ponts disulfures et une cystéine libre (Dugaiczky et al., 1982, ; Brown et Shockley, 1982). Constituée, comme les albumines d'autres espèces de trois domaines homologues (Brown, 1976), elle est produite exclusivement par le foie. Elle représente habituellement plus de la moitié des protéines plasmatiques. Sa fonction est double : assurer la régulation de la pression oncotique sanguine et transporter de nombreuses substances endogènes et exogènes. Elle possède en outre une forte affinité pour certains métaux, des hormones, des acides gras, la bilirubine, le calcium, et les drogues. Rappelons que les noyaux  $\beta$ -lactames des antibiotiques réagissent chimiquement avec les acides aminés de type lysine, sérine, voire histidine et cystéine, des protéines.

## II-1- Etude par Mécanique Quantique :

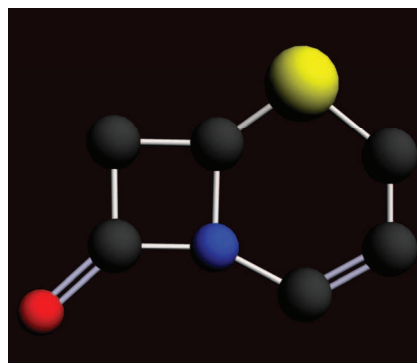
Comme nous l'avons vu, la mécanique quantique est une technique extrêmement précise qui autorise les ruptures et formations de liaison, mais dont les temps de calculs sont très longs et nécessitent des systèmes chimiques beaucoup plus petits qu'en Mécanique Moléculaire. Il existe plusieurs sous famille de méthodes QM et les résultats les plus fins sont issus de méthodes tenant compte de la corrélation inter-électronique ; ce sont soit les méthodes post Hartree-Fock, soit la méthode Density Functional Theory. Les méthodes post HF étant beaucoup plus lentes pour un résultat équivalent, nous avons sélectionné la méthode DFT.

Afin d'alléger le système, nous n'avons conservé que les deux noyaux centraux des pénicillines et des céphalosporines, notés noyau-P pour le modèle des pénicillines et noyau-C pour les céphalosporines. De la même façon, nous avons réduit les modèles de la lysine et de la sérine. Pour la lysine nous avons choisi la méthylamine car la chaîne latérale de cet acide aminé est constituée de 4 atomes de carbone consécutifs reliés par de simples

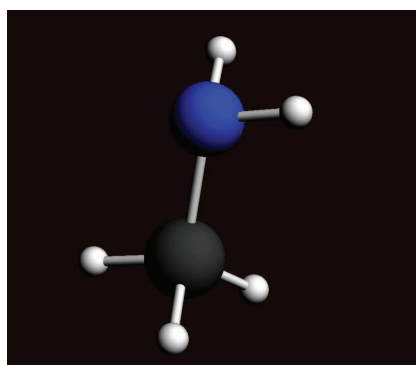
liaisons, et qu'il n'y a aucun effet électronique délocalisé. Pour la sérine, nous avons utilisé comme modèle et pour les mêmes raisons, le méthanol. La figure 16 a ci-dessous présente l'ensemble de nos modèles.



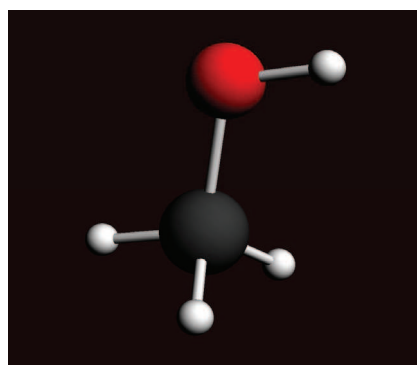
**Noyau-Pénicilline (P)**



**Noyau-Céphalosporine (C)**



**Modèle Lysine (méthylamine)**



**Modèle Sérine (méthanol)**

Figure 16 a : Modèles utilisés pour l'étude par mécanique quantique

Les réactions ici étudiées se décomposent en trois étapes chimiques clefs correspondant chacune à une énergie propre :

- Les **réactifs** : étape qui correspond à la somme des énergies des réactifs (ici l'énergie de la partie médicament additionnée à l'énergie de la partie lysine ou sérine)
- L'**intermédiaire réactionnel** : ou état de transition, très instable pour lequel l'énergie du système est la plus grande des trois états.
- Les **produits** : étape qui correspond à la partie médicament fixée sur la partie acide aminé. Cet état thermodynamiquement stable est celui pour lequel le système étudié est de plus basse énergie.

Le schéma ci-dessous présente classiquement la courbe énergétique des réactions chimiques.

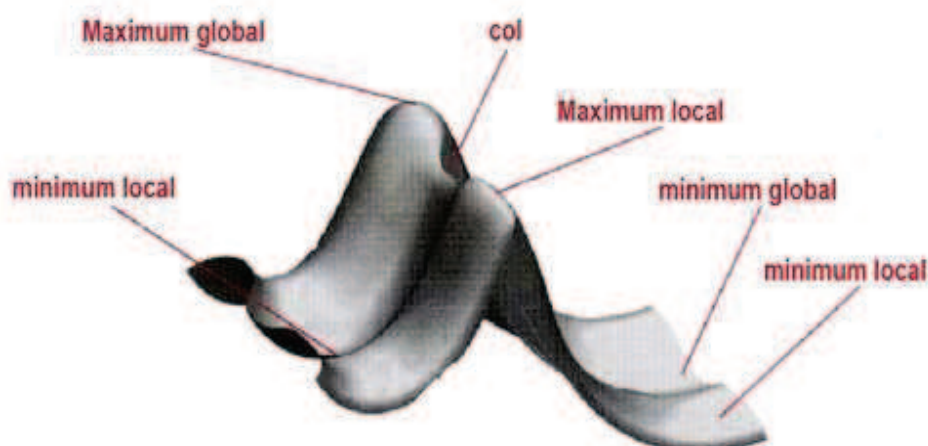


Schéma 5: courbe énergétique classique des réactions chimiques.

La différence d'énergie entre les réactifs et l'état de transition correspond à l'énergie d'activation nécessaire pour que la réaction chimique se produise. Sachant que plus l'énergie d'activation est forte, plus la réaction est difficile à réaliser (du point de vue cinétique), nous avons comparé ces énergies d'activation dans le cas du noyau-P vis-à-vis des fonctions  $\text{NH}_2\text{CH}_3$  (méthylamine) ou  $\text{CH}_3\text{OH}$  (méthanol) versus le noyau-C face aux fonctions  $\text{NH}_2\text{CH}_3$  ou  $\text{CH}_3\text{OH}$ .

Le doublet non liant de l'azote ou l'oxygène va attaquer le carbone chargé partiellement positivement de la fonction carbonyle (du cycle bêta-lactame) et déplacer ainsi le doublet sur l'oxygène du cycle. Dans un second temps, le doublet va revenir sur le carbone pour détruire la liaison amide du cycle  $\beta$ -lactame. Le schéma 7 représente la réaction étudiée.

Il faut remarquer que la réaction entre l'acide aminé et l'antibiotique résulte d'une attaque nucléophile. De ce fait, si c'est une lysine qui réagit, elle sera dans un état non protoné. En effet, dans le cas où elle serait sous forme protonée, ( $\text{R-NH}_3^+$ ), le doublet non liant porté par l'hétéroatome ne serait plus disponible pour attaquer le carbone du cycle bêta-lactame, et la réaction serait impossible. Ainsi, l'environnement immédiat en acides

aminés est important et doit être suffisamment basique pour que les lysines soient non protonées.

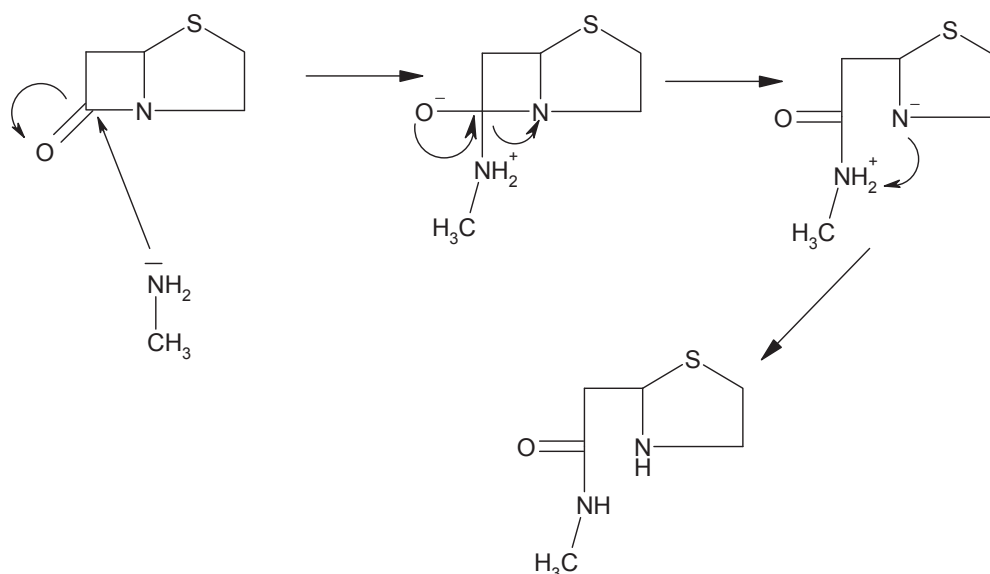


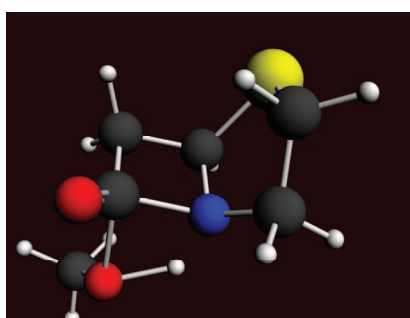
Schéma 6 : Réaction chimique entre le noyau-P et la méthylamine

Notre étude de QM a fait appel au logiciel Amsterdam Density Functional. Nous avons utilisé la DFT avec une base atomique de Triple Potentiel Zêta avec 2 fonctions de polarisation (TZ2P).

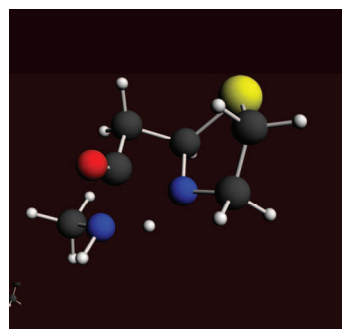
Nous avons modélisé indépendamment les réactifs et les produits et minimisé énergétiquement chacun des modèles en utilisant la méthode quantique décrite précédemment. Pour la géométrie de l'intermédiaire réactionnel le calcul est plus complexe. Nous avons fait, dans un premier temps, une étude de cet intermédiaire par une recherche linéaire de transfert des atomes : « Linear transit ». Pour notre cas, nous avons défini une contrainte de distance entre le carbone du carbonyle et l'azote ou l'oxygène de l'acide aminé. Cette distance est divisée en 10 étapes. Pour chacune d'entre elles, le système est optimisé géométriquement. Dans un second temps, pour la conformation de plus haute énergie, nous avons fait calculer le spectre vibrationnel et avons vérifié que la fréquence imaginaire unique (fréquence négative) correspondait bien au mouvement réactionnel. Une fois ce calcul effectué, nous avons utilisé l'algorithme de recherche d'intermédiaire réactionnel. Cet algorithme, afin de trouver la conformation de l'intermédiaire réactionnel qui correspond à un col du point de vue énergétique, va

minimiser toutes les directions de l'espace exceptée la direction de la fréquence négative qu'il va maximiser. Dans une dernière étape, nous avons utilisé l'algorithme de suivi de chemin réactionnel dans les deux sens (IRC) pour vérifier que nous avons bien trouvé le seul intermédiaire réactionnel.

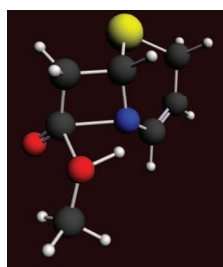
La géométrie des intermédiaires réactionnels est illustrée au niveau de la figure 16 b. On remarque que l'attaque nucléophile par l'oxygène du méthanol (cas de la sérine) ou de l'azote de la méthylamine (cas de la lysine) est dirigée vers le carbone du cycle bêta-lactame. L'hétéroatome (oxygène du méthanol ou azote de la méthylamine) se situe au dessus ou en dessous du plan selon un angle d'environ 80°. D'autre part, un des atomes d'hydrogène de la fonction amine (méthylamine) ou l'hydrogène de la fonction hydroxyle (méthanol) se place entre les atomes de carbone et d'azote du cycle bêta-lactame, en vue de l'insertion lors de la rupture de la liaison amide du cycle. La distance entre le carbone du carbonyle (cycle bêta-lactame) et l'oxygène du méthanol est de 2,08 Å Elle est de 1,94 Å entre le carbone (carbonyle) et l'azote de la méthylamine.



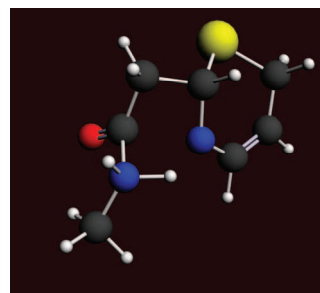
**Intermédiaire pénicilline (P) –  
méthanol**



**Intermédiaire pénicilline (P) –  
méthylamine**



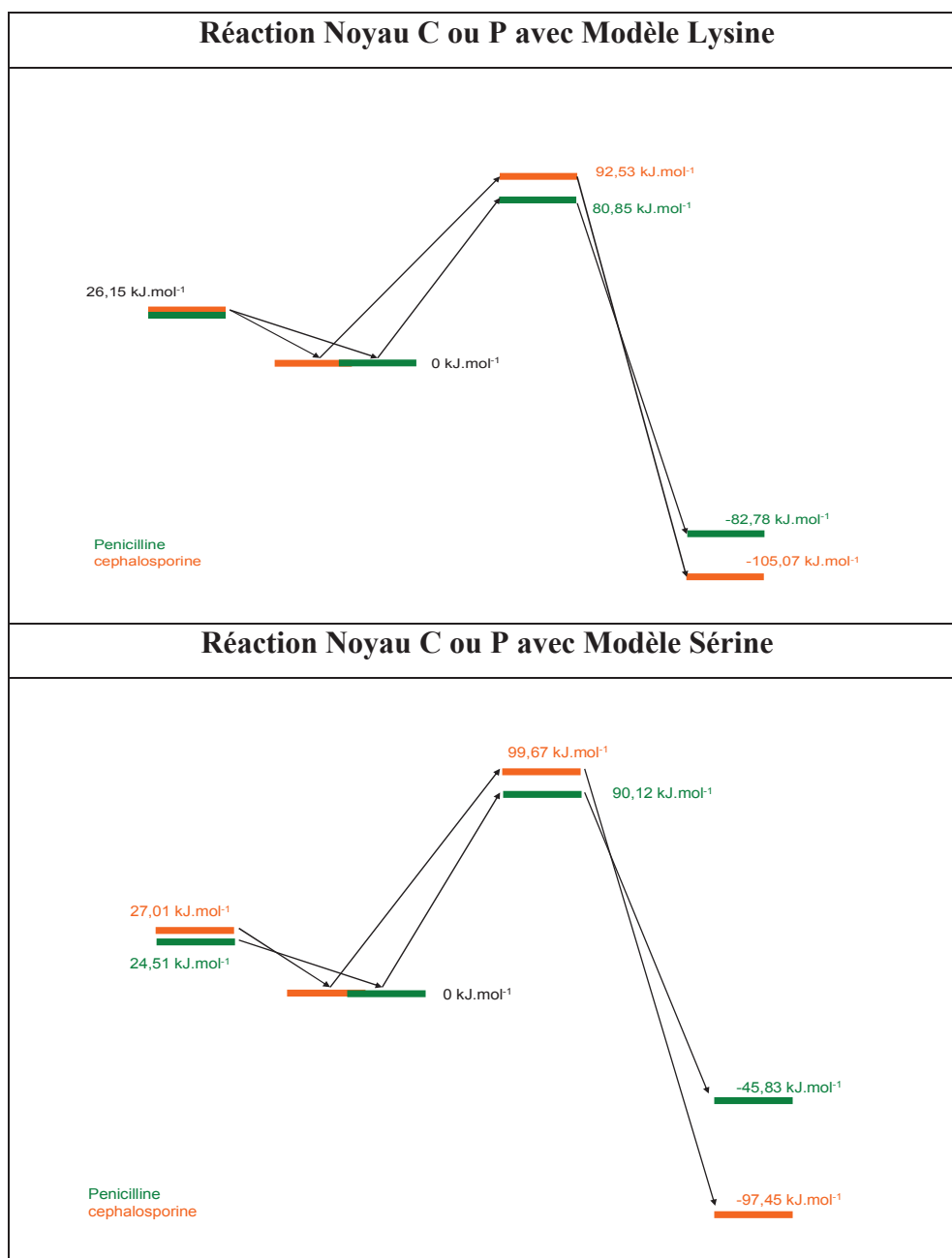
**Intermédiaire céphalosporine (C)  
– méthanol**



**Intermédiaire céphalosporine (C)  
– méthylamine**

Figure 16b : Structure des intermédiaires réactionnels calculés par mécanique quantique

Dans les schémas qui suivent, sont représentés les diagrammes énergétiques pour les réactions chimiques du noyau-C et du noyau-P avec les modèles de lysine et de sérine. Les énergies sont exprimées en  $\text{kJ}\cdot\text{mol}^{-1}$ . Nous avons quatre états d'énergie (de gauche à droite) correspondant aux réactifs, aux réactifs au contact, à l'intermédiaire réactionnel, et aux produits. Le vert code pour le noyau pénicilline, le orange pour le noyau céphalosporine.



Schémas 7 : Diagrammes énergétiques des réactions impliquant les noyaux-P et -C et les modèles de lysine et sérine.

Sur le schéma précédent, l'énergie est ramenée à zéro lors du contact entre les deux réactifs. En effet, les calculs DFT étant effectués dans le vide, les interactions électrostatiques favorisent le rapprochement des molécules, ce qui explique une énergie plus basse lorsque les molécules sont au contact.

L'analyse des schémas ci-dessus montre que la réaction est thermodynamiquement favorisée dans les quatre cas, comme en témoigne le fait que l'énergie des produits est plus faible que celle des réactifs. On remarque également que les deux produits issus des réactions avec les céphalosporines sont plus stables que dans le cas des pénicillines. Ceci est probablement lié à un meilleur recouvrement du doublet non liant de l'azote (méthylamine :  $\text{NH}_2\text{CH}_3$ ) ou de l'oxygène (méthanol :  $\text{CH}_3\text{OH}$ ) sur le cycle à 6 atomes qui dispose en alpha de la double liaison CC, alors que ce recouvrement n'existe pas pour les pénicillines.

Les résultats montrent que d'une part, l'énergie d'activation relative à la fixation de  $\text{NH}_2\text{CH}_3$  sur le noyau-P est plus faible que celle sur le noyau-C. De même, l'énergie d'activation de la fixation de  $\text{CH}_3\text{OH}$  sur le noyau-P est plus faible que celle sur le noyau-C. D'autre part, l'énergie d'activation dans le cas des fonctions de type lysine ( $\text{NH}_2\text{CH}_3$ ) est toujours plus faible que si l'on modélise les réactions sur une fonction de type sérine ( $\text{CH}_3\text{OH}$ ).

Les différences ici mentionnées sont, sur un plan physiologique, nettement significatives. En effet, nous avons calculé qu'il faut apporter au système  $11,7 \text{ kJ.mol}^{-1}$  de plus pour que la méthylamine fixe le noyau-C qu'il n'en faut pour qu'elle fixe le noyau-P. De même, il faut apporter  $9,6 \text{ kJ.mol}^{-1}$  de plus pour que le méthanol fixe le noyau-C qu'il n'en faut pour qu'il fixe le noyau-P. Enfin, si l'on compare les énergies d'activation sérine versus lysine, on constate qu'il faut  $9,3 \text{ kJ.mol}^{-1}$  de plus pour que le méthanol fixe le noyau-P qu'il n'en faut dans le cas de la méthylamine ; et de la même façon il faut  $7,2 \text{ kJ.mol}^{-1}$  de plus pour que le méthanol fixe le noyau-C qu'il n'en faut dans le cas de la méthylamine. Le tableau 9 présente la synthèse des données calculées.



Energies d'activation en kJ.mol <sup>-1</sup>	Noyau-P	Noyau-C
Modèle Lysine	80,85	92,52
Modèle Sérine	90,12	99,67

Tableau 9 : Energies d'activation des réactions de liaisons des noyau-C et-P sur la méthylamine et le méthanol.

De ces résultats, nous pouvons conclure que d'une part, les pénicillines sont plus réactives que les céphalosporines, et d'autre part que les lysines sont plus réactives que les sérines.

### II-3- Etude de la réaction de fixation sur l'albumine :

Au cours des études précédentes sur le sujet, nous avons montré que la réactivité des molécules de type pénicilline ou céphalosporine devait être différente. Nous avons classé les antibiotiques de la famille des bêta-lactamines non plus en fonction de leurs structures chimiques (pénicillines, céphalosporines...) mais en fonction de leurs champs de forces moléculaires, et obtenions 4 familles. Il est apparu que certaines familles de molécules présentaient un taux très homogène d'allergénicité, alors que d'autres regroupant des molécules extrêmement proches au niveau de leurs champs moléculaires avaient des taux très différents d'allergénicité. Cette variabilité nous a poussé à émettre l'hypothèse que le noyau commun des pénicillines devait posséder une réactivité différente du noyau commun des céphalosporines. Nous avons exploré la réactivité de ces deux familles de molécules grâce à la chimie quantique.

Afin d'obtenir des résultats les plus proches de la réalité physiologique, nous avons cherché à étudier le système complet c'est-à-dire composé de la molécule antibiotique et de la protéine d'hapténisation complète. Pour cela, nous avons fait appel à une approche que nous nommons : la méthode mixte ou méthode QM/MM, utilisant à la fois les méthodes quantiques et la mécanique moléculaire.

Cependant, dans une phase préliminaire, il nous faut connaître, le ou les sites susceptibles de fixer de manière covalente la drogue. Nous avons donc analysé les acides aminés accessibles sur l'albumine sérique humaine puis avons calculé l'énergie d'interaction entre le médicament et le(s) site(s) étudié(s). Cette partie du travail fait appel à la technique de docking.

### II-3-1 Docking :

Nous avons vu précédemment que la Protéine Fixant la Pénicilline était la cible des antibiotiques de type bêta-lactamine. Le médicament vient en effet se fixer de manière covalente sur une fonction chimique de la PFP. Ainsi, et en postulant que le caractère immunogène de l'antibiotique nécessite sa fixation stable à une protéine endogène, nous avons pris appui sur le site de reconnaissance et de fixation de la  $\beta$ -lactamine au sein de la PFP, dont nous avons récupéré la structure cristallographique sur la Protein Data Bank (1QMF, Gordon et al. 2000).

Sur la PFP, ce site possède la forme d'une tranchée légèrement coudée en son centre. Le fond en est tapissé notamment par deux résidus sérine et une lysine. C'est très probablement cette lysine qui va réagir avec le cycle  $\beta$ -lactame et déclencher ainsi la réaction de transamidation. La molécule antibiotique est stabilisée dans le site via une autre sérine qui crée une liaison hydrogène avec la fonction ammonium quaternaire. Le cycle phényle des pénicillines va donner une interaction de type «  $\pi$  stacking » avec un acide aminé de type tryptophane. De l'autre côté de la molécule, la fonction acide carboxylique crée plusieurs liaisons hydrogène avec le squelette de la protéine. Le modèle du site actif en interaction avec la molécule de cefuroxime (céphalosporine de deuxième génération) est exposé sur la figure 17.

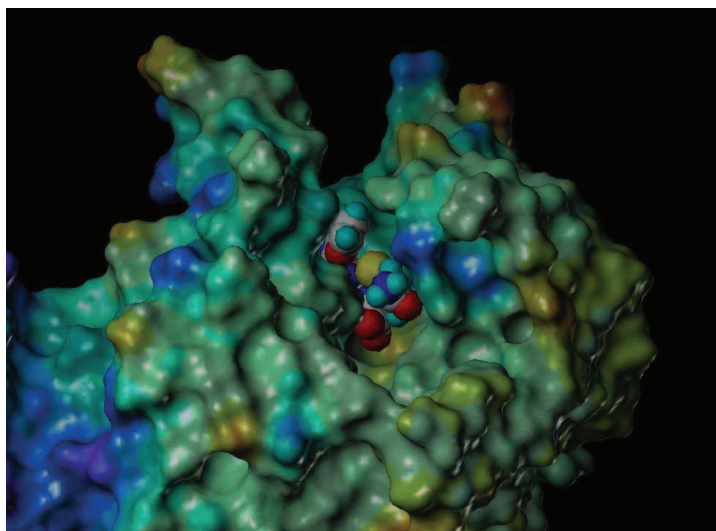


Figure 17 : Site actif de la PFP en interaction avec la molécule de cefuroxime

La structure de l'albumine humaine est déposée à la PDB sous le code 1AO6 (Sugio et al., 1999). Nous avons dans un premier temps détecté toutes les lysines, sérines et histidines accessibles au solvant grâce au programme Definition of Secondary Structure of Proteins (DSSP), en définissant un seuil d'accessibilité supérieur à 30% de la surface d'accès au solvant. Nous obtenons ainsi 41 lysines, 12 sérines et 11 histidines. Sachant que la molécule se fixe obligatoirement via l'un de ces 3 acides aminés, nous avons soumis une simulation de docking pour chacun des résidus accessibles. Ces analyses ont été effectuées avec le logiciel Sybyl X1.1 et s'appuient sur le module de docking « Surflex-Dock » qui est l'un des standards de l'industrie pharmaceutique, grâce notamment à sa fonction de score nommée « CScore ». Elle correspond à un consensus de plusieurs fonctions d'évaluations telles que : ChemScore, PMF\_score, D\_Score, G\_score, qui sont toutes des fonctions empiriques avancées capables de prendre en compte différents potentiels et d'approximer la variation d'énergie libre d'interaction. Nous n'entrerons pas dans les détails de ce type de fonctions, qui, malgré leur complexité mathématique, ne parviennent pas à donner un résultat assez fidèle. Aussi, Tripos Inc. a eu l'idée de toutes les réunir et d'en faire la moyenne. La figure 18 représente l'évaluation de l'affinité d'une molécule par certaines fonctions qui constituent le CScore.

<b>LUDI</b>	$\Delta G_{bind} = \Delta G_{H-bond} \sum_{H-bond} f(\Delta R, \Delta \alpha) + \Delta G_{ionic} \sum_{ionic} f(\Delta R, \Delta \alpha) +$ $\Delta G_{hydrophobic} \sum_{hydrophobic}  A_{hydrophobic}  + \Delta G_{rotor} \cdot N_{rotor} + \Delta G_0$ $A_{hydrophobic} = \text{molecular surface area}$
<b>F-Score</b>	$\Delta G_{bind} = \Delta G_{H-bond} \sum_{H-bond} f(\Delta R, \Delta \alpha) + \Delta G_{ionic} \sum_{ionic} f(\Delta R, \Delta \alpha) + \Delta G_{aromatic} \sum_{aromatic} f(\Delta R, \Delta \alpha)$ $+ \Delta G_{contact} \sum_{contact} f(\Delta R, \Delta \alpha) + \Delta G_{rotor} \cdot N_{rotor} + \Delta G_0$
<b>Chem-Score</b>	$\Delta G_{bind} = \Delta G_{H-bond} \sum_{H-bond} f(\Delta R, \Delta \alpha) + \Delta G_{metal} \sum_{metal} f(\Delta R, \Delta \alpha) +$ $\Delta G_{lipo} \sum_{lipo} f(\Delta R) + \Delta G_{rotor} \sum_{rotor} f(P_{ni}, P'_{ni}) + \Delta G_0$

Figure 18 : Formules d'évaluation de l'affinité d'une molécule utilisée par la fonction consensus CScore

Le programme permet également de prendre en compte la flexibilité des chaînes latérales des acides aminés, ce qui représente une augmentation des temps de calcul. Afin d'obtenir des simulations les plus exhaustives possibles, nous avons choisi de prendre en compte cette flexibilité au niveau de CScore, et avons augmenté le nombre de poses générées : pour chaque acide aminé catalytique, le médicament est placé suivant 30 positions différentes et la meilleure énergie d'interaction sera conservée. Pour toutes les simulations de docking, nous avons créé une sphère de sélection de 5 Å autour du résidu étudié, puis avons généré la zone d'interaction avec les sondes appelées « protomol ». Elles servent de point d'accrochage au fragment lors de la reconstruction « ad-hoc » du ligand dans la cavité suivant l'algorithme de Surflex-Dock. Une densité standard de sondes a été utilisée (0.5), cependant nous avons étendu la zone au maximum (10 Å) afin d'éviter les effets de bord. Les résultats sont exprimés sans unité pour la fonction CScore et exposés dans le tableau 10. La colonne 1 donne le numéro de la lysine étudiée. La colonne 2, le type de résidu. La troisième colonne présente l'accessibilité (en pourcentage par rapport à l'accessibilité maximum du résidu), la colonne 4 permet de sélectionner uniquement les acides aminés dont l'accessibilité dépasse les 30%, enfin, la dernière colonne donne les valeurs de CScore.

N° résidu	Type résidu	% Acc DSSP	N° résidu si accessibilité >30%	CScore
12	K	63	12	2,68
20	K	27		
41	K	44	41	5,41
51	K	46	51	4,84
64	K	28		
73	K	28		
93	K	45	93	5,30
106	K	9		
136	K	26		
137	K	46	137	5,20
159	K	39	159	4,98
162	K	39	162	5,64
174	K	28		
181	K	29		
190	K	33	<b>190</b>	6,02
195	K	48	<b>195</b>	4,98
199	K	17		
205	K	65	205	4,54
212	K	37	<b>212</b>	6,57
225	K	47	225	2,65
233	K	38	233	4,72
240	K	55	240	5,47
262	K	56	262	4,90
274	K	20		
276	K	56	276	5,05
281	K	51	281	2,78
286	K	23		
313	K	80	313	3,27
317	K	69	317	4,38
323	K	46	323	5,57
351	K	52	351	6,01
359	K	79	359	3,37
360	K	1		
369	K	22		
372	K	63	372	4,57
378	K	63	378	3,97
389	K	60	389	4,41
402	K	58	402	4,67
413	K	24		
414	K	5		
432	K	34	<b>432</b>	6,66
436	K	50	436	5,91
439	K	81	439	5,73

444	K	55	444	5,13
466	K	47	466	5,64
475	K	55	475	3,11
500	K	75	500	4,77
519	K	54	519	4,82
524	K	52	524	3,56
525	K	6		
534	K	9		
536	K	20		
538	K	100	538	2,31
541	K	85	<b>541</b>	4,32
545	K	55	<b>545</b>	5,68
557	K	62	557	2,96
560	K	62	560	5,38
564	K	84	564	5,42
573	K	71	573	4,91
574	K	79	574	2,47

Tableau 10 : CScore d'interaction entre l'amoxicilline (centrée sur la lysine numéro X) et l'albumine.

Nous ne présentons ici que les résultats de docking sur les lysines car nous avons démontré que celles-ci étaient plus réactives et devraient donc constituer la majorité des sites d'hapténisation. L'analyse des meilleurs scores (entrées grises) fait apparaître que quatre sites se détachent nettement : ceux des lysines 432, 212 et un peu plus faiblement les sites 190 et 351. Nous avons analysé par mécanique quantique le positionnement du médicament vis-à-vis d'une lysine et nous nous basons sur ces résultats pour étudier l'interaction de l'amoxicilline avec les lysines des différents sites. Nous remarquons que les sites 432, 212, 190 et 351 présentent les CScores les plus forts, ce qui signifie qu'ils sont les plus susceptibles en terme d'énergie d'interaction de réagir avec la drogue. Nous observons que ces sites sont très bien adaptés au ligand en terme de conformation et que l'ancrage est renforcé par un réseau de liaisons hydrogène entre les acides aminés et les fonctions  $\text{COO}^-$  et  $\text{NH}_3^+$ . Cependant, les sites 212 et 190 sont encore plus appropriés car d'une part l'énergie apportée par l'interaction molécule/lysine, et qui servira à activer la réaction chimique de fixation, est grande et que d'autre part, la position des lysines catalytiques vis-à-vis du médicament est optimale.

Les figures ci-après montrent l'amoxicilline en interaction avec le site K190 (figure 19) et K212 de l'albumine sérique humaine (Figure 20). Dans les deux cas, la lysine est en

contact étroit avec le cycle bêta-lactame et de ce fait idéalement placée pour que la réaction de fixation se produise. De nombreuses liaisons hydrogène sont développées en direction de la protéine à partir de la fonction acide carboxylique et de l'ammonium quaternaire de la drogue. La fonction hydroxyle portée par le cycle benzénique n'est engagée que dans le cas des sites 212 et 432. Les quatre sites étudiés de la protéine sont très hydrophiles et relativement basiques du fait du nombre important d'acides aminés de type arginine et lysine.

S'agissant du site K432, nous avons constaté qu'il était très ouvert. De ce fait, malgré un bon CScore, une réaction de fixation semble moins probable du fait d'une moins bonne stabilité du ligand dans son site. Cependant, il nous aurait fallu effectuer une simulation de dynamique moléculaire afin d'étudier les mouvements de la molécule ainsi que ceux de la protéine et mesurer la distance entre l'azote et le cycle bêta-lactame. Nous sommes, dans ce cas, confrontés aux limites des méthodes de docking, pour lesquelles la protéine réceptrice est rigide, même si notre logiciel tient compte de la flexibilité des chaînes latérales ce qui constitue une performance très importante au regard des programmes plus anciens.

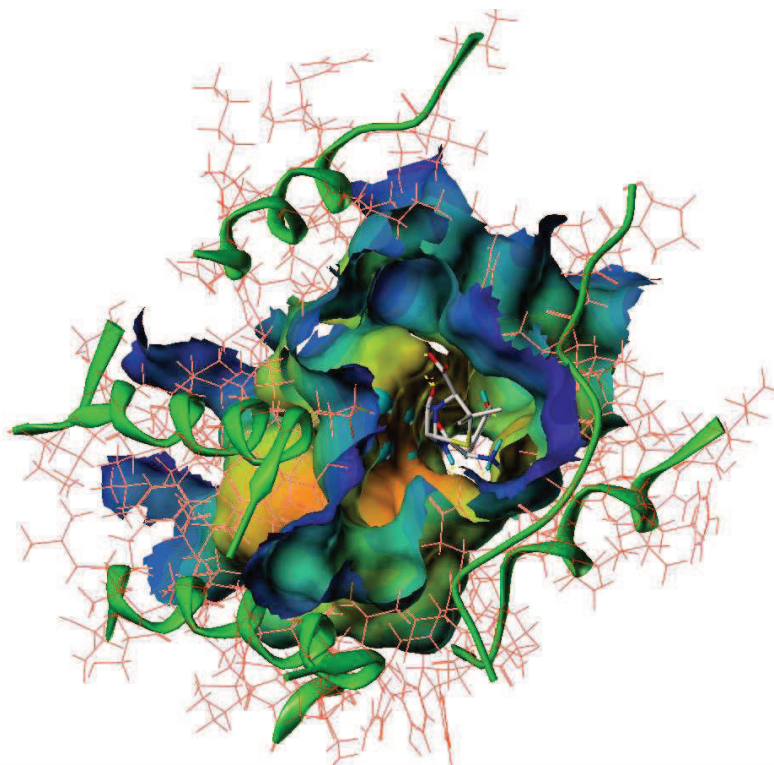


Figure 19 : Interaction entre l'amoxicilline et le site K190

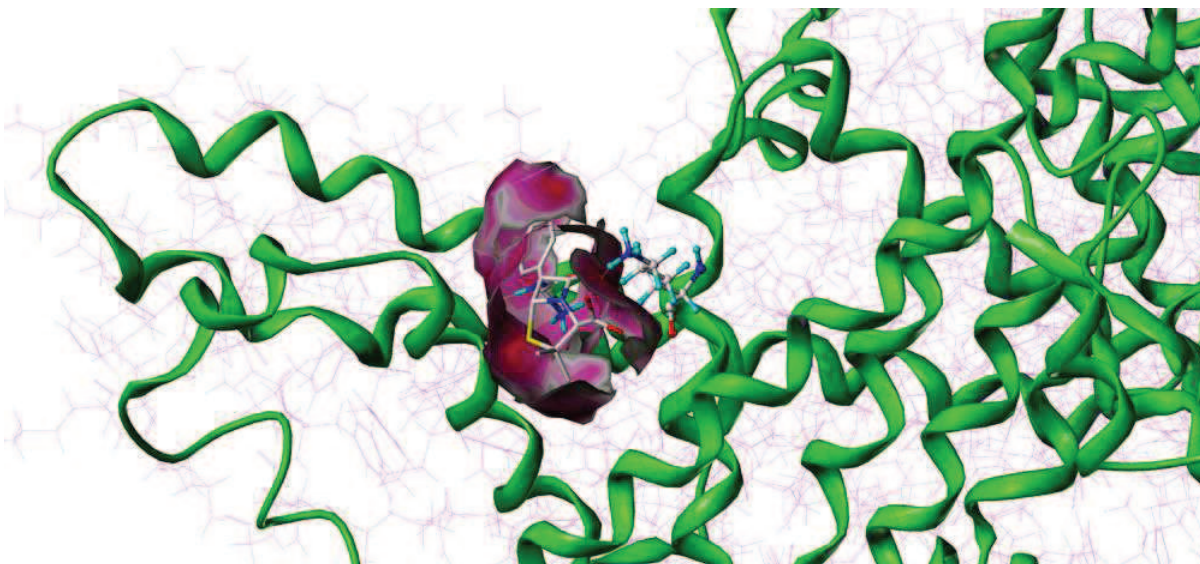


Figure 20: Interaction entre l'amoxicilline et le site K212

### II-3-2 Etude mixte Mécanique Quantique / Mécanique moléculaire :

Comme nous l'avons vu, la mécanique quantique offre l'avantage d'autoriser la formation de liaisons, ce dont nous avons besoin pour évaluer la réactivité des noyaux pénicilline et céphalosporine face aux modèles lysine et sérine. Cependant, et du fait que les temps de calculs des méthodes QM sont extrêmement longs, nous avons réduit considérablement notre système d'étude en nous affranchissant de la majeure partie de l'albumine et des chaînes latérales des bêta-lactamines. Nous avons mentionné plus haut le principe des méthodes mixtes QM/MM qui, dans notre cas, sont parfaitement adaptées. En effet, via la QM/MM, nous pouvons modéliser la réaction de fixation de la  $\beta$ -lactamine (amoxicilline ou ceftriaxone) complète sur un acide aminé précis de l'albumine entière. Nous présentons dans cette section les résultats concernant la lysine 190.

Nous obtenons, de la même manière que lors de nos expériences de QM, les énergies des 3 états clef et en déduisons alors l'énergie d'activation qu'il est nécessaire d'apporter au système pour que la fixation de l'antibiotique se fasse sur la protéine de transport.

Les atomes participant à la réaction chimique sont inclus dans la partie quantique tandis que les autres appartiennent à la partie mécanique. Nous avons inclus dans le domaine QM une partie de la chaîne latérale de la lysine 190 ( $\text{NH}_2\text{CH}_2$ ) ainsi que



l'ensemble des atomes de l'amoxicilline ou de la ceftriaxone. L'énergie des atomes du domaine MM est calculée par l'intermédiaire du champ de forces C.H.A.R.M.M. et l'énergie des atomes de la partie quantique utilise le même paramétrage que dans la phase précédente d'étude QM pure. Les calculs ainsi obtenus sont très proches des résultats de mécanique quantique. La variation d'énergie d'activation entre les céphalosporines et les pénicillines est identique pour l'analyse par QM/MM à celle obtenue par la QM. Les différents acides aminés qui interagissent avec l'antibiotique (autre que le résidu qui attaque le cycle) ne sont pas inclus dans la partie QM. Aussi, il est logique d'obtenir des résultats corroborant l'analyse quantique. D'autre part, nous ne pouvons évaluer ni l'effet d'activation électronique, ni l'abaissement de l'énergie d'activation, attribués aux acides aminés qui entourent le médicament.

Nous observons que l'énergie d'activation nécessaire à réaliser la réaction de fixation sur K190 sera plus importante dans le cas des céphalosporines comme la ceftriaxone ici testée, prouvant ainsi que ce type de  $\beta$ -lactamine est moins réactif que les  $\beta$ -lactamines de type pénicilline comme l'amoxicilline utilisée dans ce cas. La figure 21 ci-dessous représente l'albumine liant l'amoxicilline.

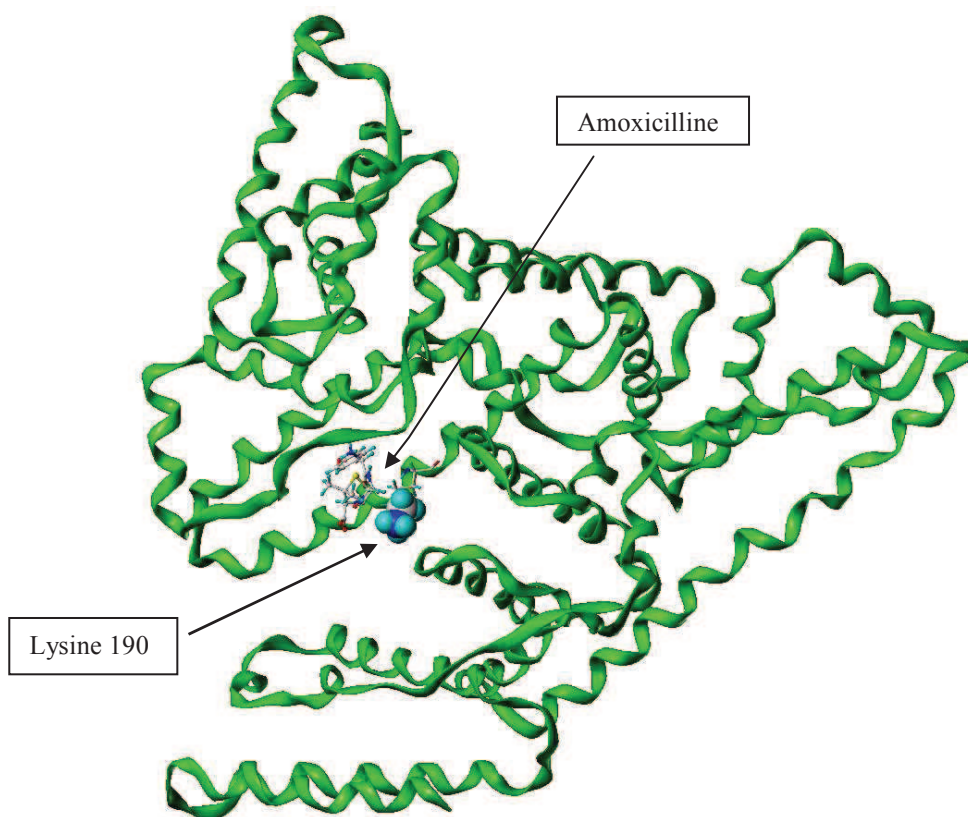


Figure 21: Modèle 3D de l'amoxicilline fixée sur la lysine 190 de l'albumine.

Nous ne présentons pas ici les analyses effectuées sur les autres lysines sélectionnées par notre étude de docking car elles sont basées sur la même approche méthodologique et de ce fait sont très similaires.

## II-4- Recherche d'autres protéines d'hapténisation : approche SuMo :

Partant de l'hypothèse que la bêta-lactamine doit se fixer de façon covalente à une protéine d'hapténisation, nous nous sommes intéressés à rechercher quelles protéines humaines étaient capables de prendre en charge le médicament, qui par la suite, serait reconnu soit par les IgE, soit par les LT. Pour cela, nous avons réalisé une étude au moyen de notre logiciel SuMo, dont les capacités de comparaison de structures tridimensionnelles des protéines correspondaient à nos besoins.

Nous savons que la cible naturelle des bêta-lactamines est la PFP. Le site d'interaction entre la bêta-lactamine et la PFP nous sert de point de départ et nous souhaitons retrouver l'ensemble des protéines humaines possédant un site similaire à celui de la PFP. Le logiciel SuMo est tout à fait adapté à ce genre de requête.

Nous avons donc pris comme référence le site de la Protéine Fixant la Pénicilline co-résolue avec la cefuroxime et dont le code PDB est 1QMF. Le site a été défini comme l'ensemble des acides aminés bordant le site de co-cristallisation à une distance minimum de 10 Å.

Nous avons créé une base de données contenant tous les modèles de protéines sériques humaines que nous avons pu obtenir. Pour cela, nous avons récupéré dans la base de données de séquence Uniprot, les 6020 séquences de protéines sériques humaines et les avons modélisé au moyen de notre logiciel Modeome. Il permet de modéliser par homologie de séquences (il s'appuie sur des structures résolues expérimentalement et déposées à la PDB), et de façon automatique, un très grand nombre de modèles (voir chapitre 2). Sur l'ensemble des séquences, nous n'avons pu modéliser que 1536 protéines

(les autres ne présentant pas une homologie de séquence suffisante). Ces 1536 protéines constituent notre “Serum DataBank”.

Nous demandons donc à SuMo de retrouver au sein de la Serum databank les sites les plus proches de notre référence afin d’obtenir une liste de protéines candidates pour prendre en charge le médicament.

Nous obtenons une liste de protéines susceptibles d’être hapténisées par le chimique, ainsi que le site de fixation de la bêta-lactamine sur chaque protéine. Le tableau 11 donne la liste des premières protéines retrouvées, classées selon le score SuMo.

La première colonne donne les identifiants des protéines c’est-à-dire le code PDB si la structure a été résolue expérimentalement ou le numéro du modèle créé grâce à Modeome si aucune structure n’a été déposée à la PDB. La colonne 2 donne le nombre d’objets communs avec la « query » (la référence), la colonne 3 donne le score SuMo et permet donc le classement des protéines retrouvées et la dernière colonne présente le nom des protéines.

<b>Identifiant PDB ou modéome</b>	<b>Nombre d’objets communs</b>	<b>Score SuMo</b>	<b>Nom des protéines</b>
851	5	3,5	Human Dimethylglycine dehydrogenase mitochondrial precursor
851	4	3	Human Dimethylglycine dehydrogenase mitochondrial precursor
2J2P	6	3	L-Ficolin complexed to N-Acetyl-Cystein
2J3G	6	3	Ficolin 2 precursor
1155	4	3	Human Serum albumin
2J2P	5	2	L-Ficolin complexed to N-Acetyl-Cystein
900	3	2	Human Multidrug resistance-associated protein 5
387	3	2	Human Complement C4-A precursor
215	3	2	Human Potassium-transporting ATPase alpha chain 2
170	3	2	Human AP-2 complex subunit alpha-2
854	3	2	Human Mitogen-activated protein kinase kinase kinase 12
388	3	2	Human Complement C4-B precursor
1403	3	2	Human Synaptotagmin-like protein 4
223	4	2	Human Potassium-transporting ATPase alpha chain 1
1097	3	2	Human Receptor-type tyrosine-protein phosphatase U precursor
215	3	2	Human Potassium-transporting ATPase alpha chain 2

170	3	2	Human AP-2 complex subunit alpha-2
1196	4	2	Human ATPase Na <sup>+</sup> /K <sup>+</sup> transporting alpha 4
223	4	2	Human Potassium-transporting ATPase alpha chain 1
854	3	2	Human Mitogen-activated protein kinase kinase kinase 12
608	3	2	Human Guanine nucleotide-binding protein alpha-13
1039	3	2	Human Polymeric immunoglobulin receptor precursor
1514	3	2	Human Vacuolar ATP synthase catalytic subunit A
917	3	2	Human Myosin-Ic
2J3U	5	2	L-Ficolin complexed to galactose
2J0Y	5	2	L-Ficolin complexed to B-13-D-Glucan

Tableau 11 : Résultats de la comparaison entre le site de fixation de la pénicilline de la PFP et les protéines de la Serum Databank

Notons que certaines protéines apparaissent plusieurs fois. Ceci est lié au logiciel qui peut proposer différentes superpositions pour une même entrée. Ainsi, les protéines que nous obtenons présentent toutes la particularité de posséder un site proche de celui de la PFP, et sont donc potentiellement candidates pour la fixation covalente de la bêta-lactamine. La figure 22 présente la superposition du site de référence (en gris) avec celui de la L-Ficolin (en violet). Les objets SuMo communs sont représentés par des boules colorées et superposées.

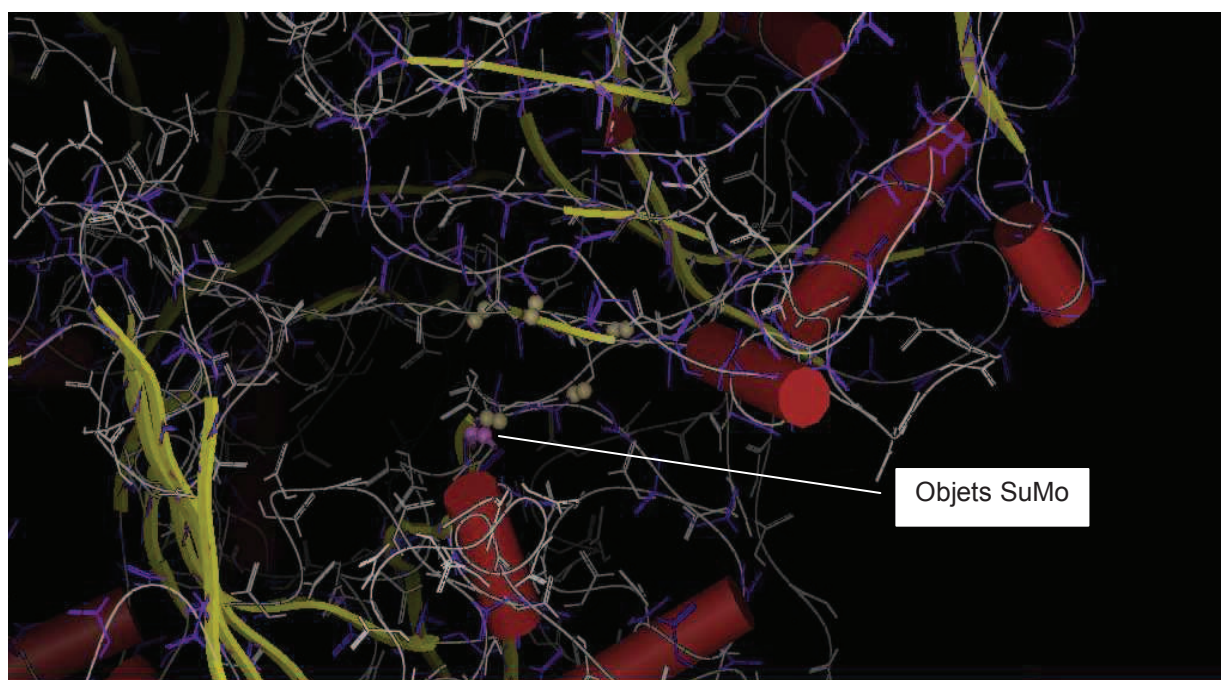


Figure 22 : Superposition du site de fixation des pénicillines de la PFP avec un site de la L-Ficolin

Ces résultats seront exploités dans le chapitre 6 qui détaillera également les séquences peptidiques contenant l'acide aminé catalytique (fixant covalamment le médicament) pour les protéines ici trouvées.

## II-5- Discussion :

Partant du postulat que les bêta-lactamines pour être immunogènes doivent se fixer à une protéine endogène, notre approche comparative entre la PFP et les protéines sériques que nous avons modélisées a montré plusieurs cibles potentielles d'hapténisation. Cependant, l'albumine étant une protéine présente en grande concentration dans l'organisme, qu'elle est en outre très souvent impliquée dans des phénomènes de transport et notamment celui des drogues, nous avons choisi de débiter avec cette protéine. D'autre part, nous nous sommes focalisés sur l'amoxicilline et la ceftriaxone qui représentent respectivement les pénicillines et les céphalosporines. Par des méthodes de Mécanique Quantique, nous avons étudié la réactivité des noyaux de type pénicilline et de type céphalosporine avec des modèles simplifiés de lysine et de sérine. Nous avons conclu que les pénicillines possédaient une réactivité supérieure à celle des céphalosporines, que ce soit pour une liaison sur une lysine ou sur une sérine. Ceci est parfaitement en accord avec les données cliniques que l'on peut trouver dans la littérature. En effet, si la réactivité des pénicillines est plus forte, cela signifie qu'elles réagiront plus sur les acides aminés que les céphalosporines, de ce fait auront une capacité accrue à se fixer de façon covalente sur l'albumine, ou une autre protéine, et donc pourront potentiellement déclencher plus d'accidents allergiques. De plus, nous avons constaté que les lysines sont plus réactives que les sérines ce qui est en accord avec leurs structures chimiques respectives.

Par la suite, nous avons cherché à valider notre hypothèse, basée sur le fait qu'une protéine ne va fixer, aux conditions physiologiques, que très peu de molécules médicamenteuses, voir une seule. En effet, cette liaison nécessite un site catalytique approprié qui ne sera pas présent sur toute la protéine. D'autre part, il semble logique de trouver un site plus particulièrement apte à lier l'antibiotique, car dans le cas contraire, c'est-à-dire si tous les acides aminés de toutes les protéines du patient pouvaient fixer de façon identique le médicament, celui-ci n'arriverait jamais sur sa cible naturelle... Nous avons postulé que l'albumine transportait les bêta-lactamines et avons recherché des sites

d'interactions par une approche de docking. Nous avons analysé les acides aminés de type lysine et sérine accessibles sur l'albumine, et donc susceptibles de créer une liaison covalente avec l'amoxicilline. Nos résultats montrent alors que cette fixation se fera préférentiellement sur les lysines 212, 190, 432 et 351 de l'albumine. Différentes données issues de la littérature sont en accord avec ces résultats *in silico*. L'équipe de Sugio (Sugio et al., 1999) a démontré que l'albumine humaine lie de façon covalente les acides gras via sa lysine 351. L'équipe de K.B. Park a également mentionné au cours d'une communication orale (Drug Hypersensitivity Meeting 4, Rome, avril 2010) que la liaison covalente en condition physiologique de la pénicilline sur l'albumine se faisait au niveau des lysines 190 et 212. D'autre part, l'équipe de J.M. Wal conclut que l'hapténisation se ferait sur les lysines 190, 195, 199, 432, 541 et 545. (Wal et al., 1975, Yvon M.P. et al., 1988, Yvon M.P. et al., 1989, Yvon M.P. et al., 1990). Les CScores (présentés dans le tableau 10) des lysines 195, 199, 541 et 545 sont relativement élevés (entre 4,32 et 6,02). Concernant le site 199 cependant, le CScore est extrêmement faible car la lysine est totalement enfouie et de ce fait inaccessible.

Notons également que nos résultats de docking montrent que les acides aminés de type lysine interagissent plus fortement que les sérines ou les histidines, ce qui est en accord avec la réactivité chimique de ces acides aminés.

Dans une troisième étape, et grâce aux méthodes mixtes de QM/MM, nous avons étudié la molécule antibiotique entière fixée de façon stable et covalente sur la lysine 190 du site d'interaction de l'albumine (ainsi que sur le résidu 212 ; résultats non présentés). Nous obtenons des résultats identiques à ceux de mécanique quantique pure, et donc sommes en accord avec les données de la littérature. En effet, nous voyons que l'amoxicilline qui est une pénicilline est plus réactive que la ceftriaxone qui est une céphalosporine de troisième génération. D'autre part, la lysine 190 est plus réactive que les sérines. Notons toutefois qu'il existe un biais lié à la fréquence de prescription des antibiotiques : l'amoxicilline étant largement plus employée que la ceftriaxone.

La validation du nombre de molécule(s) de bêta-lactamine(s) fixée(s) sur l'albumine ainsi que les acides aminés catalytiques cibles, passera par des expériences que nous détaillerons dans les chapitres 5 et 6 et qui consistent notamment à produire des bio-conjugués d'albumine couplée avec de l'amoxicilline radiomarquée. Rappelons que ces

expériences sont à réaliser au pH physiologique. En effet, nous ne discutons pas la possibilité de fixer beaucoup plus de ligands si l'on se place dans des conditions favorables (saturation de l'albumine avec un nombre maximum d'antibiotiques), ce que nous avons par ailleurs fait et présentons dans les applications des chapitres 5 et 6. Cette phase expérimentale devra également être complétée par l'identification des sites de fixation du médicament sur l'albumine.

D'autre part, notre hypothèse selon laquelle la liaison du médicament sur sa cible protéique ne peut se faire que si celle-ci présente un site avec une conformation particulière peut trouver une « contre théorie ». Elle énonce que l'hapténisation pourrait être due à une capacité accrue de certains acides aminés (particulièrement les lysines) à fixer le chimique. Cette capacité serait liée à une constante d'acidité de l'acide aminé considéré modifiée par l'environnement. Afin d'étudier cette théorie, nous avons prédit les pKa des acides aminés de type lysine, acide aspartique, acide glutamique et histidine de la structure 1AO6 de l'albumine. Nous ne présentons, dans le tableau 13, que les pKa des acides aminés suspectés de pouvoir fixer covalamment une molécule d'amoxicilline. Ces prédictions se sont appuyées sur le programme Yasara. Il effectue dans un premier temps une minimisation énergétique du modèle de l'albumine sérique humaine, puis rajoute des molécules d'eau et des ions sodium et chlorure autour de la protéine en les plaçant au point du potentiel électrostatique le plus fort afin d'atteindre une concentration de 0,9 % par rapport à l'eau. Une simulation de dynamique moléculaire est ensuite effectuée par le logiciel dans un ensemble périodique à température (300°K) et pression constante (1 atmosphère). Durant cette simulation, le programme effectue un calcul de pKa par acide aminé en fonction d'une formule empirique, présentée sur le figure 23, et basée sur le calcul du potentiel par la méthode de Poisson-Boltzmann, l'accessibilité du résidu et son nombre de liaisons hydrogène. Le programme se base également sur une série de 227 mesures de pH d'acides aminés dans 27 structures différentes. (Krieger et al., 2006).

$$pK_a = \text{Model } pK_a + \sum_{\text{Ionizable atoms}} [-A \times \text{Ewald}E_i + B \times \text{HB}_i] + \text{Sign}(\text{HBSum}) \times C \times \text{SurfaceLoss}$$

Figure 23 : Formule décrivant le calcul du pKa d'un acide aminé

La phase suivante consiste à comparer le pKa prédit par le logiciel au pKa de référence de l'acide aminé ce qui règle alors son état de protonation. Il faut noter que ce programme semble minimiser la variation de pKa, mais que d'après les auteurs, il est capable de reproduire les courbes de titration. Le tableau 12 présente dans la première colonne le type de résidu, en colonne deux la position du résidu dans la protéine, la colonne 3 donne le pKa prédit, la colonne 4 le pKa de référence et la dernière colonne correspond à la différence entre pKa prédit et pKa de référence.

Résidu	N° résidu	pKa(Prédit)	pKa(référence)	Différence
<b>LYS</b>	<b>190</b>	<b>9,95</b>	<b>10,76</b>	<b>0,81</b>
LYS	195	10,36	10,76	0,40
LYS	199	10,31	10,76	0,45
LYS	212	10,74	10,76	0,02
LYS	432	10,13	10,76	0,63
LYS	525	10,52	10,76	0,24
LYS	534	10,73	10,76	0,03
LYS	541	10,31	10,76	0,45
LYS	545	10,26	10,76	0,50

Tableau 12 : Résultat du calcul des pKa des acides aminés de type lysine de l'albumine sérique humaine

Nous avons donc analysé les pKa prédits de tous les acides aminés (type acide aspartique et glutamique, lysine, histidine) de l'albumine mais n'avons pas trouvé de différence significative de leurs valeurs, excepté pour le résidu lysine 190. La lysine 190 présente en effet une variation de la valeur de son pKa de 0,8. C'est la plus forte variation observée sur l'ensemble des données étudiées, et elle reste relativement faible. Cette différence peut trouver une explication dans le fait que la lysine 190 possède dans son environnement immédiat deux arginines, deux lysines et un acide glutamique. Au vu de ces résultats, nous ne pensons pas que l'environnement d'un résidu puisse, à lui seul, conditionner la réaction de fixation covalente. Le plus important, d'après nos calculs de docking et de QM/MM, est d'avoir une conformation spéciale et adaptée à la molécule pour favoriser sa fixation. Nous remarquons d'ailleurs que la lysine 190 est placée au fond d'une poche, ce qui peut en faire un acide aminé de choix pour la fixation de la molécule antibiotique. En effet, le site naturel retrouvé sur la PFP forme un sillon, et un site qui serait « protégé » comme celui de la lysine 190, avec de nombreuses interactions faibles potentielles (et donc une bonne énergie d'interaction), est plus adapté au phénomène



d'hapténisation. Cependant, il faut garder en tête que la lysine doit être sous forme non protonée, et donc qu'elle doit être dans un environnement suffisamment basique.

Le chapitre 6 discutera des différentes protéines que notre approche SuMo a retrouvées, et nous détaillerons les sites de liaison au médicament.

A ce stade du projet de thèse, deux objectifs présentés respectivement dans les chapitres 5 et 6 sont apparus. Concernant les HSA de type immédiat, nous utilisons les résultats présentés ci-dessus afin d'explorer au niveau moléculaire l'interaction entre les IgE et le médicament porté par une protéine. D'autre part, pour les HSA de type retardé, nous étudierons les différentes possibilités d'interactions entre le TCR et le médicament.

L'ensemble des modèles que nous avons générés, ainsi que les résultats obtenus doivent être confrontés à des résultats expérimentaux. L'analyse, la critique, la validation expérimentale ainsi que les applications potentielles seront présentées dans les deux derniers chapitres de ce travail de thèse.

# Chapitre 5 : Hypersensibilité allergique immédiate aux bêta-lactamines

Le chapitre 1 présente les caractéristiques générales de ce type d'hypersensibilité allergique. Ces réactions, de type 1 selon la classification de Gell et Coombs, font intervenir les immunoglobulines de type E, et sont caractérisées par un déséquilibre de la balance Th1/Th2 en faveur de la voie Th2 humorale. Lors de la phase de sensibilisation, sans symptôme, les IgE spécifiques du médicament vont être produites et vont se fixer via leur fragment constant Fc sur les mastocytes tissulaires et sur les basophiles du sang circulant (qui possèdent à leur surface des récepteurs à haute affinité pour le Fc). Lors d'un deuxième contact avec le même antigène, ce dernier va être spécifiquement reconnu par les IgE (pontage de deux IgE), ce qui va provoquer une dégranulation explosive des mastocytes et basophiles, à l'origine de la pathologie.

Dans ce chapitre, nous avons plusieurs objectifs :

- Valider la classification des bêta-lactamines par notre dendrogramme (chapitre 4).
- Etudier les mécanismes moléculaires d'interaction entre le médicament et les IgE.
- Appliquer et valider expérimentalement nos résultats grâce à une collaboration avec l'équipe du Pr. J. Bienvenu, qui vise à étudier la reconnaissance spécifique du médicament fixé à l'albumine par les IgE d'un patient souffrant d'HSA Immédiate à ce médicament.

## I- Validation du dendrogramme :

La première partie de ce travail de thèse, correspondant à la création du dendrogramme hiérarchique et à l'étude 3D-QSAR, s'est appuyée sur des données cliniques obtenues grâce à l'Unité Fonctionnelle Immunologie clinique et allergologie du Centre Hospitalo-Universitaire Lyon Sud.

Pour la construction du dendrogramme (figure 14, chapitre 4), nous avons modélisé, aligné puis classé 42 antibiotiques de type  $\beta$ -lactamine. Cette approche présente l'avantage de s'affranchir de la répartition classique connue, basée sur la nature du noyau accolé au cycle  $\beta$ -lactame, mais également de paramètres tels que la fréquence de prescription du médicament ou le pourcentage d'allergénicité. Les molécules sont analysées et classées en fonction de leurs champs moléculaires, permettant ainsi de prendre en compte l'ensemble de la molécule et non plus uniquement le motif central. Nous avons déterminé six familles (dont deux sont sous-représentées en terme de nombre) à partir desquelles nous étudions les réactivités croisées entre différentes molécules de  $\beta$ -lactamine. D'autre part, nous proposons une molécule de substitution en cas d'hypersensibilité allergique à une ou plusieurs  $\beta$ -lactamine(s). Il est extrêmement important de garder en mémoire que tous les calculs qui aboutissent au dendrogramme sont effectués sans aucune donnée clinique ou biologique.

Une fois le dendrogramme construit, et dans le but d'en valider la pertinence et les capacités de classification, nous avons utilisé les données de patients qui subissent des tests cutanés de type « prick test ». Ces tests, réalisés au CHU Lyon Sud, permettent de définir une ou plusieurs réaction(s) d'HSA de type Immédiat. Dix sept molécules sont ainsi testées. Les données cliniques ont été recueillies lors des « semaines HSI » sur une cohorte de 42 patients. Nous avons anonymé les données, et présentons les résultats dans le tableau 13. La première colonne correspond à la famille du dendrogramme à laquelle appartient le médicament testé, dont le nom apparaît dans la colonne 2. La dernière colonne présente les 42 patients numérotés de 1 à 42. Il faut noter que les familles 2 et 4 n'apparaissent pas car elles ne sont constituées que de une ou deux molécules respectivement, qui ne sont pas disponibles pour les tests cutanés. D'autre part, concernant la famille 6, seule une molécule a pu être testée.

famille N°	Nom de la Molecule	Numéro du Patient																									
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	Cefixime																										
1	Cefpodoxime																										
1	Ceftazidine																			X							
1	Ceftriaxone																										
1	Céfuroxime			X																							
3	Amoxicilline	X			X	X	X	X	X	X	X	X	X		X	X	X				X	X	X	X		X	X
3	Ampicilline	X				X					X				X	X	X				X	X		X		X	
3	Bacampicilline	X				X					X	X		X		X	X				X					X	
3	Céfaclor	X															X									X	
3	Céfalexine	X															X										
3	Céfatrizine	X			X	X								X			X	X							X		
5	Oxacilline		X							X	X						X				X						
5	Péni G	X				X			X	X	X			X	X		X				X					X	
5	Péni V									X				X							X						
5	Pipéracilline	X								X						X	X				X					X	
5	Ticarilline									X							X				X						
6	Céfazoline																X		X		X						

famille N°	Nom de la Molecule	Numéro du Patient															
		27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1	Cefixime																
1	Cefpodoxime																
1	Ceftazidine																
1	Ceftriaxone											X					
1	Céfuroxime						X				X						
3	Amoxicilline	X	X	X	X	X	X	X	X	X		X	X	X	X	X	
3	Ampicilline	X	X			X						X				X	
3	Bacampicilli	X	X			X							X	X		X	
3	Céfaclor																
3	Céfalexine	X				X											
3	Céfatrizine					X	X										
5	Oxacilline					X		X							X		
5	Péni G					X								X	X	X	
5	Péni V																
5	Pipéracilline					X		X				X		X			
5	Ticarilline	X	X									X					
6	Céfazoline	X															

Tableau 13: Tableau des données cliniques de 42 patients souffrant d'HSAl, sur une batterie de 17 molécules

Les croix correspondent à un test cutané positif. Les molécules sont classées en fonction de leur appartenance aux familles définies par le dendrogramme. Pour chacune des familles, nous avons calculé le pourcentage de molécules présentant un test cutané positif sur le nombre de molécules de la famille, et cela pour les 42 patients. La coloration des cases permet de représenter les pourcentages obtenus : Le blanc correspond à un pourcentage compris entre 0 et 4%, le gris clair entre 5 et 39%, le gris moyen entre 40 et 60%, le gris foncé entre 61 et 90% et en noir un résultat supérieur à 90%. Le tableau 14 résume ces résultats. En première colonne, la famille issue du dendrogramme et en colonne 2, les 42 patients.

Famille ▼	Numéro de patient																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
3	100	0	0	33	33	50	17	17	17	17	50	33	17	33	33	50	100	17	0	50	33	17	33	17	67	17
5	40	20	0	0	0	20	0	0	20	40	100	0	0	40	20	20	80	0	0	100	0	0	0	0	40	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	100	0	100	0	0	0	0	0

Famille ▼	Numéro de patient															
	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1	0	0	0	0	0	20	0	0	0	0	40	0	0	0	0	0
3	67	50	17	17	17	84	33	17	17	17	0	33	33	33	17	50
5	20	20	0	0	0	60	0	40	0	0	0	40	0	40	40	20
6	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Tableau 14 : Pourcentages de molécules positives par famille et par patient.

L'analyse du tableau ci-dessus nous confirme la pertinence du dendrogramme obtenu par modélisation moléculaire et plusieurs observations ressortent:

- **Il n'existe pas de test positif chez un même patient pour des molécules de la famille 1 et 3 ou de la famille 1 et 5.** Cette règle se vérifie pour 41 sur 42 patients, soit **97,6%** de fiabilité. (Seul le patient 32 échappe à cette règle). Autrement dit, un patient n'aura jamais un test positif à une molécule de la famille 3 s'il a un test positif à une molécule de la famille 1, et inversement. De même, si un patient a un test positif à une molécule de la famille 1, il n'aura aucun test positif pour des molécules de la famille 5, ni de la 3. Inversement, un patient qui a un test positif à une molécule de la famille 5 n'aura jamais un test positif pour une molécule de la famille 1 (mais pourra présenter un test positif à une molécule de la famille 3). Ainsi, si l'on considère que

les tests cutanés sont fiables et caractérisent une réaction d'HSA, on peut dire qu'aucun patient ne sera allergique à une molécule de la famille 1 si il est allergique à une molécule de la famille 3 ou 5. On peut également conclure qu'il n'existe pas de réaction d'HSA Immédiate croisée entre les molécules des familles 1 et 3, ni entre les molécules des familles 1 et 5. Ainsi, les familles du dendrogramme sont elles parfaitement corrélées avec les données patients. Ces observations laissent penser que notre approche est pertinente puisque les familles que nous déterminons par modélisation moléculaire ne croisent pas dans la majorité des cas (1 ne croise ni avec 3, ni avec 5). Le regroupement en six familles obtenu par modélisation moléculaire semble donc plus fidèle que la répartition classique pour laquelle nous ne trouvons pas de règle exclusive comme avec le dendrogramme.

- Nous avons poussé plus loin nos analyses et émettons la proposition suivante : **Si un patient présente un test positif à moins de 40% des molécules d'une même famille, alors, ce patient ne présentera de tests cutanés positifs qu'aux molécules de cette famille.** Par exemple, le patient 2 présente des tests positifs pour 20% des molécules de la famille 5, et on constate que ses tests sont négatifs pour les molécules des familles 1, 3 et 6. Par contre, le patient 6 a 50% de molécules de la famille 3 qui déclenchent une positivité des tests, et on voit qu'il a également des tests positifs pour des molécules de la famille 5 (pour 20% d'entre elles). Cette hypothèse se valide dans 40 cas sur 42 cas. En effet, les patients 9 et 15 présentent des résultats positifs à plus d'une famille sans dépasser le seuil de 40%.
- L'analyse du tableau révèle également que les patients qui ont réaction positive à la famille 6 auront des tests positifs pour les molécules d'au moins une des trois autres familles (1, 3 et 5).
- Nous remarquons une hétérogénéité dans la répartition du nombre de molécules par familles. En effet, la famille 1 est constituée de 5 molécules, la famille 3 présente 6 molécules, la famille 5 comprend cinq membres, la famille 6, un seul. Pour cette famille, les autres molécules n'étaient pas disponibles à l'hôpital.
- On remarque que 22 des 42 patients (soit 52,4%) présentent un pourcentage de molécules avec tests positifs inférieur à 40% pour chaque famille. Les 20 autres patients (soit 47,6%) ont une réponse positive à plus de 40% des molécules pour au moins une des quatre familles. Cela nous permet de dégager deux ensembles distincts de patients et ainsi de pouvoir proposer une stratégie de réintroduction du médicament sur des bases rationnelles.

- Concernant le premier groupe de patients (22 sur 42) qui présentent des pourcentages inférieurs à 40%, on remarque que seulement deux d'entre eux (patients 9 et 15) auront des tests positifs à plus d'une famille. Ces patients correspondent aux deux exceptions à la règle que nous avons suggérée selon laquelle, si un patient à un pourcentage inférieur au seuil de 40% pour une famille donnée, alors, il n'aura de tests positifs que pour des molécules de cette famille. Ces deux patients 9 et 15 montrent en effet des réactions croisées entre les molécules des familles 3 et 5. Notons toutefois que les tests cutanés sont toujours négatifs pour les molécules des familles 1 et 6, ce qui laisse des possibilités de traitement de substitution sans risque pour le patient. Pour les 20 patients sur 22 qui répondent à notre règle, 18 ont des tests positifs à des molécules de la famille 3 seulement, un patient (numéro 3) n'est positif qu'à des molécules de la famille 1 et le patient 2 qu'à des molécules de la famille 5. Les patients 2 et 3 sont donc particulièrement importants pour soutenir la pertinence du dendrogramme.
- Concernant les 20 patients sur 42 qui présentent une réponse positive supérieure à 40% dans au moins une famille, on remarque qu'ils sont tous positifs aux molécules de plus d'une famille (sauf le patient 37). D'autre part, une analyse plus approfondie révèle que pour tous ces patients, s'il n'y a pas de réponse positive à la molécule de la famille 6, les molécules de la famille 1 constituent un choix totalement sécurisé de substitution. En effet, pour ces patients, le pourcentage de molécules positives dans la famille 1 est toujours de zéro sauf pour le patient 19. Ce dernier est en effet positif à la ceftazidime de la famille 1 (céphalosporine de troisième génération) ainsi qu'à la cefazoline de la famille 6 (céphalosporine de première génération). On en déduit donc que tous les patients positifs aux molécules du groupe 6 risquent toujours d'avoir des tests positifs dans les autres familles, et il n'y a alors pas de choix sûr de substitution. Cette analyse est valable dans les quatre cas observés du tableau (patients 17,19, 21 et 27).

Ceci démontre clairement la fiabilité et la pertinence de notre répartition puisque notre règle présente un taux de validité de 39/42 patients soit **92,8%**.

A partir de ce dendrogramme, nous souhaitons proposer un programme de réintroduction lorsqu'un patient souffre d'HSA Immédiate à une ou plusieurs bêta-lactamines et l'analyse des données cliniques nous permet de dégager une stratégie rationnelle. Voici le principe :

- Si le patient à un prick-test **positif à la céfazoline** (famille 6), il est trop risqué de lui redonner une bêta-lactamine. En effet, la famille 6 croise avec les membres des familles 1, 3 et 5 (règle valable dans **100%** es cas). Dans cette situation, nous proposons de **changer de classe** médicamenteuse.
- Si le test de la molécule de **céfazoline** est **négatif** :
  - Si le patient est **positif** pour une ou plusieurs molécules de la **famille 1**, alors une molécule de la **famille 3 ou 5** devrait être une bonne option de substitution (règle valable dans **97,6%**)
  - Sachant que 19 patients présentent des tests positifs pour les familles 3 et 5, si le patient est **positif** à une ou plusieurs molécules de la **famille 3 et/ou 5**, il faudra opter pour des molécules de la **première famille** (règle valable dans **97,6%** des cas)

## II- Reconnaissance moléculaire des bêta-lactamines par une IgE :

Nous étudions les réactions d'hypersensibilité allergique de type immédiat, et de ce fait, comme il est décrit dans le chapitre 1, ce sont les immunoglobulines E qui vont être à l'origine de la reconnaissance spécifique du médicament, donc vont être responsables des évènements physiopathologiques caractéristiques des HSA Immédiates. Nous avons cherché à modéliser les interactions impliquées lors de la reconnaissance de l'épitope par le ou les paratopes.

La littérature décrit que la réaction d'HSA de type immédiat est liée à la reconnaissance de la bêta-lactamine par deux IgE qui, via le médicament, vont être pontées, et de ce fait déclencher la dégranulation des basophiles et mastocytes qui portent ces immunoglobulines. Nous nous sommes intéressés aux différentes possibilités de déclenchement des HSAI. En effet, une molécule de bêta-lactamine est petite, et il nous semblait difficile qu'une seule molécule médicamenteuse puisse être reconnue par deux paratopes d'IgE. D'autre part, nos études décrites dans les chapitres 4 et 6, se sont orientées sur les protéines responsables du transport des drogues. Nous parlons de protéines d'hapténisation qui fixent de manière covalente la molécule de bêta-lactamine au niveau du carbone du carbonyle (cycle bêta-lactame), et principalement via des lysines et des sérines.



Nous nous sommes donc interrogés sur la présence ou non de la protéine lors de la reconnaissance par les IgE. En effet, la protéine hapténisée par la drogue offrirait une plus grande « surface » de présentation de l'allergène aux IgE.

Afin de mieux comprendre l'origine de la dégranulation des mastocytes et basophiles, nous souhaitons donc générer plusieurs modèles d'interaction mettant en jeu, le médicament, l'IgE, avec ou non la protéine ou le peptide d'hapténisation. Voici les différentes options étudiées :

- Le médicament, seul, reconnu par une ou deux IgE (scénario que nous expliquerons dans les perspectives de ce chapitre)
- Le médicament fixé sur une protéine d'hapténisation reconnu par une ou deux IgE
- Le médicament fixé sur un peptide d'hapténisation reconnu par une ou deux IgE (voir en fin de ce chapitre, dans les perspectives).

D'autre part, il est tout à fait possible que les IgE soient pontées suite à la reconnaissance des deux molécules antibiotiques. Celles-ci seraient soit seules, soit fixées sur une seule protéine, soit fixées sur deux protéines, soit fixées sur un peptide ou deux peptides distincts.

Nous allons présentés les différents acteurs impliqués :

## II- 1- Modèle d'IgE :

Le chapitre 1 présente la structure d'une IgE. Grâce à son Fab (Fragment Antigen Binding), partie hypervariable, l'anticorps pourra reconnaître spécifiquement l'antigénique.

Comme pour l'ensemble des travaux de modélisation moléculaire, nous devons partir d'un modèle résolu expérimentalement, ou modélisé par homologie de séquences. Nous avons donc cherché au sein de la base de données protéique : PDB un modèle d'IgE qui nous servirait à analyser finement les interactions moléculaires entre l'anticorps et la drogue, ici, l'amoxicilline. Malheureusement, à l'heure actuelle, aucune structure d'IgE spécifique de l'amoxicilline n'a été résolue. La PDB recense 980 structures de Fab spécifiques de différents ligands, mais aucun Fab spécifique de l'amoxicilline, ni du déterminant majeur de la pénicilline (BPO).

Cette absence de modèle expérimental constitue un obstacle majeur à notre étude et nous avons donc dû nous contenter de mener nos calculs avec des structures non spécifiques de l'amoxicilline. Nous avons constaté que les Fab présentaient une surface de reconnaissance plane capable d'interagir avec le médicament lui-même fixé sur une protéine porteuse.

## II-2- Modèle du médicament :

Cette partie du travail a été réalisée et présentée dans le chapitre 4. Nous nous sommes intéressés ici à la molécule d'amoxicilline, puisqu'elle est la plus prescrite, mais la suite de ce travail de thèse analysera les autres molécules de la famille des bêta-lactamines, et notamment les céphalosporines de troisième génération qui semblent engendrer moins de réactions d'hypersensibilité allergique.

## II-3 Protéines d'hapténisation :

Comme nous l'avons dit, l'albumine sérique humaine étant très présente dans le sérum humain (elle représente habituellement 60% des protéines plasmatiques) et qu'elle est connue pour transporter des drogues, nous avons décidé, dans un premier temps, d'étudier l'hapténisation de l'albumine par l'amoxicilline. Le chapitre précédent présente la structure de cette protéine qui a été résolue expérimentalement par Sugio et déposée à la PDB sous le code 1A06.

Nous cherchons donc à savoir si les IgE reconnaissent l'amoxicilline fixée sur l'albumine. Les interactions avec les anticorps sont-elles limitées au médicament ? Quelle est l'importance et le rôle de la protéine hapténisée?

Notre approche SuMo a permis de proposer d'autres protéines susceptibles de fixer de manière covalente une molécule de bêta-lactamine. La suite de ce travail reprendra donc les analyses menées pour l'albumine humaine mais avec d'autres protéines hapténisées.

## III- Analyse de l'albumine hapténisée :

### III-1- Les bioconjugués :

#### III-1-1 Synthèse :

Les bioconjugués correspondent à des protéines d'albumine sur lesquelles plusieurs molécules médicamenteuses sont fixées. Les manipulations expérimentales de fixation de l'amoxicilline sur l'albumine sérique humaine ont été réalisées par le Dr. C. Marminon du laboratoire de Chimie Organique de la faculté de Pharmacie de Lyon. Le protocole suivi correspond à celui que J.M. Wal décrit dans une publication de 1975 (Wal et al., 1975). Pour cette phase du travail, nous nous sommes placés à un pH non physiologique afin de fixer un maximum de molécules d'amoxicilline. Ainsi, la synthèse a été réalisée en milieu basique à un pH de 10,4 en présence d'un tampon carbonate. Par la suite, l'albumine est mise en solution avec de l'amoxicilline selon un rapport de 1 pour 500. Le milieu basique déplace les équilibres de pH des chaînes latérales des lysines et ces dernières présentent de ce fait des fonctions  $\text{NH}_2$  à l'origine des attaques nucléophiles et donc des ouvertures des cycles bêta-lactame. La protéine hapténisée est alors dialysée durant 5 jours afin de libérer le milieu de l'excédent d'amoxicilline et de tampon.

Il est également possible de réaliser ces bioconjugués au pH physiologique afin d'étudier la réalité biologique du phénomène d'hapténisation. Ces expériences seront présentées dans les perspectives de ce chapitre ainsi que dans la seconde partie du chapitre 6. Nous observons que si l'on respecte le pH du milieu sanguin, une seule (voire deux) molécule sera fixée.

Nous sommes intéressés au premier procédé dans le but de produire une protéine pouvant servir dans les tests de diagnostic des hypersensibilités allergiques immédiates à l'amoxicilline. Nous appelons « albumine sale » la protéine hapténisée par environ 10 amoxicillines et « albumine propre » celle qui n'a fixé qu'une seule (ou 2) molécule d'amoxicilline. La validation expérimentale de la fixation des drogues a été réalisée par le laboratoire de spectrométrie de masse de l'IBCP grâce à la technique de l'électrospray. Ceci nous permet de déterminer le nombre d'amoxicilline fixé par protéine (10 en moyenne) en

faisant la différence avec la masse de la albumine seule (sans ligand). Les résultats arriveront très prochainement.

### III-1-2 Analyse par modélisation moléculaire :

Nous sommes partis du modèle de l'albumine sérique humaine déposée à la PDB sous le code 1AO6 et avons créé un modèle de l'albumine sur laquelle nous avons fixé de manière covalente 10 molécules d'amoxicilline au niveau des acides aminés suivants : K190, K195, K199, K212, K432, K541, K524, K534, K545 et S193. La modélisation de la fixation des antibiotiques nécessite tout d'abord l'ouverture du cycle bêta-lactame puis la création de la liaison covalente avec l'azote de la lysine cible (on réalise alors une liaison de type amide), ou avec l'oxygène de la sérine (liaison ester). Afin de positionner au mieux l'amoxicilline face au résidu de fixation, nous nous sommes appuyés sur nos résultats de docking qui, comme nous le mentionnons dans le chapitre précédent, permettent de déterminer la meilleure pose du médicament dans le site catalytique. Il faut remarquer que pour la position K199, nous avons dû « forcer » la fixation car l'acide aminé 199 est très enfoui au sein de la protéine. A l'heure actuelle, nous n'avons malheureusement pas été en mesure de valider les sites de fixation de manière expérimentale. Une fois la liaison amide ou ester créée, nous avons contrôlé que la molécule était dans une conformation correcte et n'engendrait pas de gêne stérique vis-à-vis de la protéine. La structure complète a subi une minimisation énergétique, en utilisant le champ de forces Tripos et les charges partielles « Gasteiger-Marsilli ». La figure 24 représente le modèle de l'albumine (en ruban) hapténisée par les molécules d'amoxicilline (en boules).

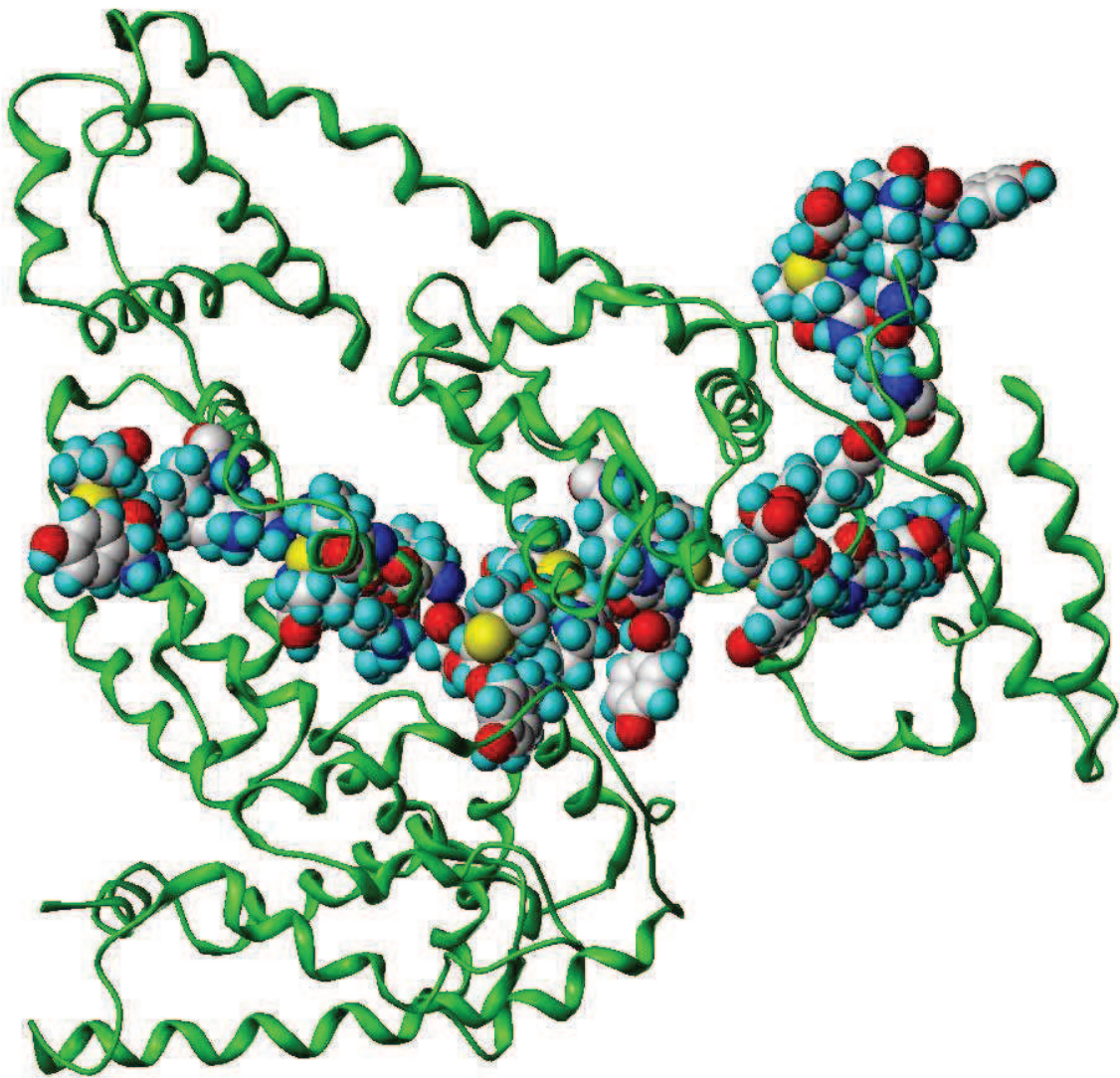


Figure 24 : Albumine sérique humaine hapténisée par 10 molécules d'amoxicilline

On remarque que la fixation des molécules ne présente pas une répartition homogène sur la surface de la protéine. Ceci valide notre première hypothèse selon laquelle la réaction chimique de fixation de la molécule ne se fait pas n'importe où, sur tous les acides aminés accessibles en surface d'une protéine, mais qu'un environnement particulier est requis. Dans les conditions les plus extrêmes, qui favorisent donc la liaison covalente et permettent de saturer l'albumine, seulement une dizaine de molécules sont fixées. Au pH physiologique, une seule molécule sera fixée. Le fait de se placer dans des conditions favorables à la fixation de l'antibiotique, c'est-à-dire à pH basique (au moyen d'un tampon carbonate) n'entraînera pas de réactions sur de « nouveaux » sites. Sur un plan thermodynamique, chaque réaction de fixation admet une variation d'énergie libre ( $\Delta G_r$ ), et le fait de se placer dans des conditions favorables permet de déplacer les équilibres et d'atteindre plus de sites. De ce fait, pour les

sites les plus susceptibles de réagir, la réaction sera quasi-totale. Ceci explique notre choix de travailler avec de l'« albumine sale » qui nous garantit que le site préférentiel (en conditions physiologiques) aura réagi de façon totale.

Nous avons choisi d'analyser de façon plus précise les positions 190 et 212, qui sont celles que nous avons trouvées grâce à notre approche de docking, et qui sont également décrites dans la littérature. Notons que les résultats que nous apportons ci-après sont issus de calculs réalisés pour les 10 sites d'hapténisation mais que nous ne présentons que ceux qui nous ont semblé les plus aptes à fixer le médicament en condition physiologique.

Concernant la lysine 190, elle est placée au milieu du canal central de l'albumine, c'est-à-dire enfouie et protégée. Cette lysine est placée dans un site qui forme une poche. Il est donc compréhensible que ce site soit plus favorable à l'hapténisation puisque dans ce cas, la molécule d'amoxicilline est stabilisée, permettant ainsi l'initiation de la réaction chimique. Cependant, l'analyse des surfaces d'accès au solvant de l'albumine libre (sans amoxicilline) et de l'albumine qui a fixé une molécule d'amoxicilline en position 190 orientent vers une conclusion opposée. En effet, il est très difficile de savoir où se place la molécule. La figure 25 représente la surface de l'albumine libre en bleu et celle de l'albumine hapténisée en vert. La différence en terme d'interactions hydrophobes de type Van der Waals n'est pas significative, et donc ne peut expliquer une reconnaissance de la drogue par les anticorps. La molécule d'amoxicilline est complètement enfouie, donc non accessible.

HSA  
HSA Amox K190

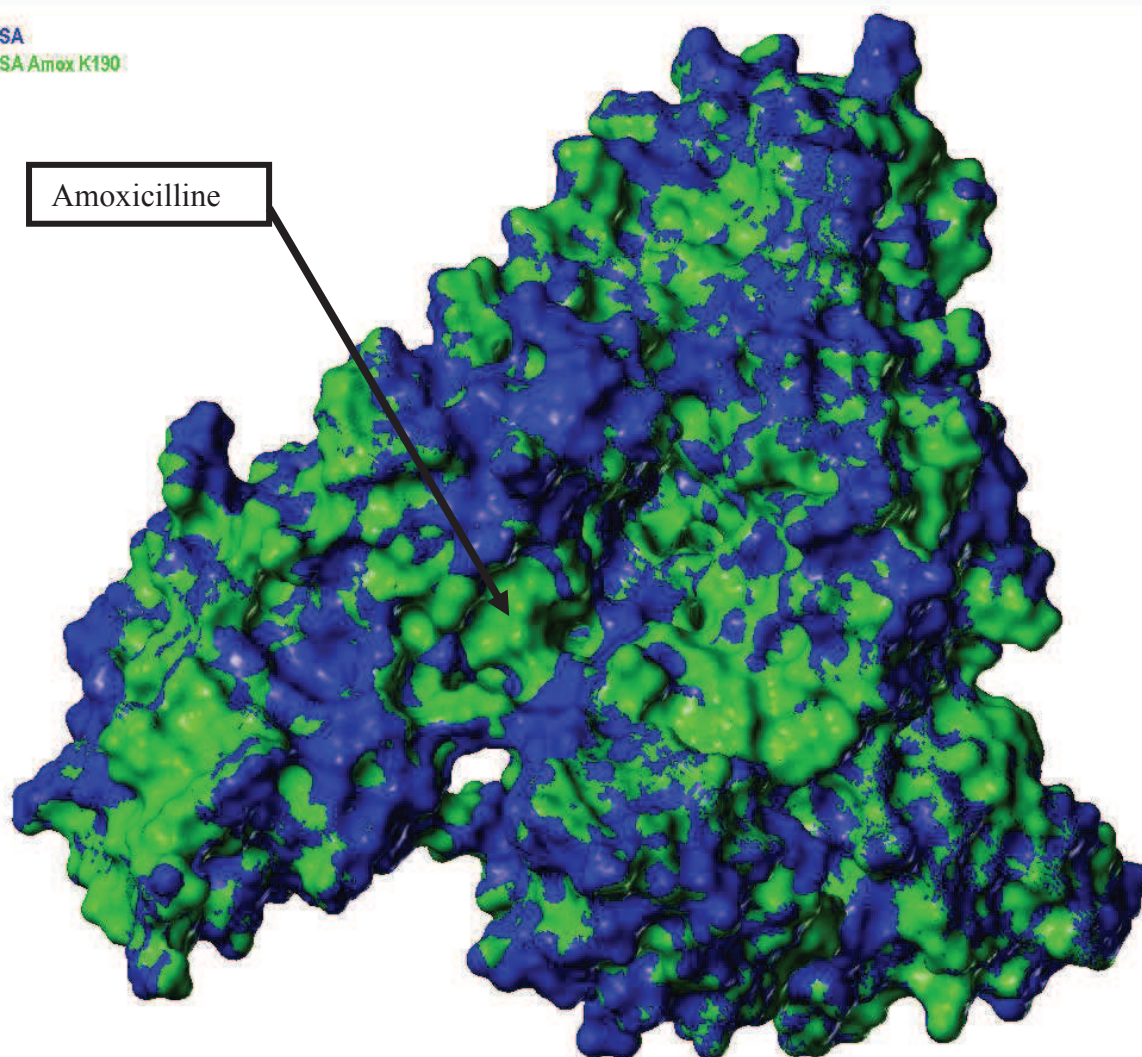
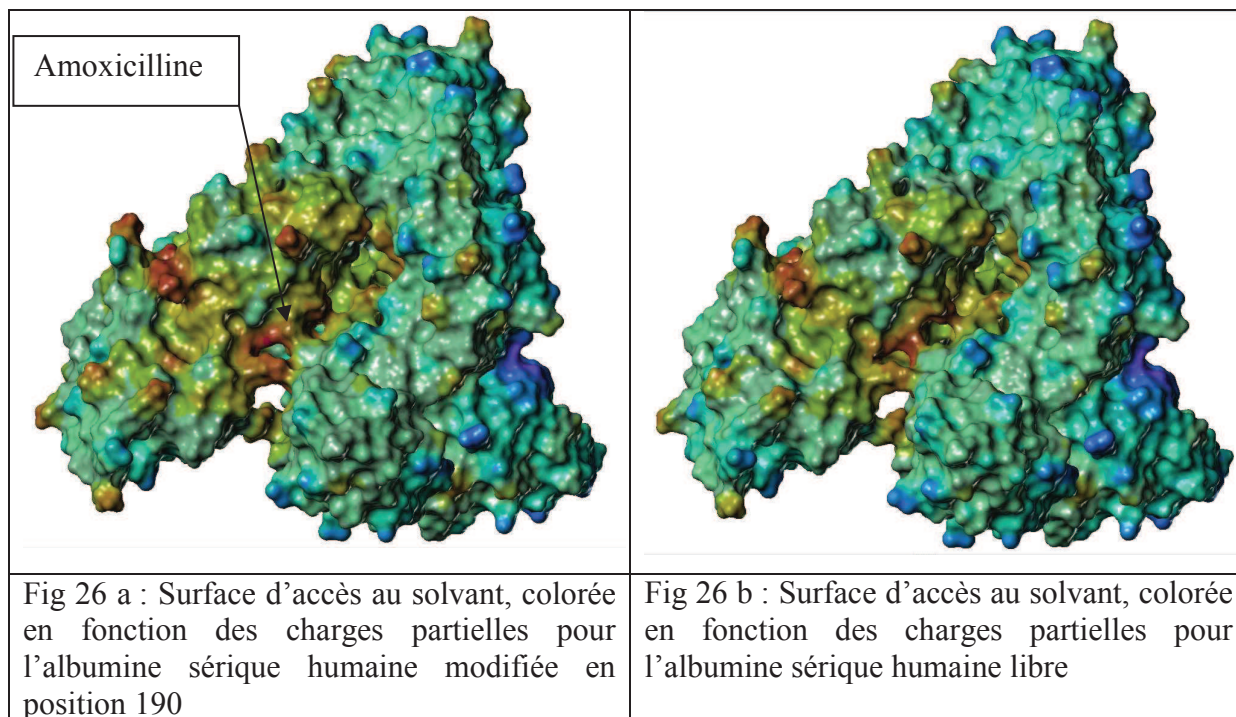


Figure 25 : Surface d'accès au solvant de l'albumine sérique humaine libre et de l'albumine sérique humaine hapténisée en position 190

Nous avons donc, par la suite, évalué en terme électrostatique les modifications liées à la molécule d'amoxicilline au niveau de la zone de fixation et qui auraient pu être un facteur déterminant lors de la reconnaissance spécifique par l'IgE. A partir des mêmes calculs que pour la figure 25, nous avons coloré la surface en fonction de la charge partielle des atomes en surface de la protéine et en contact avec le solvant. Les charges partielles sont calculées par l'algorithme Gasteiger-Marsilli avec une constante diélectrique de 80, qui permet de mimer l'effet de l'eau. Les figures 26 a et 26 b montrent les surfaces colorées en fonction des charges partielles de l'albumine modifiée en position 190 et de l'albumine libre.



Nous remarquons, là encore, que les différences sont extrêmement faibles. Elles sont dues à la fonction hydroxyle du groupement phénol de l'amoxicilline (qui pointe le plus vers l'extérieur de la protéine). Il sera donc difficile pour une IgE d'interagir spécifiquement avec un aussi petit motif de reconnaissance.

L'analyse de la position 212 est, elle, par contre, totalement différente. La lysine 212 est située dans une poche peu profonde (ce qui justifie qu'elle possède le plus fort CScore en docking) et sa chaîne latérale reste très accessible. Ainsi, les surfaces des deux modèles de l'albumine libre (en bleu) ou hapténisée en position 212 (en vert), présentées sur la figure 27, montrent clairement que l'amoxicilline est très exposée, et donc accessible aux IgE. S'agissant des charges partielles, nous constatons également que l'antibiotique est parfaitement exposé. En effet, la fonction carbonyle est accessible au solvant et constitue une région chargée négativement alors que sur l'albumine libre, la zone a une charge positive liée du fait de la fonction ammonium de la lysine 212. Les figures 28 a et b présentent les résultats concernant les surfaces d'accès au solvant colorées en fonction des charges partielles pour l'albumine hapténisée (28 a) et libre (28 b)



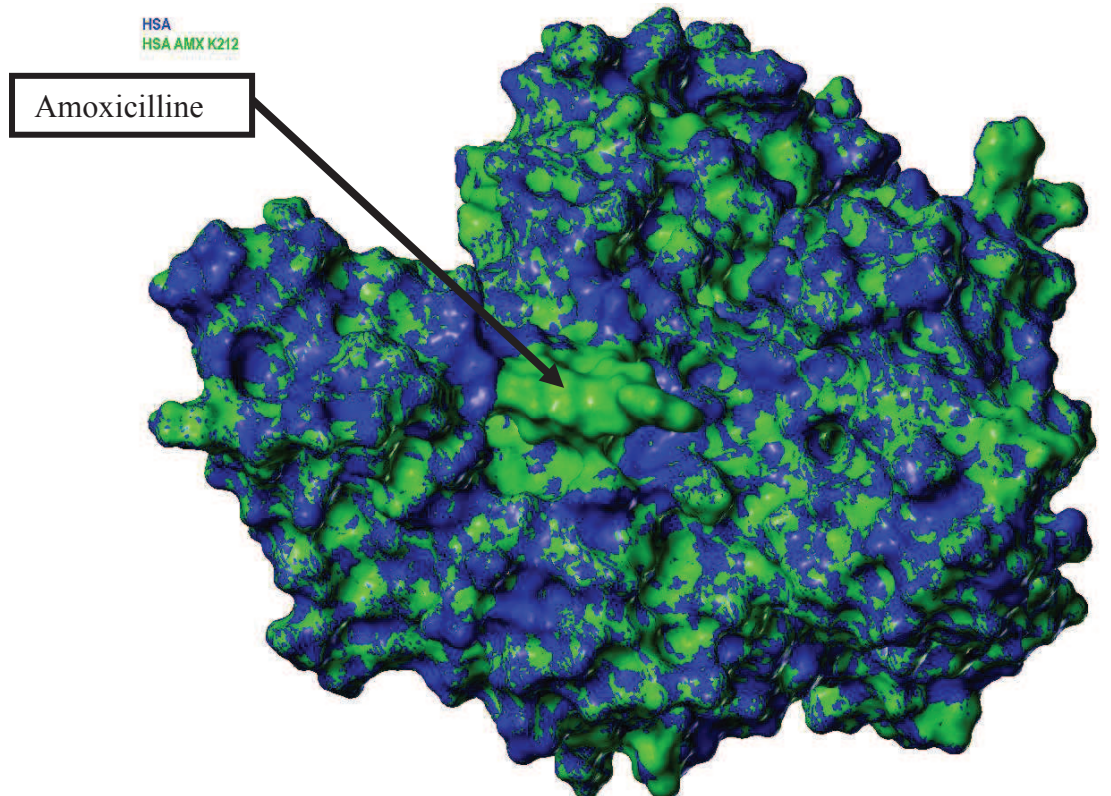
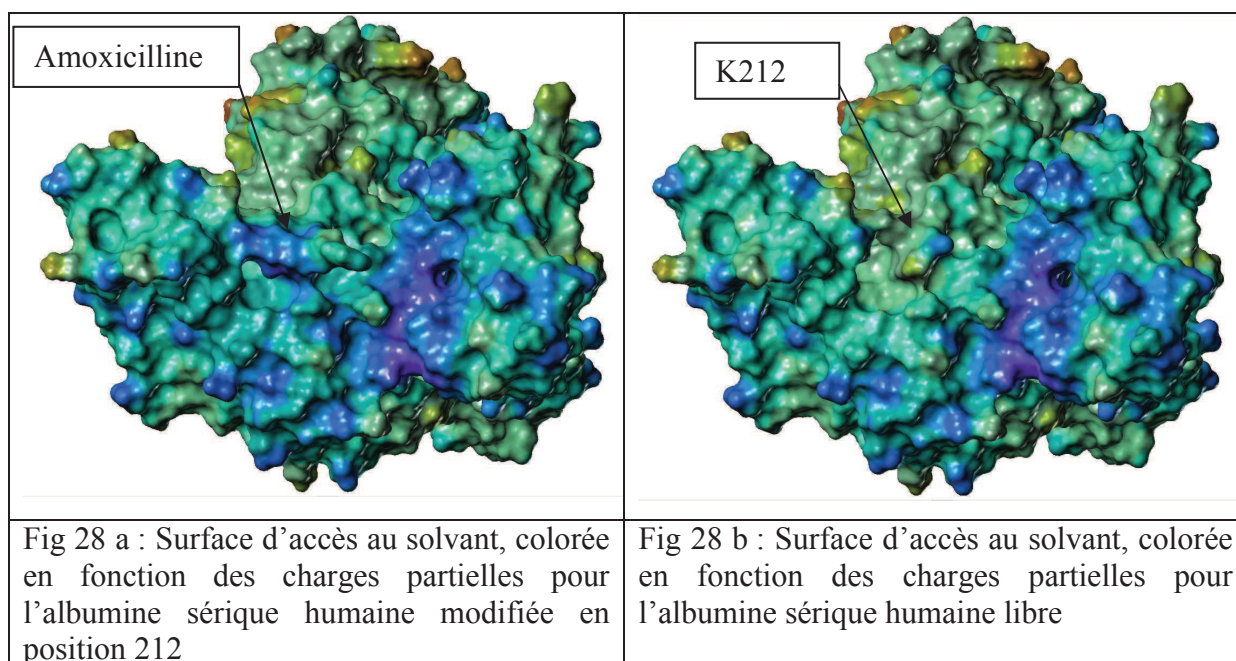


Figure 27 : Surface d'accès au solvant en terme hydrophobe de l'albumine sérique humaine libre et de l'albumine sérique humaine hapténisée en position 212



Ainsi, nous pensons que le site de la lysine 212 est le plus susceptible de fixer l'amoxicilline pour la présenter aux immunoglobulines E spécifiques.

### III-1-3- Interaction IgE-médicament :

En l'absence de données structurales sur le fragment Fab d'une IgE spécifique de l'amoxicilline, il nous est impossible de construire un complexe fidèle mettant en jeu le Fab, l'antibiotique, et la protéine hapténisée. Toutefois, nous avons analysé un certain nombre de structures de complexes IgE / Antigène issues de la PDB. Nous remarquons que, pour tous ces complexes, la reconnaissance de l'épitope se fera de manière relativement plane. Nous avons mesuré que le Fab interagissait avec des surfaces d'environ  $700 \text{ \AA}^2$  (surfaces elliptiques dont le grand rayon mesure une trentaine d'angströms)

Partant de ces observations, nous ne pensons pas que le Fab puisse aller reconnaître une molécule dans le sillon de l'albumine (site lysine 190). Par contre, le site de la lysine 212 présente une surface globalement plane bordée par des acides aminés chargés qui permettraient de créer de nombreuses liaisons hydrogène et donc de stabiliser un complexe anticorps/antigène.

D'autre part, sachant que le Fab reconnaît une surface plane de  $700 \text{ \AA}^2$ , il paraît impossible que la même molécule d'amoxicilline (qui elle, présente une surface de moins de  $100 \text{ \AA}^2$ ) soit reconnue par deux Fab simultanément. Nous pensons que un seul Fab va reconnaître spécifiquement une molécule d'amoxicilline, elle-même portée par une protéine telle que l'albumine.

### III-2- Applications :

Dans le premier chapitre, nous évoquons différents types de tests diagnostiques des réactions d'HSA Immédiates. Parmi ceux-ci, on trouve le TAB : Test d'Activation des Basophiles qui s'appuie sur le fait que les basophiles possèdent à leur surface des récepteurs de forte affinité pour le fragment constant des IgE. Ainsi, la cellule est recouverte d'anticorps spécifiques du médicament impliqué dans la pathologie allergique, et lorsque ce même médicament pénètre dans l'organisme, il sera reconnu spécifiquement par les IgE en surface des basophiles circulants ce qui provoquera leur dégranulation. Différents marqueurs permettent de caractériser l'activation des basophiles. Il est entre autre possible de mesurer la libération de médiateurs (histamine, IL 4, IL 13 notamment), ou l'expression de marqueurs tels que le CD63 ou le CD203c. Le principe du TAB est donc de reproduire *in vitro* les

conditions ayant conduit aux phénomènes allergiques cliniques observés chez un patient exposé à un allergène. Pour cela il faut incuber des basophiles sensibilisés à l'allergène en présence de ce dernier. Malheureusement, la mise au point des ces TAB est difficile et même s'ils peuvent avoir une excellente spécificité, leur sensibilité reste trop faible. Aussi, dans le but d'améliorer la sensibilité de ces tests, nous avons utilisé les bioconjugués albumine-amoxicilline. L'idée est d'optimiser la reconnaissance de l'amoxicilline par les IgE du patient et donc d'augmenter la sensibilité des tests.

Les bioconjugués synthétisés par le Dr. C. Marminon vont donc servir à activer les basophiles de patients souffrant d'HSA Immédiate à l'amoxicilline. Cette partie du travail est en cours de réalisation et effectuée par l'équipe du Pr. J. Bienvenu au laboratoire d'Immunologie du CHU Lyon Sud.

## **IV- Perspectives :**

Nos études par modélisation sont bien avancées et attendent d'être validées expérimentalement et appliquées.

Ainsi, nous avons lancé des expériences de test d'activation des basophiles avec de la « albumine sale ». Nous pourrions effectuer ces manipulations avec des peptides issus de la dégradation de cette albumine hapténisée afin d'étudier la nécessité ou non d'une protéine entière lors de la reconnaissance spécifique par les IgE.

Nous nous sommes également demandé si la protéine hapténisée par le médicament subissait une dégradation enzymatique avant la reconnaissance spécifique par les anticorps. Autrement dit, est-il possible que les IgE interagissent directement avec l'antigène fixé covalamment sur un petit peptide issu de la digestion de la protéine porteuse ? Afin de simuler une telle situation il faudrait générer des peptides à partir de l'albumine hapténisée par l'amoxicilline. Nous pourrions repérer les lysines suspectées de fixer covalamment la drogue, et sélectionner autour du résidu cible une vingtaine d'acides aminés. Cela constituerait un peptide hapténisé que nous pourrions mettre en interaction avec des IgE.

D'autre part, nous souhaiterions explorer l'hypothèse selon laquelle les IgE interagissent directement avec l'amoxicilline, sans qu'elle soit portée par une protéine. Pour cela, il faudrait bloquer l'ouverture du cycle bêta-lactame afin que le médicament ne puisse se

lier à une protéine, et voir si cette molécule seule peut être reconnue par l'IgE. Toutefois, les travaux que nous avons menés jusqu'à présent laissent penser que ce scénario n'est pas plausible puisqu'une taille minimum semble requise pour être reconnue par le Fab.

Concernant le nombre et la position des molécules médicamenteuses, une validation par spectrométrie de masse serait nécessaire. L'idée serait de synthétiser des bioconjugués d'albumine-amoxicilline et de les analyser, après digestion par des protéases, par spectrométrie de masse afin d'identifier l'acide aminé qui a fixé le médicament. Si nous travaillons en augmentant la concentration en amoxicilline, nous pourrions connaître les sites préférentiellement hapténisés. Ce type d'expérience serait extrêmement intéressant pour valider les sites de fixation que notre étude par modélisation moléculaire a trouvés.

L'ensemble des résultats que nous présentons ici pourra être appliqué à d'autres protéines d'hapténisation ainsi qu'à d'autres molécules de type bêta-lactamine, voire d'autres classes thérapeutiques.

Enfin, rappelons que cette phase du travail sera complétée et validée lorsqu'une IgE spécifique de l'amoxicilline aura été résolue expérimentalement par cristallographie de rayons X ou par RMN.

# Chapitre 6 : Hypersensibilité allergique retardée aux bêta-lactamines

Rappelons que, dans ces réactions de type IV selon la classification de Gell et Coombs, l'emballement du système immunitaire face à l'antigène (ici, la bêta-lactamine), va se faire au profit de la voie cellulaire dite Th1 (ce sont les lymphocytes CD4 helper de type 1 qui produiront des cytokines orientant le système vers une voie cellulaire) et conduira notamment à la production incontrôlée de lymphocytes T CD8 cytotoxiques spécifiques du médicament.

Au cours du chapitre 1 (paragraphe IV-2-), nous avons décrit les éléments principaux de ces hypersensibilités retardées. Lorsque le médicament pénètre dans l'organisme, il sera pris en charge par une CPA, qui le présentera, au sein de la niche à peptide du CMH aux LTCD4 puis, lorsqu'ils auront été générés, aux LTCD8 cytotoxiques effecteurs. Contrairement à la reconnaissance de l'antigène par les IgE, l'interaction entre l'antigène et le récepteur des lymphocytes T n'est jamais directe et nécessite que cet antigène soit présenté par une molécule de CMH de classe 1 (pour les LTCD8) ou 2 (pour les LTCD4).

Nous nous sommes intéressés au complexe molécule de CMH / peptide / bêta-lactamine / TCR, et avons cherché à analyser et comprendre les interactions moléculaires conduisant à l'activation spécifique des lymphocytes en réponse à la présence antigénique.

Le paragraphe III-2-1 du premier chapitre a également introduit les notions d'haptène, pro-haptène, pré-haptène et le concept défini par le Pr. W. Pichler sous le nom de P-i Concept. Les bêta-lactamines ayant un faible poids moléculaire, pour qu'elles soient immunogènes et donc potentiellement responsables de réactions d'allergies retardées, elles doivent subir un processus d'hapténisation. Il correspond à la fixation du médicament sur un acide aminé, et nous avons développé dans le chapitre 4 la réactivité des bêta-lactamines ainsi que les cibles protéiques potentielles pour l'hapténisation. Une fois que le médicament est fixé (sur une lysine ou une sérine notamment), la protéine sera captée par la CPA, et le processus

d'apprêtement conduira à la présentation du peptide (issu de la dégradation de la protéine hapténisée) sur lequel le médicament est toujours fixé (liaison covalente). Par la suite, un lymphocyte T spécifique interagira avec la CPA et reconnaîtra spécifiquement la bêta-lactamine fixée au peptide présenté dans la niche à peptide du CMH. Ce phénomène d'hapténisation - apprêtement, largement décrit dans la littérature a été nuancé par W. Pichler, qui a étudié la capacité de certains antigènes (dont des médicaments) à être reconnu par des lymphocytes T sans que la phase d'apprêtement soit nécessaire. Dans ce cas, le médicament est présenté dans la niche à peptide du CMH au LT, mais il n'y a pas de liaison covalente entre le peptide et le médicament, et la reconnaissance par le TCR serait semblable à une activation pharmacologique classique de récepteurs.

Dans ce dernier chapitre, nous allons présenter les différents scénarios envisagés pour décrire l'interaction entre les quatre acteurs que sont la molécule de CMH, le peptide de la niche, le médicament et le TCR.

Nous avons également mis en place une collaboration avec l'équipe du Pr. Marc Pallardy, qui vise à étudier les peptides présentés aux LT dans le cas d'hypersensibilité allergique retardée aux pénicillines. Nous présenterons les objectifs et les résultats de cette étude dans le paragraphe II.

## **I- Différents acteurs pour différents scénarios :**

Nous avons étudié, grâce à plusieurs outils de modélisation, les interactions moléculaires entre la molécule de CMH, le peptide qui a fixé de manière covalente ou non la bêta-lactamine et le TCR.

Pour cela, dans un premier temps, nous avons collecté les modèles de ces différents éléments. Celui du médicament a été créé et optimisé par mécanique moléculaire (chapitre 4). Pour le CMH, et le TCR, nous avons recherché un modèle expérimental dans la PDB. Nous avons pu trouver 158 structures de CMH de classe 1 et 2 ainsi que 44 structures de TCR et avons choisi le complexe déposé à la PDB sous le code 1FYT correspondant à un TCR co-cristallisé avec une molécule de CMH de classe 2 ainsi qu'un peptide antigénique. Les

modèles seront détaillés dans le paragraphe I-1. A partir de cette structure, nous avons étudié les interactions possibles entre CMH, médicament et TCR.

Présentons maintenant les différentes possibilités de reconnaissance que nous avons envisagé :

1. Mode classique après **hapténisation**- apprêtement : dans ce cas, la bêta-lactamine va être prise en charge par une protéine et se fixera de manière covalente, après ouverture du noyau bêta-lactame, sur un acide aminé cible. Nous avons présenté dans le chapitre 4 la réaction de fixation ainsi que les sites potentiels et la collaboration que nous détaillerons plus loin validera de manière expérimentale nos hypothèses. Ce scénario va donc faire intervenir un super complexe entre le CMH, le peptide hapténisé (placé dans la niche à peptides) et le TCR.
2. **P-i concept** : Ici, nous analyserons les interactions non covalentes entre le CMH, le peptide de la niche, le médicament et le TCR.
3. **New P-i concept** : Au cours de ce travail de thèse, une autre hypothèse nous a été suggérée. Elle part du principe que les interactions faibles décrites dans le modèle du P-i Concept ne sont pas suffisantes pour maintenir le complexe CMH-TCR. Nous détaillerons ce que nous appelons le « New P-i » dans le paragraphe I-2-3, et verrons qu'il s'appuie sur la présence d'un ion métallique, capable de renforcer les interactions CMH-TCR.
4. **Absence de peptide** : nous avons également modélisé un complexe entre le CMH, le médicament directement placé dans la poche du CMH (mais sans le peptide) et le TCR. Cette hypothèse nous est venue lorsque nous avons débuté nos modèles de complexes et avons constaté le peu de place laissée à l'antigène entre le CMH et le TCR.

Pour chaque scénario envisagé, le type d'analyse moléculaire est basé sur le même schéma : Nous créons les modèles des différents partenaires du complexe, puis nous le soumettons à une dynamique moléculaire afin de vérifier la stabilité et la viabilité du complexe (en effet, il n'y a que peu d'espace entre un TCR et le CMH). Ensuite, nous analysons les interactions entre les partenaires (liaisons faibles de type hydrogène ou hydrophobe, position du médicament...).

Chaque simulation de dynamique moléculaire est effectuée à l'aide du logiciel NAMD, avec le champ de forces CHARMM27. Les simulations sont réalisées en phase aqueuse décrite, grâce à des molécules d'eau de type TIP3P. Les différents complexes sont modélisés et minimisés avec le logiciel Sybyl X qui utilise le champ de forces Tripos et les charges partielles de l'algorithme de Gasteiger et Marsilli. La structure est ensuite préparée, au moyen du logiciel VMD, pour la simulation de dynamique moléculaire, qui elle, est effectuée avec NAMD. Le temps de simulation est de 10 nanosecondes avec un pas d'intégration de 1 femtoseconde. Les analyses sont effectuées à une température constante de 300 K, et sous une pression constante de 1 atm. La température est contrôlée par le thermostat de Berensen et la pression par le piston de Langevin. Toutes les 20 picosecondes, une conformation est échantillonnée. L'ensemble des simulations a été calculé sur les supercalculateurs de l'IBCP « zèbre » et « jaguar ».

Pour chacune des simulations de dynamique moléculaire, une analyse est faite grâce au RMS 2D (Root Mean Square). Le but est de détecter des variations de conformation de la structure. Pour cela, chaque conformation échantillonnée sera comparée à toutes les autres. Pour chacun des couples de conformations, les structures sont alignées sur les atomes lourds et la distance quadratique moyenne (RMS) est calculée. Le résultat se présente sous forme d'une matrice carrée et colorée, qui permet de déterminer quelles sont les conformations stables lors de la simulation de dynamique. A partir de ces conformations clefs, nous évaluons les interactions de type liaison hydrogène entre le peptide et le CMH, ainsi que le peptide et le TCR.

## I-1- Les modèles :

### I-1-1- CMH :

Nous avons choisi de partir d'une structure de CMH de classe 2 (HLA-DR1, chaîne bêta), co-résolue avec un récepteur T et un peptide antigénique, par Hennecke *et al.*, en 2001. Le code PDB est 1 FYT, et le complexe CMH-peptide-TCR a été résolu par cristallographie de rayons X (résolution de 2,60 Å).

La figure 29 (paragraphe I-1-4) présente le complexe formé par la molécule de CMH classe 2, le peptide antigénique et le TCR.



### I-1-2- Médicament :

Le chapitre 4 présente les différents modèles d'antibiotiques étudiés. Nous avons modélisé les 42 bêta-lactamines, et pour les raisons déjà citées, nous nous sommes focalisés sur l'amoxicilline. Il sera intéressant dans la suite de ce travail d'étendre chaque analyse aux autres bêta-lactamines, voire à d'autres classes thérapeutiques.

### I-1-3- Peptide :

La structure 1FYT présente, entre le CMH et le TCR, un peptide antigénique. Il est positionné dans la niche à peptides du CMH et sur lequel nous fixons covalamment le médicament.

Ce peptide provient du virus *Influenza A*. C'est l'hémagglutinine chaîne HA 1. Sa longueur est de 13 résidus. La séquence est la suivante : PKYVKQNTLKLAT.

Dans le paragraphe suivant, à travers notre collaboration avec l'équipe du Pr. Pallardy, nous détaillerons les autres séquences peptidiques issues de notre analyse SuMo et susceptibles d'être hapténisées.

### I-1-4- TCR :

La structure 1 FYT présente les chaînes alpha et bêta d'un récepteur de lymphocyte T. La figure 29 présente le modèle du complexe entre le TCR placé au dessus, le peptide au centre, et le CMH positionné en dessous.

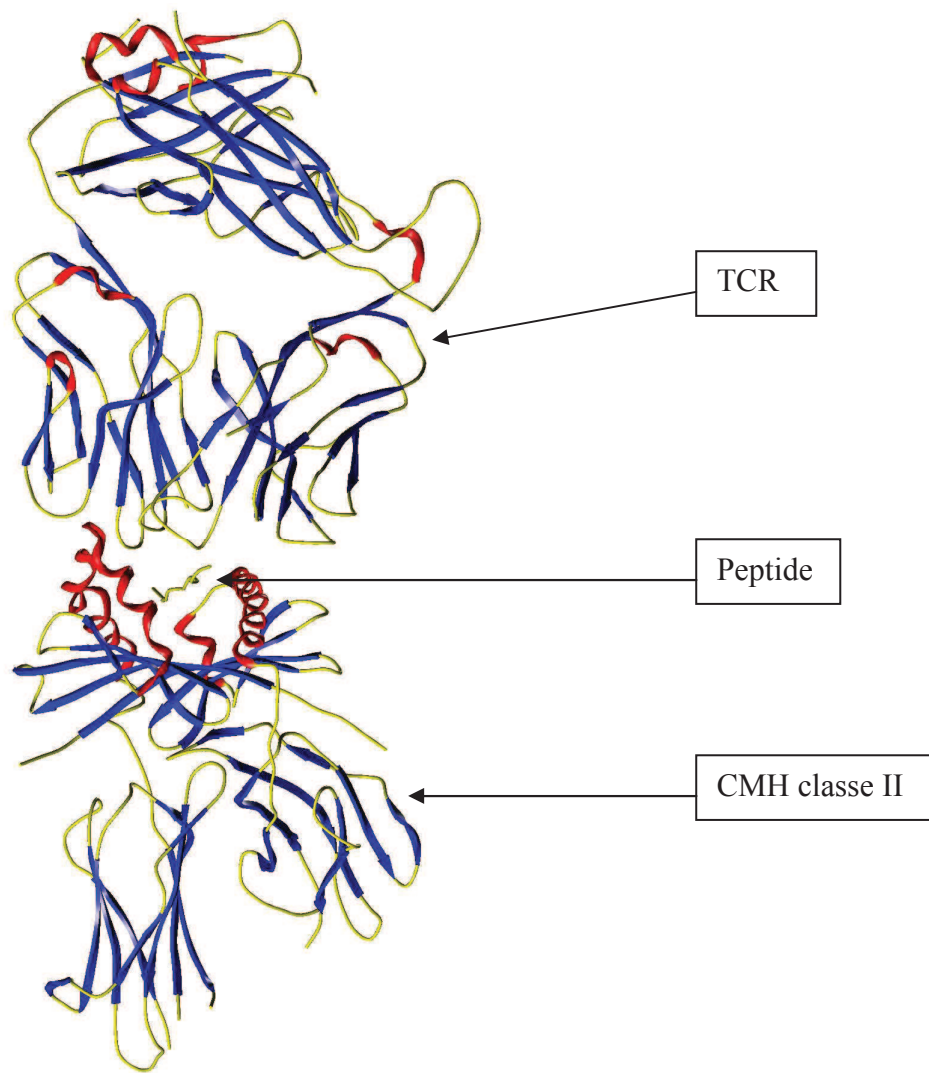


Figure 29 : Complexe entre une molécule de HLA-DR1, le peptide hémagglutinine et le TCR

Notons que le choix de ce complexe est lié d'une part au fait qu'il a été co-résolu avec les trois acteurs que nous voulions étudier, d'autre part que nous voulions travailler sur des CMH de classe 2, et enfin qu'il fallait un peptide reconnu par un grand nombre de CMH. La prédiction de la reconnaissance d'un peptide au sein de la poche d'une molécule de CMH est réalisée au moyen d'un outil de prédiction appelé PROPRED disponible en ligne à l'adresse : <http://www.intech.res.in/raghana/propred/>

## I-2- Les scénarios :

### I-2-1- Mode classique :

Dans cette situation, et du fait de son faible poids moléculaire, la bêta-lactamine doit, pour être reconnue par le système immunitaire, se fixer de manière covalente à une protéine porteuse. Par la suite, une CPA va internaliser la protéine hapténisée, la dégrader puis présenter au sein de sa niche à peptides (qui dans le cas de CMH classe II présente des peptides de 13 acides aminés au moins), les peptides issus de la dégradation et qui sont ou non hapténisés par le médicament. Les lymphocytes T CD4, via leur TCR, vont venir interagir spécifiquement avec l'ensemble CMH/peptide ou CMH/peptide hapténisé.

Nous avons ainsi ajouté, au complexe 1 FYT, une molécule d'amoxicilline et avons vérifié la stabilité de l'assemblage final. La séquence du peptide dans la structure est la suivante : **PKYVKQNTLKLAT**. Cette séquence de 13 acides aminés possède 3 lysines en positions 2, 5 et 10 mais aucune sérine. Nous avons étudié, au niveau de la zone de contact entre le TCR et le peptide, la place nécessaire pour insérer une molécule de type bêta-lactamine. Il nous paraît très peu probable d'utiliser la position lysine numéro 10 car celle-ci est située dans la partie la plus resserrée du sillon du CMH et cela ne laisse pas de place pour y rajouter une petite molécule. La position 2 est quant à elle beaucoup trop extérieure, et la molécule n'interagirait que très peu avec le TCR. La position 5 est, par contre, optimale en terme de disposition spatiale, et les boucles flexibles du TCR qui reconnaissent l'antigène autorisent l'intercalation d'une petite molécule. Il faut noter que nous ne disposons pas de la séquence d'un TCR spécifique de l'amoxicilline et, à fortiori, pas de sa structure. Toutefois, en utilisant le même modèle de TCR pour toutes les simulations, nous espérons minimiser ce point négatif.

L'analyse par RMS-2D du complexe (figure 30) révèle une structure particulièrement stable. En effet, le RMS moyen est de 1,4 à 1,6 Å au sein de la seule conformation qui existe, et ce tout au long de la simulation (10ns).

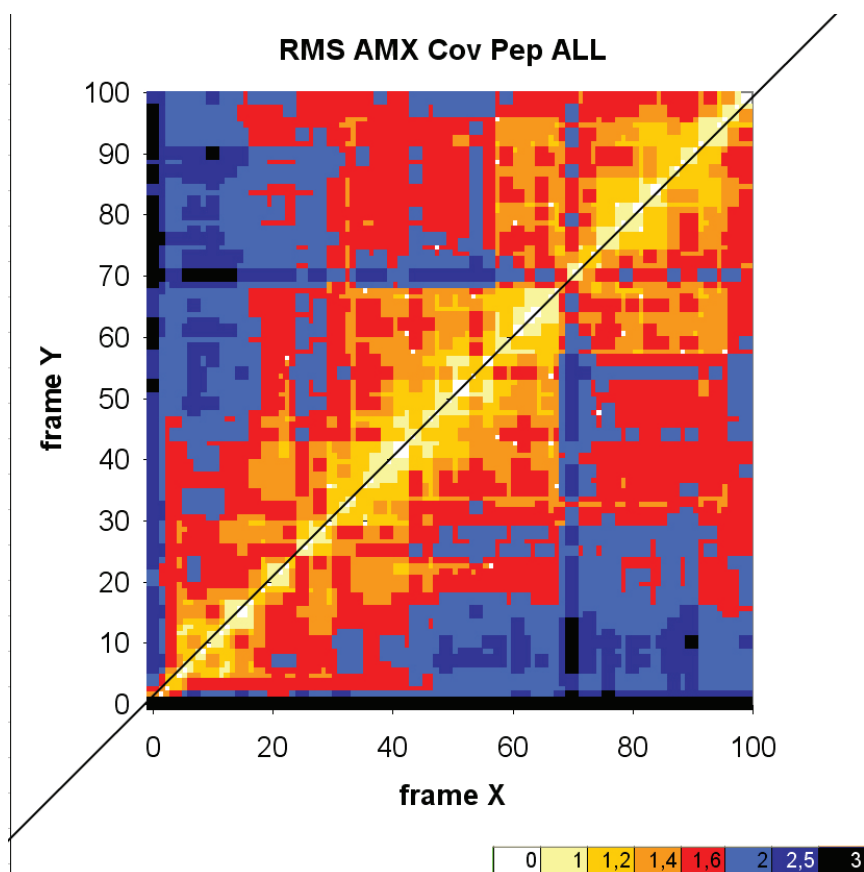


Figure 30 : Analyse par RMS-2D du complexe formé par HLA-DR1, le peptide hapténisé par l'amoxicilline et le TCR

Nous avons analysé la conformation numéro 40 de la trajectoire (environ à la moitié de la simulation), et nous constatons que la molécule chimique a pivoté de façon à se positionner entre les boucles du TCR et à développer des interactions de type liaisons hydrogène à partir de ses fonctions  $\text{NH}_3^+$ ,  $\text{COO}^-$  et hydroxyle du groupement phénol. L'étude des liaisons hydrogène montre que celles-ci persistent durant environ 2 ns et qu'elles sont récursives. Ainsi, le médicament crée des liaisons hydrogène avec 3 à 4 acides aminés autour de ses fonctions chimiques. La figure 31 représente la conformation 40 de la simulation. Le TCR est placé au dessus, la molécule d'amoxicilline est enchâssée entre les boucles du TCR et le CMH est placé dans la partie inférieure.

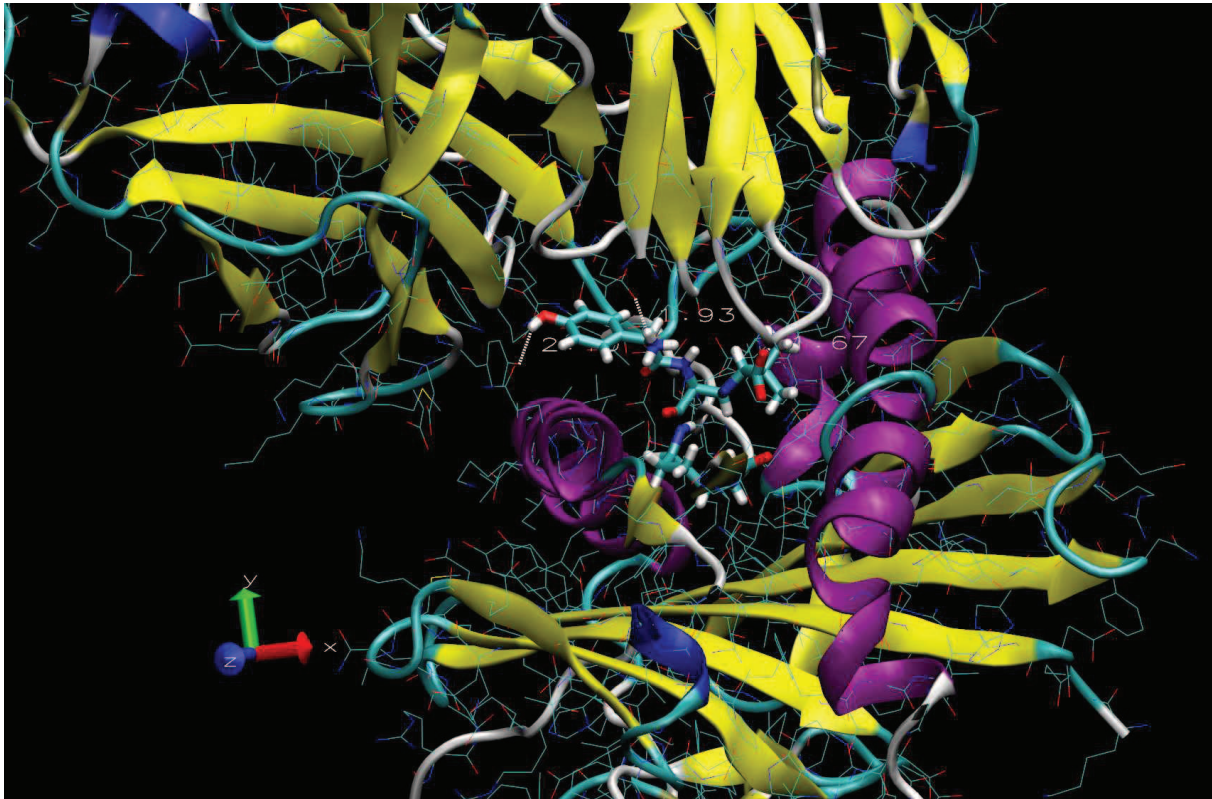


Figure 31 : Conformation stable issue de la dynamique moléculaire du complexe HLA-DR1, peptide hapténisé par l'amoxicilline et TCR

Cette simulation montre donc que la molécule développe quatre liaisons hydrogène avec le TCR sans que la conformation du peptide soit modifiée, et donc sans perturber l'interaction peptide-CMH. Nous concluons d'après cette analyse, que la molécule d'amoxicilline contribue à la stabilisation du complexe CMH/TCR.

## I-2-2- P-i concept: Pharmacological interaction with the immune system:

### I-2-2-1- Principe

Plusieurs éléments poussent à penser que le processus d'hapténisation n'est pas un mode d'action unique dans la reconnaissance de l'allergène par les lymphocytes T. En effet, des expériences montrent qu'il peut y avoir des réactions qui surviennent très rapidement après la prise d'un médicament chez un individu non sensibilisé, ou sensibilisé. Ce court délai n'est pas compatible avec le temps nécessaire au processus d'hapténisation-apprêtement. D'autre part, W.Pichler a mené des études basées sur le blocage des fonctions présentatrices

des CPA. Ainsi, même en absence d'apprêtement, on peut avoir activation des LT. Ce concept est très étudié notamment pour le sulfaméthoxazole (SMX), la lidocaïne, ou la carbamazépine par exemple. Ces molécules chimiques peuvent directement activer les lymphocytes T, sans apprêtement ni métabolisme, en se liant de façon non covalente, comme le ferait n'importe quel médicament sur son récepteur (concept dit pharmaco-immunologique ou P-i concept), sur les molécules HLA de classe II des CPA ou sur les TCR.

Le p-i concept n'est pas en opposition avec les concepts d'haptène et de pro-haptène, mais en complémentarité. Il est basé sur l'analyse de clones T chez des patients allergiques. Ainsi, cliniquement, le p-i concept peut expliquer la survenue rapide de manifestations immuno-allergiques vis à vis d'un médicament. L'ensemble de ces observations et expériences a mené W. Pichler à définir un autre type de reconnaissance du médicament par le TCR que la reconnaissance classique mettant en jeu hapténisation et apprêtement. Le p-i concept énonce que l'allergène (ici, l'antibiotique) interagirait via des liaisons non covalentes avec les molécules de CMH d'une part, et d'autre part avec le TCR. Les liaisons entre TCR et médicaments seraient nombreuses ce qui permettrait un contact étroit et renforcé.

Toutefois, pour le moment, il n'y a pas eu validation du p-i concept avec les bêta-lactamines.

Notre point de vue de modélisateur, et l'ensemble des modèles que nous avons créé, nous a poussé à émettre une nouvelle théorie : le New p-i concept que nous présenterons dans le paragraphe I-2-3.

#### I-2-2-2- Modélisation

Nous avons utilisé la même structure cristallographique pour l'ensemble des simulations, à savoir le complexe 1FYT. Nous avons inséré l'amoxicilline (mais sans liaison covalente contrairement au cas précédent) au niveau de l'acide aminé numéro 5 du peptide. L'ensemble a été soumis à une dynamique moléculaire puis analysé de la même façon que précédemment. L'étude du RMSD-2D révèle que le complexe se stabilise modérément en deux sous-familles. Le RMS moyen est de 2 Å. Nous constatons que ces deux sous-familles sont assez différentes, puisqu'elles possèdent des RMS moyens respectifs de 2,5 et 3 Å.

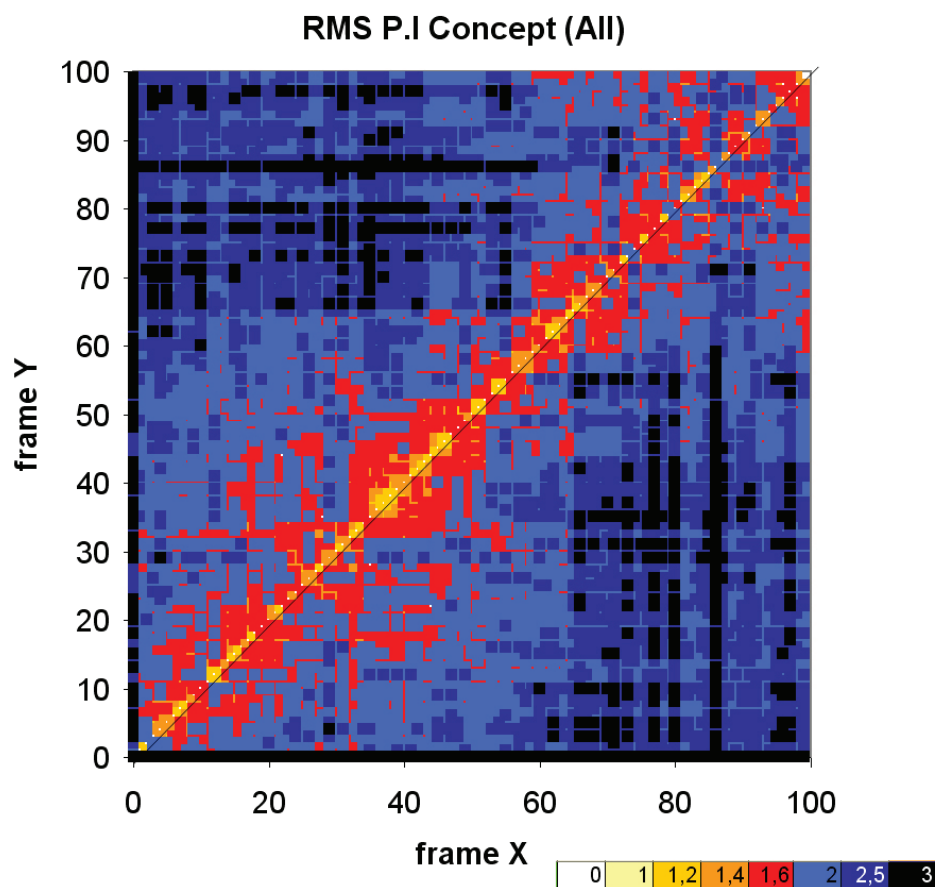


Figure 32 : Analyse par RMS-2D du complexe formé par HLA-DR1, le peptide en interaction non covalente avec l'amoxicilline et le TCR

L'analyse des conformations représentatives des deux sous-familles montre que la position de la molécule d'amoxicilline évolue au cours de la simulation. Dans les deux cas, la molécule s'est stabilisée grâce à des liaisons hydrogène développées à partir de ses fonctions  $\text{NH}_3^+$ , OH (du groupement phénol) et  $\text{COO}^-$ , en direction du TCR et des hélices qui bordent le CMH. La figure 33 représente le complexe dans sa conformation représentative de la première sous-famille. Quatre liaisons hydrogène (trois en direction du TCR et une vers l'hélice droite du CMH) assurent un bon enchâssement de l'antibiotique et contribuent à stabiliser le complexe CMH/TCR. Le TCR est placé dans la partie supérieure, l'amoxicilline au centre et le CMH en bas (hélices colorées en violet).

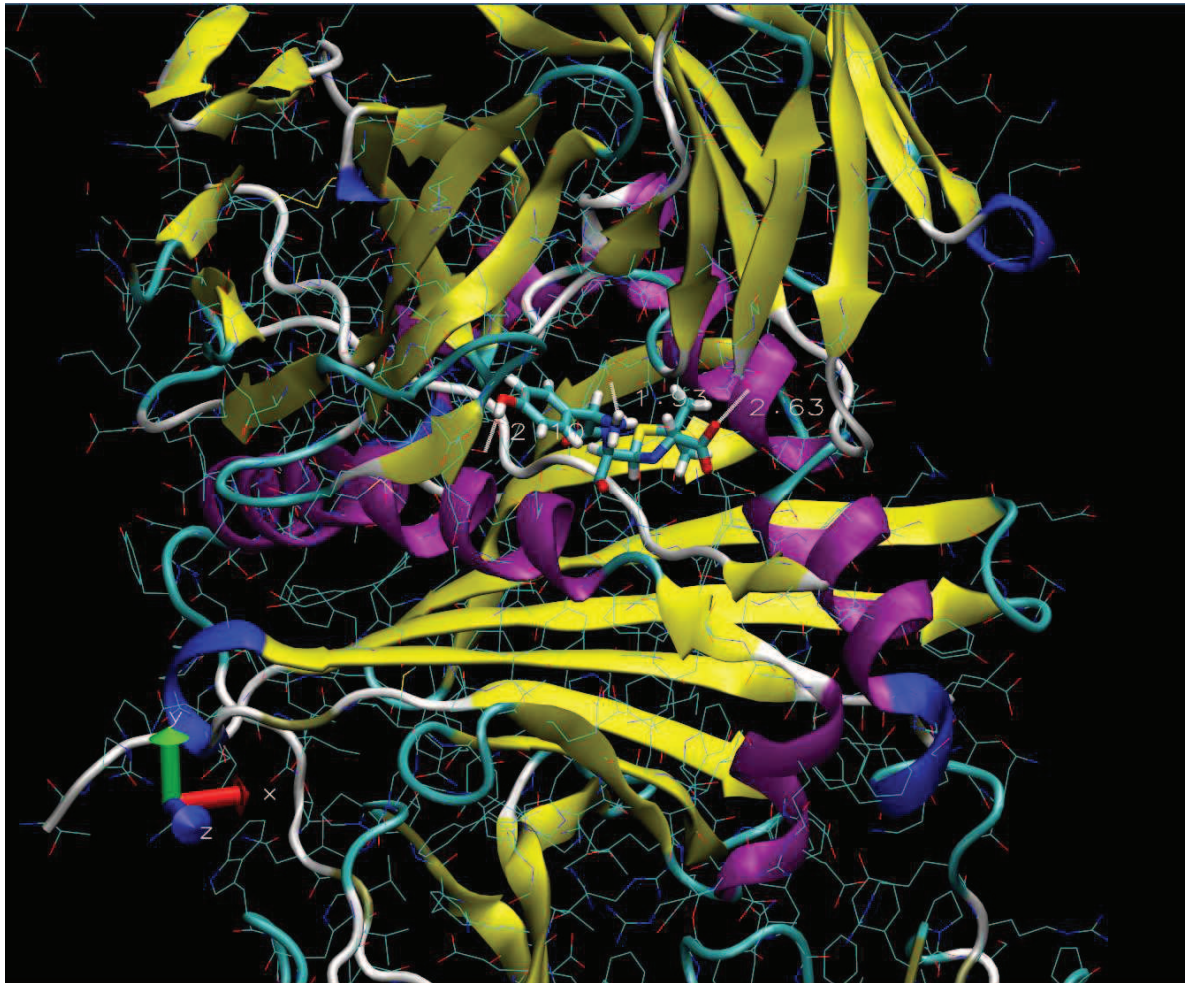


Figure 33 : Complexe formé entre HLA-DR1, le peptide en interaction non covalente avec l'amoxicilline et le TCR selon la configuration de la première sous-famille conformationnelle

La simulation montre que la présence de l'amoxicilline augmente la stabilité du complexe CMH-TCR, grâce à la formation de plusieurs liaisons hydrogène. Aussi, ce scénario est, pour nous, possible. Toutefois, il faut remarquer que dans le cas du p-i concept, le complexe sera beaucoup moins fort et stable que lorsqu'il y a hapténisation. En effet, dans le « scénario hapténisation », le médicament développe 4 liaisons hydrogène avec le TCR, de plus il est lié de façon covalente au peptide qui lui-même interagit avec le CMH par un réseau de liaisons hydrogènes. Dans le cas du p-i concept, le médicament développe seulement 3 liaisons hydrogène avec le TCR, aucune liaison avec le peptide et une seule liaison hydrogène avec le CMH.



## I-2-3- “New P-i”

### I-2-3-1- Principe

Ce concept a été étudié en collaboration avec le groupe de travail sur le New p-i. Les travaux, menés par modélisation moléculaire et chimie quantique, sur des interactions moléculaires mises en jeu dans le scénario du « p-i concept », pour l’amoxicilline, montrent que ces interactions seraient insuffisantes en nombre et en énergie pour aboutir au déclenchement d’une réaction immunitaire. Nous savons par ailleurs, que les médicaments qui répondent au « p-i-concept » sont des ligands du cuivre et du fer (avérés pour certains, probables pour les autres). Nous ne remettons pas en question le p-i concept mais avons émis l’hypothèse que l’interaction entre le CMH, le peptide, le médicament et le TCR ferait également intervenir un ion métallique, qui permettrait de renforcer les interactions entre les éléments du complexe. La figure 34 schématise ce que nous appelons le New p-i (M représente le métal).

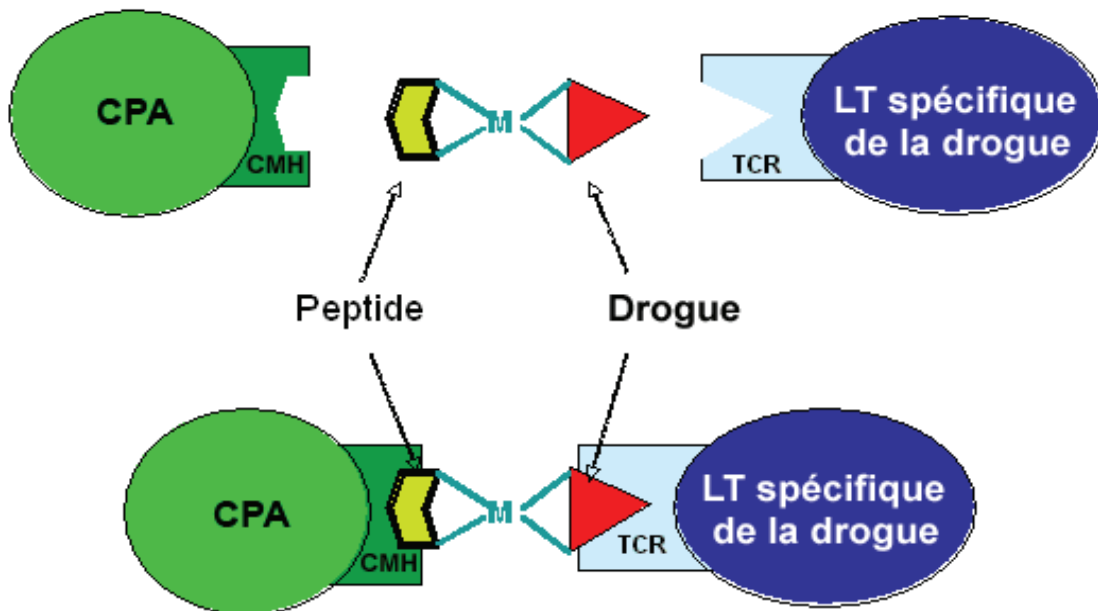
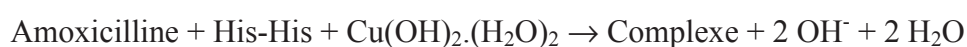


Figure 34 : Vue schématique du New « pharmacological interaction with immune receptors » Concept

Dans cette situation, les liaisons créées par le métal seront datives. Ce sont des liaisons non covalentes qui sont plus fortes que des liaisons de type hydrogène ou hydrophobe (telles que celles intervenant dans le p-i concept). Ces liaisons datives, fortes, du métal impliqué dans le complexe mixte « peptide/métal/médicament », maintiendraient l'ensemble de manière assez solide, et s'apparenteraient donc à la liaison covalente que l'on retrouve dans la situation classique d'hapténisation. Dans ce cas, l'ion métallique serait enfermé dans une cage constituée par le peptide et le médicament, amenant ainsi le complexe mixte « peptide/ion métallique/médicament » à ressembler à un haptène « Peptide/liaison covalente/médicament ». Ainsi donc, vu des récepteurs, ce complexe ressemblerait à un haptène ordinaire et l'activation des LT spécifiques du médicament serait alors possible.

### I-2-3-2- Modélisation

Pour cette partie nous avons dû réaliser une étape préalable afin de modifier le peptide. En effet, celui-ci doit développer des interactions de type liaisons datives afin de former le complexe peptide / métal / antibiotique. Pour cela, nous avons modélisé un complexe mettant en jeu le cuivre II et différents acides aminés afin de savoir lesquels étaient capables de créer un tel complexe ainsi que les positions qu'ils adoptaient. L'étude a été réalisée avec le logiciel de mécanique quantique ADF qui utilise la DFT (Density Functional Theory) (Parr, 1989) ainsi que la méthode GGA/PW d'Approximation des Gradients Généralisée et l'algorithme de Perdew/Wang (Perdew, 1996). L'orbitale atomique de base ayant servi pour les calculs correspond à une base en double potentiel zêta (DZ). Le résultat montre que seul le résidu histidine peut induire une liaison dative avec le cuivre II. L'ion  $\text{Cu}^{2+}$  présentant une géométrie dite « plan carré ». Il est nécessaire que deux liaisons datives soient générées vers le peptide ainsi que deux autres vers la molécule. Les deux histidines peuvent être soit consécutives soit séparées par une glycine. La distance mesurée, après minimisation énergétique, entre le  $\text{Cu}^{2+}$  et les oxygènes de l'amoxicilline est de 2,3 Å. Elle est de 2,1 Å avec les azotes de l'histidine. La séquence sans glycine reste la meilleure en terme géométrique. Nous avons calculé la variation d'enthalpie de formation des différents composés et en déduisons que la réaction de complexation est thermodynamiquement favorisée ( $-96 \text{ kJ.mol}^{-1}$ ) et peut s'écrire :



La géométrie du complexe après minimisation énergétique est exposée sur la figure 35, qui montre parfaitement la géométrie de type plan carrée et la stabilité du complexe.

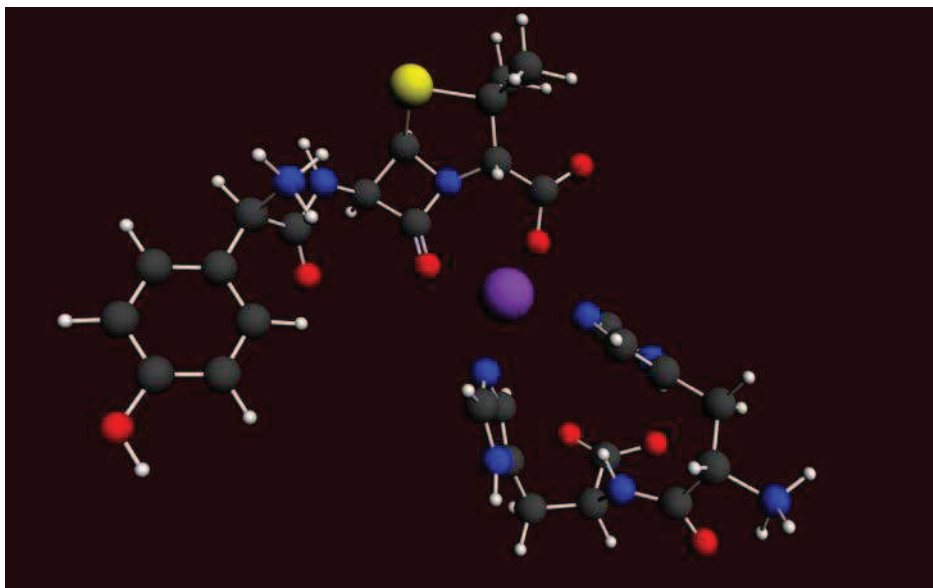


Figure 35 : Structure du complexe entre une molécule d'amoxicilline, le cuivre  $\text{Cu}^{2+}$  et le dipeptide histidine-histidine

Partant de ces résultats, nous avons modifié la séquence du peptide cristallisé au sein du complexe CMH-TCR (1FYT) afin de rajouter les histidines nécessaires. Ainsi, la séquence d'origine: PKYVKQNTLKLAT devient PKYHHQNTLKLAT. Par la suite, nous avons placé l'ion cuivre II ainsi de la molécule d'amoxicilline de manière à respecter la géométrie du complexe issu de notre calcul de chimie quantique. Nous avons modélisé le complexe CMH/peptide modifié/cuivre II/amoxicilline/TCR grâce au logiciel Sybyl X. Nous avons imposé des contraintes géométriques au système afin que les différents acteurs restent dans une conformation plan carré. Nous avons ensuite transféré la structure minimisée dans le logiciel VMD et simulé la dynamique moléculaire avec NAMD suivant le protocole décrit précédemment.

L'analyse de la trajectoire par RMS-2D, présentée sur la figure 36, révèle que le complexe se maintient et que l'interaction CMH / TCR est stabilisée. Le fait que le RMS moyen de la famille centrale soit inférieur à 2 Å montre très clairement que le super complexe est très stable.

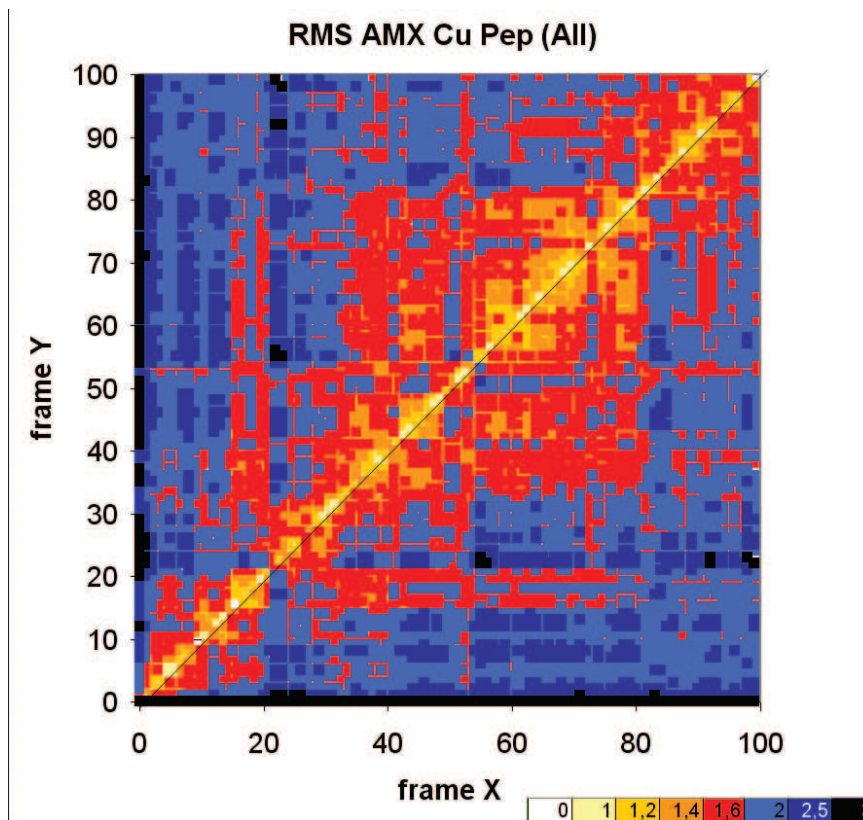


Figure 36 : Analyse par RMS-2D du complexe formé entre HLA-DR1, le peptide modifié, le cuivre, l'amoxicilline et TCR

Cependant, si l'on regarde la conformation représentative de la famille centrale (frame N°62), on remarque que la structure de type plan carré a été détruite. En effet, l'ion  $\text{Cu}^{2+}$  a préféré créer une liaison dative avec la fonction carbonyle de l'amoxicilline et des liaisons datives avec l'acide aminé Q325 de l'hélice qui borde le CMH à droite. La figure 37 montre le complexe étudié avec en haut le TCR, au centre l'amoxicilline (en bâtons), le peptide en interaction avec le cuivre (boule orange), et en bas, le CMH. Il faut noter que l'hélice alpha (colorée en violet) du CMH est partiellement déstabilisée par cette interaction, mais que les liaisons hydrogène du squelette formant cette hélice ne sont pas touchées. Elle est donc faiblement déstabilisée comme le montre la partie bleue de l'hélice. Cette analyse montre que la modification du peptide a été inefficace et que le cuivre n'interagit pas avec le peptide. Par contre, la molécule d'amoxicilline développe des interactions de type liaisons hydrogène avec le TCR par ses fonctions  $\text{NH}_3^+$  et OH (du phénol).

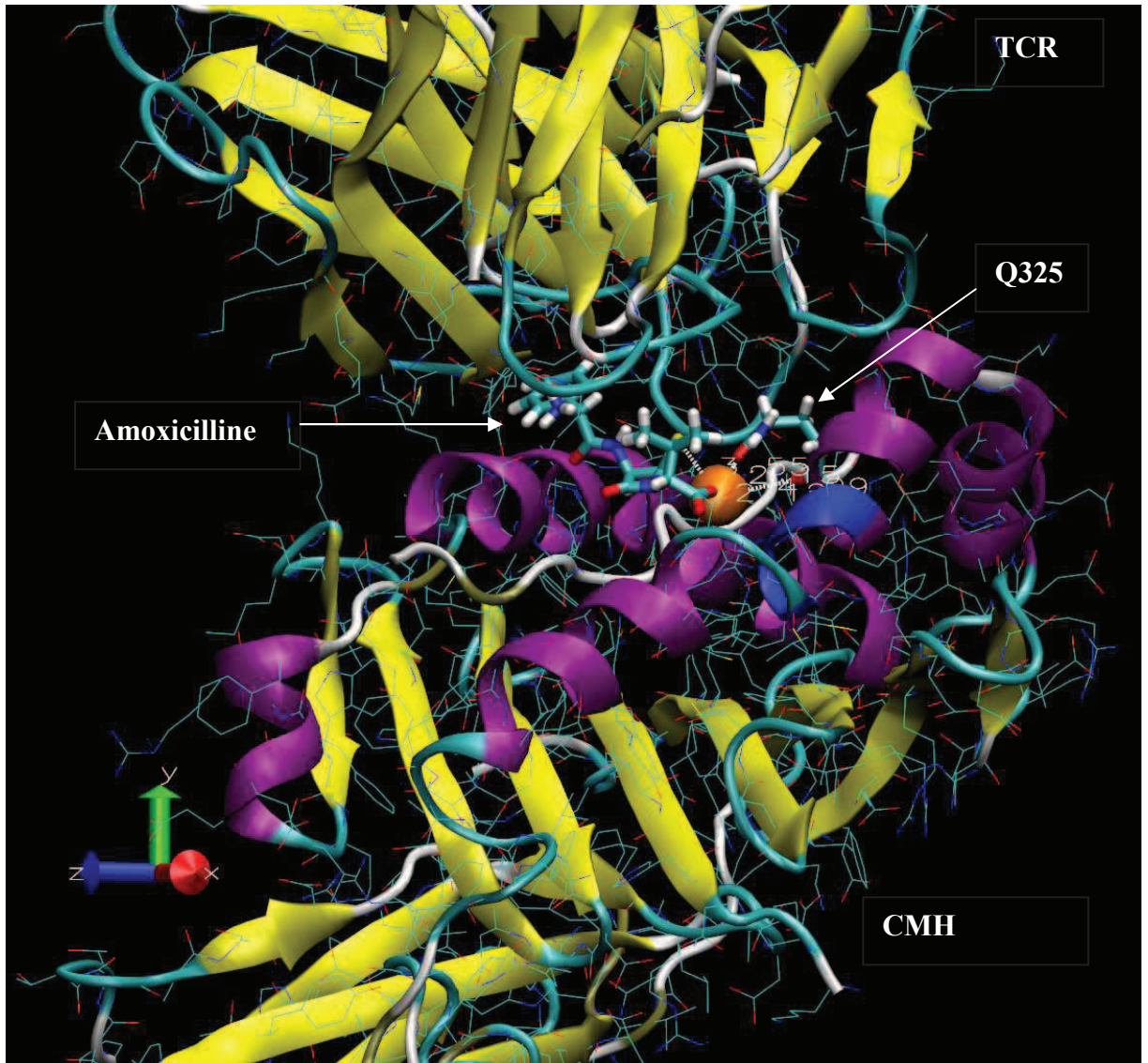


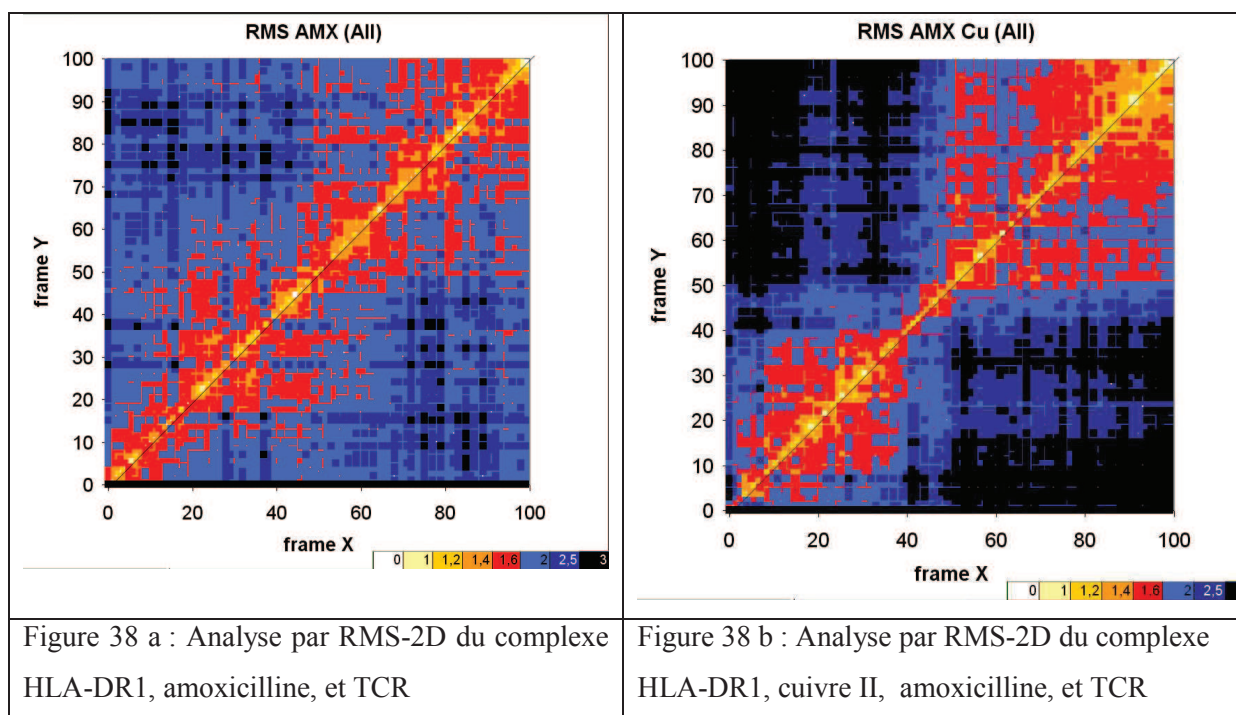
Figure 37 : Conformation du complexe entre HLA-DR1, peptide modifié, cuivre, amoxicilline et TCR.

Cette simulation nous permet de conclure que le New-p-i concept s'accompagne d'une augmentation de l'énergie d'interaction entre le CMH et le TCR. En effet, l'amoxicilline permet de créer 3 liaisons hydrogène (comme dans le p-i concept) et que le pontage de la molécule via l'ion  $\text{Cu}^{2+}$  permet d'ancrer l'antibiotique dans le CMH avec une énergie bien plus forte que celle générée par des liaisons hydrogène lors du p-i concept. Ce pontage se fait de façon indépendante au peptide, qui d'après nous, doit contribuer à la stabilisation du complexe CMH/ TCR.

## I-2-4- Absence de peptide

Une dernière hypothèse est venue au cours de cette phase du projet, en raison de nos observations concernant la place très limitée au niveau de la synapse immunologique. En effet, nos modèles théoriques tendent à montrer que le super complexe CMH-peptide-médicament-TCR est très volumineux. Nous avons donc imaginé un scénario dans lequel le TCR reconnaîtrait le médicament positionné directement dans la molécule de CMH, qui ne présenterait pas de peptide. Ainsi, le complexe de déclenchement du mécanisme immunitaire serait constitué d'une molécule de CMH (sans peptide), une molécule d'antibiotique et un TCR.

A partir de la structure 1FYT, nous avons supprimé le peptide central puis nous l'avons remplacé par une molécule d'amoxicilline. Nous avons construit deux modèles : le premier avec un ion  $\text{Cu}^{2+}$  et une molécule d'amoxicilline, le second uniquement avec une molécule d'amoxicilline. Après minimisation énergétique par le logiciel Sybyl X, nous avons effectué deux simulations de dynamique moléculaire grâce au logiciel NAMD en suivant le protocole détaillé plus haut. L'analyse des trajectoires par RMS-2D montre que les deux modèles sont instables. Les figures 38 a et 38 b présentent les RMS-2D des complexes.



S'agissant du complexe sans le métal (Figure 38 a), on constate que la structure est très instable et qu'aucune sous-famille conformationnelle ne se dessine vraiment.

L'ajout d'un ion métallique provoque de fortes modifications. En effet, deux grandes sous-familles apparaissent et sont très différentes comme le prouvent leurs RMS dont les valeurs divergent de plus de 3 Å. Ces observations nous permettent de conclure que l'ion métallique ne stabilise pas le complexe.

Nous avons analysé les structures des deux complexes mais la position de l'amoxicilline varie énormément et ne se maintient pas dans le sillon du CMH. Aussi, il nous semble peu probable que ce type de complexe puisse stabiliser suffisamment l'interaction CMH/TCR. Nous pensons donc que le scénario excluant le peptide n'est pas valable.

### I-3- Conclusions :

Le calcul d'une simple variation d'énergie ( $\Delta E$ ) par mécanique moléculaire n'est pas suffisamment fiable pour conclure de manière certaine sur les mécanismes mis en jeu. Toutefois, il est possible de faire un comparatif, en prenant des références thermodynamiques. Une liaison dative impliquant le  $\text{Cu}^{2+}$  et un carbonyle représente une énergie de  $60 \text{ kJ.mol}^{-1}$  (Terreux, 2006). Une liaison hydrogène a une énergie voisine de  $40 \text{ kJ.mol}^{-1}$ . Une liaison covalente correspond à plusieurs centaines de  $\text{kJ.mol}^{-1}$ . Nous avons récapitulé succinctement les interactions mises en jeu pour les trois scénarios pertinents: haptène, p-i et new p-i. Le but est d'évaluer l'énergie d'interaction (entre le TCR et le CMH) qui est apportée par la molécule médicamenteuse. Le système peut se voir comme une chaîne constituée de différents maillons retenant d'un côté le TCR, et de l'autre le CMH. La résistance d'une chaîne correspondant à la résistance du maillon le plus faible, essayons de déterminer les résistances pour nos trois hypothèses.

Le tableau 16 ci-après résume notre analyse des interactions calculées entre les différents acteurs impliqués dans l'hapténisation, le p-i concept, ou le new p-i concept.

Interaction	Scénario		
	haptène	p-i	new p-i
TCR / Molécule	4 liaisons hydrogène	3 liaisons hydrogène	3 liaisons hydrogène
Molécule / CMH	Rien	1 liaison hydrogène	1 liaison dative
Molécule / Peptide	Liaison covalente	rien	rien
Peptide / CMH	Nombreuses liaisons hydrogène	Nombreuses liaisons hydrogène	Nombreuses liaisons hydrogène
Résistance du maillon le plus faible	4 liaisons hydrogène	1 liaison hydrogène	1 liaison dative
Energie (kJ.mol <sup>-1</sup> )	80	40	60

Tableau 16 : Analyses des interactions mises en jeu lors d'une reconnaissance de l'amoxicilline par le système CMH/TCR

L'analyse des trois scénarios montre bien que le cas le plus favorable est le mode haptène. Le concept p-i permet d'augmenter d'une liaison hydrogène la chaîne d'interactions entre le CMH et le TCR. Le dernier scénario du new p-i est intéressant au niveau énergétique, puisque nous sommes presque au niveau énergétique du mode haptène.

Les conclusions concernant cette phase du projet doivent absolument tenir compte du fait que l'ensemble des modèles et analyses a été réalisé à partir d'une structure non spécifique de l'amoxicilline. Aussi faudra-il reprendre ce travail lorsqu'un complexe CMH, amoxicilline et TCR spécifique de l'amoxicilline sera résolu expérimentalement.

Cependant, nous avons avec le modèle 1FYT, une bonne idée des phénomènes intervenant lors de la reconnaissance de l'antibiotique par le TCR. Nos résultats nous orientent donc vers les conclusions suivantes : une reconnaissance après hapténisation apprêtement est tout à fait pertinente. Le p-i concept pour lequel il n'y a pas de liaison covalente entre l'allergène et le peptide est également plausible en terme de stabilité du complexe CMH-TCR. Le new p-i concept qui fait intervenir un ion métallique se rapproche du p-i et de ce fait peut se produire. Quant au scénario qui ne fait intervenir aucun peptide dans la niche du CMH, nous ne pensons pas qu'il puisse être impliqué dans des phénomènes d'hypersensibilités allergiques.



Ces résultats de modélisations seront par la suite validés expérimentalement.

## II- Validation expérimentale :

Dans le but de mettre en application nos outils et de valider nos résultats, nous avons mis en place une collaboration avec l'équipe du Pr. M. Pallardy, de l'Université de Paris – Sud XI. Dans cette étude, nous nous plaçons dans le scénario classique d'hapténisation suivi d'un apprêtement. L'objectif principal consiste à étudier le répertoire des LT spécifiques de la pénicilline et du SMX, et à identifier les peptides présentés à ces lymphocytes. Ce projet est en cours et nous ne présentons ici que les premiers résultats. Par ailleurs, nous ne parlerons que des analyses qui touchent à la pénicilline et prévoyons de les appliquer à l'amoxicilline très prochainement. Deux questions se sont posées : Quels sont les peptides présentés par les molécules HLA de classe II dans le cas de la pénicilline ? Existe-t-il un répertoire de LT capables de reconnaître ces peptides ? Si oui, quelle en est la fréquence ?

Afin de répondre à ces questions, la démarche expérimentale fait apparaître trois phases : 1- Synthèse de protéines et de peptides couplés à la pénicilline 2- Tests d'interaction avec les molécules de CMH, et enfin 3-Tests *in vitro* de stimulation des LT

De notre côté, nous avons recherché des protéines capables de prendre en charge le médicament, ainsi que les peptides, issus de la dégradation de ces protéines au cours de leur apprêtement, et qui possèdent donc un acide aminé catalytique responsable de la liaison covalente avec l'antibiotique. Les séquences peptidiques ainsi trouvées sont en cours de synthèse et serviront à stimuler les LT naïfs (CD4). Ce travail se déroule en deux grandes étapes :

1- Synthèse de bioconjugués : Nous souhaitons synthétiser de l'albumine sérique humaine sur laquelle sont fixées des molécules de pénicilline radiomarquées. On parle alors de bioconjugués albumine-pénicilline qui sont utilisés afin de tenter d'induire la maturation des cellules dendritiques et de stimuler des LT.

2- Identification des protéines d'hapténisation et des peptides possédant un acide aminé catalytique nécessaire à fixer le médicament. Ces peptides seront synthétisés avec la pénicilline radiomarquée (fixée covalamment) afin d'étudier la stimulation de lymphocytes T.

## II-1- Bioconjugués albumine – pénicilline :

Cette partie du travail est réalisée par l'équipe du Pr. Marc Pallardy (et notamment par Mlle Cathy Nhim), aussi, nous ne la présenterons que brièvement puisque nous ne sommes pas directement intervenus en tant que modélisateurs. Cependant, le protocole de synthèse des bioconjugués nous a permis d'avancer, comme nous le décrivons dans le chapitre 5, sur le versant immédiat des hypersensibilités allergiques. D'autre part, les résultats de cette phase sont extrêmement importants puisqu'ils nous permettent de connaître le nombre de molécules de pénicilline fixé sur une protéine (albumine), ainsi que le site d'hapténisation.

Deux protocoles sont envisagés. Tout d'abord, un travail de couplage qui reproduit les conditions physiologiques. Cette phase a déjà été réalisée par C. Nhim et elle conclut qu'au pH physiologique, une seule molécule de pénicilline sera fixée par molécule d'albumine (travail sur la human serum albumin). Ceci confirme notre hypothèse selon laquelle une protéine ne peut être hapténisée que par très peu de molécules de bêta-lactamine car l'hapténisation implique une conformation particulière du site de fixation (ce qui permet une bonne énergie d'activation nécessaire à la réaction chimique de liaison covalente).

L'autre possibilité est de travailler à pH basique afin de saturer l'albumine en pénicilline.

L'ensemble des manipulations de couplage est réalisé par le Dr. Delphine Joseph. Nous ne décrivons pas ici le protocole mis en place. Par la suite, l'analyse des positions de liaison de la pénicilline marquée sur la protéine est réalisée par le Dr. David Touboul grâce à la spectrométrie de masse. Ainsi, il est possible de connaître les acides aminés impliqués dans la fixation covalente de la pénicilline. Nous attendons actuellement les résultats de ces manipulations.

Notons qu'il nous faudra appliquer la même démarche pour valider notre étude théorique sur l'hapténisation de l'albumine par l'amoxicilline, et cela fait partie de nos objectifs à très court terme. Cela nous permettra de confirmer, ou d'infirmer nos hypothèses exposées dans le chapitre 4 et tirées de nos études de docking et QM/MM. D'une part, nous pensons qu'aux conditions physiologiques, l'antibiotique se fixera sur un seul site (ou deux) de l'albumine, préférentiellement celui de la lysine 212. D'autre part, nous pensons qu'il faut nécessairement un site dont la conformation est adaptée afin que les interactions avec la

drogue fournissent suffisamment d'énergie pour déclencher la réaction de fixation. Nous avons défini plusieurs sites susceptibles de remplir ces conditions énergétiques, (étude de docking), et en réalisant des expériences de couplage, nous pourrions valider ces sites.

Par la suite, ces bionconjugués serviront à analyser : la reconnaissance du médicament par les cellules dendritiques en tant que signal danger, la maturation de ces cellules dendritiques (qui se traduit par l'expression de différents marqueurs tels que les CD40, CD86 ou CD83), la stimulation des LT. Ces travaux sont en cours de réalisation par l'équipe du Pr. M. Pallardy et ne seront pas ici développés.

## II - 2- Bioconjugués peptide-pénicilline

Comme nous l'avons dit, nous avons recherché des protéines susceptibles de fixer le médicament, et avons isolé la séquence peptidique liant de façon covalente la drogue. Pour cela, nous nous sommes appuyés sur le logiciel SuMo et sommes partis du principe présenté dans le chapitre 4, basé sur la comparaison d'un ensemble de protéines sériques humaines avec la cible naturelle des pénicillines : la Protéine Fixant les Pénicillines. Nous demandons donc à SuMo de retrouver quelles protéines (parmi notre base de données de 1536 modèles : Serum DataBank) présentent le plus de similarité avec le site de fixation de la PFP.

Les résultats de cette recherche ont été présentés dans le paragraphe II-3 du chapitre 4 et le tableau 11 liste les premières protéines retrouvées. Les plus pertinentes étant celles qui présentent le plus grand score SuMo, nous avons choisi d'analyser plus spécifiquement les protéines suivantes :

- La **Human Dimethylglycine dehydrogenase** est la première de la liste, ce qui signifie que c'est elle qui présente un site de liaison au médicament le plus proche du site de la PFP. Son score est de 3, 5 et elle possède 5 objets SuMo communs et superposés avec ceux de la référence.

- La **Human Ficolin-2-precursor** : elle est la deuxième de la liste, *ex æquo* avec la L-Ficolin complexée avec la N-acétyl-cystéine. Nous avons choisi de prendre la protéine sans ligand. Son score est de 3 et elle a 6 objets communs, soit un de plus que la Dimethylglycine mais qui elle présente un score plus élevé.

- La **Human Serum albumine** : troisième de la liste, ce qui est en accord avec tout ce que nous résultats présentés jusqu'ici. Elle possède 4 objets communs et un score de 3.
- La **Human Multidrug resistance associated protein 5** a un score de 2, avec 3 objets SuMo
- La **Human Complement C4-A precursor** qui a le même score et le même nombre d'objets communs que la précédente.

A partir de ces protéines, nous voulions identifier les peptides responsables de la fixation du médicament. Les peptides qui vont dans une molécule de CMH classe 2 mesurent au moins 13 acides aminés mais l'interaction avec le CMH se fait avec neuf d'entre eux (les six autres encadrent la séquence qui interagit réellement). Toutefois, on peut déterminer un résidu principal d'ancrage au moyen d'un logiciel Tepitope (Cochlovius et al. 2000, Bian et al. 2004, Nielsen et al. 2010), qui prédit l'interaction entre les molécules de CMH de différents types et le peptide considéré. Pour notre étude, nous avons déterminé des peptides de 15 acides aminés (dont 9 interagissent avec le CMH), puisque pour les tests cellulaires de liaison aux molécules de CMH (réalisés après la synthèse des peptides), il faut des séquences de 15 acides aminés.

La superposition des objets communs avec ceux de la « query » nous indique l'endroit exact, dans la protéine, où peut se fixer la drogue. Nous avons ainsi identifié, pour chacune des protéines sélectionnées, l'acide aminé catalytique responsable de la fixation covalente avec l'antibiotique. Ce résidu est encadré de part et d'autre par 14 acides aminés, formant ainsi la séquence peptidique qui sera reconnue par le CMH, et qui présentera le médicament au TCR.

Ainsi, pour chaque protéine, en partant du site catalytique, une ou plusieurs séquence(s) peptidique(s) de 9 acides aminés a (ont) été identifiée(s). Pour cela, des fenêtres de 9 résidus, incluant obligatoirement l'acide aminé catalytique, ont été découpées. Notons que chaque peptide doit remplir les conditions suivantes :

- Taille : 15 résidus dont 9 interagissent avec le CMH (analyse Tepitope)
- Présence de l'acide aminé catalytique
- Reconnu par la molécule de CMH classe 2 (séquence de 9 résidus)

-Orientation de l'acide aminé catalytique en direction du TCR (puisque c'est ce résidu qui fixe l'antibiotique qui sera reconnu par le TCR).

Rappelons que la fixation de la molécule chimique sur la protéine est covalente et se fera préférentiellement sur une lysine ou une sérine.

Notre requête SuMo nous donne une liste de protéines qui présentent un site plus ou moins proche du site de fixation des pénicillines de la PFP. Nous nous sommes focalisés sur les cinq premières entrées. On parle de signature pour désigner l'ensemble des objets SuMo placés sur une protéine. Ainsi, l'analyse des différentes signatures nous permet de repérer les objets communs avec ceux de la référence (site de la PFP), nous fournissant donc l'acide aminé catalytique. La capture d'écran présentée sur la figure 40 présente les résultats que nous obtenons.

Avant d'expliquer le résultat final, analysons plus en détail ce que nous appelons la signature de la « query », c'est-à-dire du site de fixation des pénicillines de la PFP. La figure 39 détaille la première ligne de notre résultat :

SumoSignature																
acyl GLU 334 A	bb GLY 336 A	bb SER 337 A	hydroxyl SER 337 A	ami_ali LYS 340 A	hyd_ali GLU 334 A	hyd_aro TRP 374 A	ami_aro TRP 374 A	hydroxyl SER 395 A	amide ASN 397 A	hyd_ali GLN 452 A	hyd_ali GLN 452 A	bb GLY 549 A	bb THR 550 A	hyd_ali ALA 551 A	hydroxyl THR 526 A	ami_ali LYS 547 A

Figure 39 : Signature SuMo du site de référence (PFP)

La signature de la référence est constituée de 17 objets SuMo répartis dans l'espace sur les acides aminés suivants : Glu 334 (objets acyl et hydrophobe aliphatique), Gly 336 (objet bb encore appelé carbone alpha), Ser 337 (objets hydroxyl et carbone alpha), Lys 340 (objet amine aliphatique), Trp 374 (objets hydrophobe aromatique et amine aromatique), Ser 395 (objet hydroxyl), Asn 397 (objet amide), Gln (2 objets hydrophobe aliphatique), Gly 549 (objet carbone alpha), Thr 550 (objet carbone alpha), Ala 551 (objet hydrophobe aliphatique), Thr 526 (objet hydroxyl) et Lys 547 (objet amine aliphatique).

Ensuite, le logiciel SuMo superpose les objets de chacun des modèles de la base sur cette référence et donne une liste de protéines classées en fonction de leur score. Nous présentons ci-dessous les meilleurs résultats. Sur la figure 40, la colonne 1 donne l'identifiant PDB ou le numéro du modèle lorsque la protéine a été modélisée par Modeome. Les colonnes 2 et 3 donnent le nombre d'objets communs avec la référence et le score (qui permet de classer les protéines). Le nom de la protéine apparaît dans la quatrième colonne. Enfin, vient la « signature SuMo ». Elle est constituée de différents objets positionnés sur les protéines au niveau des sites proches du site de référence. Chaque type d'objet est représenté par une couleur et on sait sur quel acide aminé il est placé. Pour chaque protéine, nous avons coloré la case lorsqu'un de ses objets est superposé avec un objet de la query. Par exemple, pour la Human Diméthylglycine dehydrogénase mitochondriale, première de la liste, on constate qu'elle possède 5 objets communs : hydroxyl (au niveau de sa sérine 695, qui d'ailleurs est un acide aminé catalytique), amine aliphatique (sur sa lysine 481, aussi potentiellement catalytique), amide, et deux hydrophobe aliphatique. Nous avons représenté en gras les acides aminés catalytiques.

JOB Query : IQMF optimisée site CES				SumoSignature																	
PDB Id	Object	Score	Hit Information	aejl_GLU 334 A	bb_GLY 336 A	bb_SER 337 A	hydroxyl_SER 337 A	ami_ali_LYS 340 A	hyd_ali_GLU 334 A	hyd_aro_TRP 374 A	ami_aro_TRP 374 A	hydroxyl_SER 395 A	amide_ASN 397 A	hyd_ali_GLN 452 A	hyd_ali_GLN 452 A	bb_GLY 549 A	bb_THR 550 A	hyd_ali_ALA 551 A	hydroxyl_THR 526 A	ami_ali_LYS 547 A	
851	5	3	<b>2GD_HUMAN Dimethylglycine dehydrogenase mitochondrial precursor</b>				<b>695</b>	<b>481</b>													
851	4	3	2GD_HUMAN Dimethylglycine dehydrogenase mitochondrial precursor				<b>410</b>	<b>107</b>													
2J2P	6	3	L-FICOLIN COMPLEXED TO N-ACETYL-CYSTEINE (150MM)				<b>91(18)</b>														
2J3G	6	3	<b>L-FICOLIN</b>				<b>91(18)</b>														
1155	4	3	<b>56G89_HUMAN Serum albumin - Homo sapiens (Human)</b>					<b>130</b>													
2J2P	5	2	L-FICOLIN COMPLEXED TO N-ACETYL-CYSTEINE (150MM)				<b>91(18)</b>														
900	3	2	<b>RP5_HUMAN Multidrug resistance-associated protein 5 - Homo sapiens (Human)</b>					<b>1040</b>													
387	3	2	<b>O4A_HUMAN Complement C4-A precursor - Homo sapiens (Human)</b>					<b>375</b>													
215	3	2	T12A_HUMAN Potassium-transporting ATPase alpha chain 2 - Homo sapiens (Human)					<b>744</b>													
170	3	2	P2A2_HUMAN AP-2 complex subunit alpha-2 - Homo sapiens (Human)					<b>825</b>													
854	3	2	3K12_HUMAN Mitogen-activated protein kinase kinase kinase 12					<b>238</b>													
388	3	2	O4B_HUMAN Complement C4-B precursor - Homo sapiens (Human)					<b>375</b>													
1403	3	2	YTL4_HUMAN Synaptotagmin-like protein 4 - Homo sapiens (Human)					<b>410</b>													
223	4	2	TP4A_HUMAN Potassium-transporting ATPase alpha chain 1 - Homo sapiens (Human)					<b>709</b>													
1097	3	2	TPFU_HUMAN Receptor-type tyrosine-protein phosphatase U precursor					<b>1403</b>													
215	3	2	T12A_HUMAN Potassium-transporting ATPase alpha chain 2 - Homo sapiens (Human)					<b>744</b>													
170	3	2	P2A2_HUMAN AP-2 complex subunit alpha-2 - Homo sapiens (Human)					<b>825</b>													
1196	4	2	7Z4I9_HUMAN ATPase Na <sup>+</sup> /K <sup>+</sup> transporting alpha 4 - Homo sapiens (Human)				374														
223	4	2	TP4A_HUMAN Potassium-transporting ATPase alpha chain 1 - Homo sapiens (Human)				344														
854	3	2	3K12_HUMAN Mitogen-activated protein kinase kinase kinase 12					<b>238</b>													
608	3	2	NA13_HUMAN Guanine nucleotide-binding protein alpha-13 subunit					<b>299</b>					<b>250</b>								
1039	3	2	IGR_HUMAN Polymeric immunoglobulin receptor precursor - Homo sapiens (Human)					<b>376</b>					<b>351</b>								
1514	3	2	ATA_HUMAN Vacuolar ATP synthase catalytic subunit A - Homo sapiens (Human)										<b>583</b>								<b>585</b>
917	3	2	Y01C_HUMAN Myosin-1c - Homo sapiens (Human)																		<b>397</b>
2J3U	5	2	L-FICOLIN COMPLEXED TO GALACTOSE				<b>91(18)</b>														<b>526</b>
2J0Y	5	2	L-FICOLIN COMPLEXED TO B-13-D-GLUCAN				<b>91(18)</b>														

Figure 40 : Résultats de la comparaison du site de fixation des pénicillines de la PFP avec les modèles de la Serum Databank : recherche de protéines d'hapténisation.

Si l'on regarde l'entrée numéro quatre, on voit que l'on retrouve l'albumine sérique humaine. Ce résultat confirme les capacités de cette protéine à être impliquée dans le transport des drogues et à être hapténisée par une molécule de bêta-lactamine. Cependant, il faut noter que l'acide aminé catalytique fournit par l'analyse SuMo est la lysine 130. Or, dans le chapitre 4, nos études de docking et de QM/MM nous ont orientés sur d'autres lysines et notamment les 190 et 212. La raison pour laquelle l'étude de docking ne retrouve pas la lysine 130 est liée à l'accessibilité de ce résidu. En effet, nous n'avons analysé par docking que les acides aminés dont l'accessibilité avaient une valeur supérieure à 30%. Or, la lysine 130 est trop enfouie et son accessibilité au solvant n'est que de 8%. Il faudrait donc réduire le seuil que nous avons imposé à Surfex-Dock afin d'étudier l'énergie d'interaction entre le médicament et la lysine 130. Nous lancerons cette analyse dans un futur très proche.

A partir de ces résultats, a été généré un certain nombre de séquences peptiques de 9 acides aminés qui ont été analysées par le Pr. Bernard Maillère grâce au logiciel Tepitope. C'est un logiciel de prédiction de reconnaissance de peptides par différentes catégories de CMH de classe 2. L'équipe du Pr. Pallardy travaille notamment avec les CMH II suivants: DRB1\*0101, DRB1\*0301, DRB1\*0401, DRB1\*0701, DRB1\*1101, DRB1\*1301, DRB1\*1501, DRB5\*0101, DRB3\*0101, DRB4\*0101, DPB1\*0401, et DPB1\*0402. Il fallait donc que nos séquences peptidiques puissent être reconnues par tous ces sous types de CMH. Ainsi, nous obtenons un ou plusieurs peptides pour chaque protéine candidate. D'autre part, il nous fallait une taille totale de peptide de 15 résidus. Nous avons analysé sur modèles théoriques des interactions mises en jeu entre le CMH et le peptide et avons conclu que la meilleure solution était d'encadrer le nonamère par 3 résidus avant et 3 après. Nous présentons les séquences trouvées dans le tableau 16. En rouge sont indiqués les principaux acides aminés d'ancrage aux molécules HLA, en bleu les résidus d'interaction avec la pénicilline et en vert, ceux permettant d'avoir une séquence de 15 résidus.

<b>Human Dimethylglycine dehydrogenase :</b> (acide aminé catalytique : <b>lysine 481</b> )	SGL <b>Y</b> QRLES <b>K</b> CSMG <b>F</b> YQRLES <b>K</b> CSMG <b>F</b> HAG
<b>Human Ficolin 2 precursor :</b> (acide aminé catalytique : <b>sérine 18</b> )	AAT <b>L</b> LLS <b>F</b> LGMA <b>W</b> AL VGV <b>L</b> GAATLLLS <b>F</b> LG AT <b>L</b> LLS <b>F</b> LGMA <b>W</b> ALQ
<b>Human serum albumin :</b> (acide aminé catalytique : <b>lysine 130</b> )	Pas de peptide
<b>Human multidrug resistance associated proteine 5 :</b> (acide aminé catalytique : <b>lysine 1040</b> )	IRE <b>L</b> KRLDNITQ <b>S</b> PF VSR <b>V</b> LIREL <b>K</b> RLDNI SR <b>V</b> LIREL <b>K</b> RLDNIT R <b>V</b> LIREL <b>K</b> RLDNITQ

Tableau 16 : Neuf peptides issus des protéines candidates analysées par SuMo et reconnus par les molécules de CMH classe II

Il faut noter qu'aucun peptide issu de l'albumine sérique humaine et qui possède le résidu catalytique K130 n'est reconnu par les molécules CMH de classe 2. Ceci signifie que même si le médicament peut se fixer covalamment sur la lysine 130, un peptide hapténisé au niveau de ce résidu ne pourra pas être reconnu par une molécule de CMH II et donc il n'y aura pas activation des LTCD4.

Nous avons appliqué la même analyse par Tepitope autour des acides aminés catalytiques trouvés par notre approche de docking. Les peptides hapténisés en positions 190, 212 et 432 sont reconnus par le CMH classe 2. Pour la lysine 190, le nonamère hapténisé est : LRDEG**K**ASS. Pour la lysine 212, il a cette séquence : F**K**AWAVARL. Enfin, pour la lysine 432, le peptide hapténisé est : VEVSRNLG**K**.



La dernière condition concerne l'acide aminé catalytique qui doit obligatoirement être orienté vers le TCR afin que le médicament lui soit présenté. Pour cela, nous avons étudié les interactions de plusieurs structures de CMH classe 2 co-résolues avec leur peptide. Nous avons analysé douze structures différentes afin d'avoir une analyse moyenne des interactions possibles entre peptide et CMH. Chaque structure est analysée grâce au logiciel Sybyl 8.03 (Tripos Inc.). Le tableau 17 présente les structures étudiées en terme d'interaction CMH-Peptide. Les lettres rouges indiquent les acides aminés qui sont dirigés vers le CMH, en rouge et en gras nous avons le résidu d'ancrage principal, en orange sont représentés les acides aminés qui n'ont pas une orientation nette ni vers le TCR ni vers le CMH, en noir les résidus dirigés vers le TCR, enfin en gris ceux qui sortent de la poche et ne présentent donc pas d'intérêt.

Nom	Code PDB	Séquence du peptide
HLA-DR1 (DRA, DRB1-1)	1T5W	A A <b>Y</b> S D Q A T P L L L S P R
HLA-DR1 (DRB1*0101)	2FSE	A G <b>F</b> K G E Q G P K G E P G
HLA-DR4 (DRA*0101, DRB1*0401)	2SEB	A Y <b>M</b> R A D A A A G G A
HLA-DR2 (DRA*0101, DRB1*1501)	1BX2	E N P V <b>V</b> H F F K N I V T P R
HLA-DR1 (DRA, DRB1-1)	2IAM	<b>G</b> E L I G I L N A A K V P A D
DR1 (DRA, DRB1-1)	2IAN	G E L I G I T L N A A K V P A D
HLA-DRA1*0101/DRB5*0101	1H15	G G V <b>Y</b> H F V K K H V H E S
HLA-DR1 (DRA*0101, DRB1*0101)	1SJH	P E <b>V</b> I P M F S A L S E G
HLA-DR1 (DRA*0101, DRB1*0101)	1SJE	P E <b>V</b> I P M F S A L S E G A T
DR1 (DRA, DRB1-1)	2IPK	P K <b>W</b> V K Q N T L K L A T
HLA-DR1	1DLH	P K <b>Y</b> V K Q N T L K L A T
HLA-DR1 (DRA, DRB1 0101) HLA-A2	1AQD	V G S D <b>W</b> R F L R G Y H Q Y A

Tableau 17 : Orientation des chaînes latérales des résidus peptidiques au sein de la niche à peptide

Nous constatons que les acides aminés dirigés vers le TCR et qui sont donc potentiellement aptes à présenter le médicament au TCR, sont placés à des endroits clés de la séquence peptidique. En effet, si l'on considère que l'on compte à partir du principal résidu d'ancrage au CMH, on obtient des acides aminés orientés vers le TCR (en noir) aux positions 2, 3, 5 et 8. Parfois, à ces positions, la chaîne latérale du résidu n'est pas plus orientée vers le TCR que vers le CMH (en orange), et de ce fait, on peut dire que la position est neutre. Quoiqu'il en soit, les acides aminés 1, 4, 6 et 7 ne peuvent pas servir à présenter le médicament au TCR.

Ainsi, des ces analyses moléculaires, nous avons constaté que la dernière condition que les peptides doivent remplir est de posséder un acide aminé de fixation qui soit en position 2 ou 3 ou 5 ou 8.

Sur l'ensemble des peptides issus de notre comparaison SuMo, nous avons donc dans un premier temps sélectionné 9 peptides reconnus par le système HLA de classe 2, et enfin, sur ces 9 candidats, seuls trois d'entre eux possèdent un acide aminé catalytique en position 2, 3, 5 ou 8. Les trois candidats sont présentés ci-dessous, avec en vert, les résidus encadrant les acides aminés qui interagissent réellement avec le CMH, en rouge le résidu d'ancrage principal et en bleu l'acide aminé catalytique:

Pour la Human Ficolin 2 precursor :

A T L L S F L G M A W A L Q

Pour la Multidrug resistance associated protein 5 :

I R E L K R L D N I T Q S P F

et

R V L I R E L K R L D N I T

Ces peptides sont en cours de synthèse, puis ils seront couplés à la pénicilline (et plus tard à l'amoxicilline) afin de stimuler des LT.

En parallèle, le Pr. B. Maillère lancera des expériences de reconnaissance de ces peptides par les molécules de CMH classe 2 au moyen de tests cellulaires que nous ne détaillerons pas dans ce travail.

### III- Perspectives

Les validations expérimentales constituent pour nous modélisateurs, un point crucial sans lequel notre travail ne pourrait ni être évalué, ni évoluer. En effet, lorsque nous émettons des hypothèses à travers nos modèles, nous attendons en retour des résultats expérimentaux qui pourront nous guider, soit sur une autre voie, si nos modèles ne sont pas corrélés avec les expériences, soit dans la même voie si nos modèles sont pertinents au vu des manipulations *in vitro* et/ou *in vivo*. Le but est de faire évoluer nos modèles afin de se rapprocher au mieux de la réalité physiologique et de mieux comprendre les phénomènes mis en jeu.

Concernant les différents scénarios présentés dans ce chapitre de thèse, nous envisageons de poursuivre notre collaboration avec l'équipe du Pr. J.F. Nicolas et nous appuyons notamment sur l'aide du Dr. A. Rozières qui pourra nous apporter en partie les idées et les validations expérimentales nécessaires à l'évaluation de la pertinence de nos modèles. Une première étape consiste à bloquer le cycle bêta-lactame afin d'empêcher la réaction d'ouverture du cycle nécessaire à la fixation covalente du médicament sur sa cible protéique. Dans ce cas, nous inhiberions l'hapténisation. Il faudra alors vérifier si le médicament est toujours responsable de l'activation spécifique des lymphocytes T.

Les expériences qui ont été réalisées ou qui sont en cours actuellement ne concernent pas l'amoxicilline, ni d'ailleurs la ceftriaxone, qui sont les deux molécules que nous avons choisies pour nos études théoriques de docking et de QM / MM. Aussi, la prochaine étape sera de valider nos hypothèses avec l'amoxicilline. Avec l'aide de l'équipe du Pr. Pallardy, nous reproduirons la méthodologie expérimentale employée pour la pénicilline. Nous avons d'ailleurs déjà débuté ces validations et les résultats sont en cours.

D'autre part, pour le moment, plusieurs peptides ont été synthétisés. Ils proviennent des protéines sériques humaines que nous avons détectées par notre recherche SuMo. Nous appliquerons la même démarche pour d'autres peptides. Notamment, les peptides qui correspondent au site de fixation du médicament que notre étude de docking a retrouvé. La littérature décrit également d'autres séquences peptidiques sur l'albumine humaine et qui présentent l'acide aminé catalytique nécessaire à l'hapténisation. Les séquences ont été analysées par le Pr B. Maillère grâce à Tepitope et sont reconnues par les molécules de CMH classe2. Ainsi, nous synthétiserons les peptides (issus de l'albumine humaine) suivants :

Pour nos résultats de docking (pour la fixation de l'amoxicilline) :

- Site K 190 : LRDEG**K**ASS
- Site K212 : F**K**AWAVARL
- Site 432 : VEVS**R**NLG**K**

Pour les résultats issus de la littérature (pour la fixation de la pénicilline) :

- Site K190 : LRDEG**K**ASS
- Site S193 : pas de peptide reconnu
- Site K195 : pas de peptide reconnu
- Site K 199 : pas de peptide reconnu
- Site K212 : F**K**AWAVARL
- Site K 432 : VEVS**R**NLG**K**
- Site K541 : pas de peptide reconnu
- Site K 545 : pas de peptide reconnu
- Site K 525 (pour la fixation de la pipéracilline) : pas de peptide reconnu
- Site K 534 (pour la fixation de la pipéracilline) : LV**K**HKPKAT ;  
K**H**KPKATK ; LVELV**K**H**K**P ; VELV**K**H**K**PK

Par la suite, des molécules de pénicilline ou d'amoxicilline seront fixées et seront alors réalisés les tests de stimulation de LT et de reconnaissance cellulaire par les molécules de CMH.

L'ensemble des peptides conjugués à une molécule médicamenteuse pourrait servir à créer une puce, par la suite utilisée dans le diagnostic des réactions d'hypersensibilités allergiques aux bêta-lactamines.

A plus long terme, il sera extrêmement intéressant de résoudre expérimentalement, par une approche cristallographique, le super complexe généré au cours de la reconnaissance spécifique de la drogue par le LT et impliquant le CMH, la drogue, le peptide, l'ion éventuel et le TCR spécifique de la drogue.

# Conclusion générale

La modélisation moléculaire est une technique qui a énormément évolué ces dernières années et dont les développements actuels et futurs sont très prometteurs.

Les outils de modélisation permettent de simuler *in silico* de nombreuses réactions chimiques, d'étudier les interactions au niveau moléculaire entre différents partenaires, d'explorer des systèmes biologiques complexes... Les applications sont donc très vastes. Toutefois, il faut garder à l'esprit qu'en tant que modélisateurs, nous ne réalisons que des analyses théoriques qui nécessitent des validations expérimentales tout au long des études menées. Aussi devons nous travailler en étroite collaboration avec nos partenaires scientifiques afin de créer le modèle le plus proche de la réalité biologique.

Ce travail de thèse s'est intéressé aux réactions d'hypersensibilités allergiques liées aux antibiotiques de la famille des bêta-lactamines. Nous avons, dans un premier temps, élaboré un dendrogramme permettant de classer ces molécules en fonction de leurs champs moléculaires. Ce dendrogramme est utilisé par les médecins et scientifiques avec lesquels nous avons collaboré. D'autre part, notre analyse de 3D-QSAR révèle des zones plus ou moins impliquées dans les réactions d'HSA Immédiates.

Dans un second temps, nous avons analysé la réactivité des bêta-lactamines. Nous concluons que les pénicillines sont plus réactives que les céphalosporines et qu'elles se fixeront préférentiellement à des lysines. Par la suite, nous avons étudié l'albumine sérique humaine sur laquelle nous avons identifié les acides aminés susceptibles d'assurer la liaison covalente décrite lors des phénomènes d'hapténisation. Notre logiciel SuMo a retrouvé d'autres protéines qui pourraient également porter le médicament.

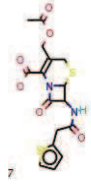
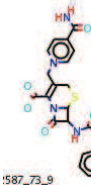
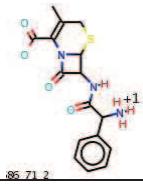
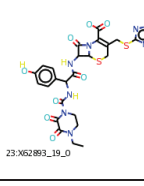
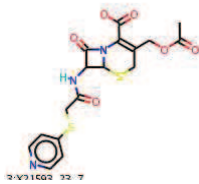
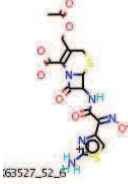
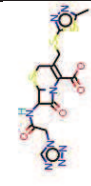
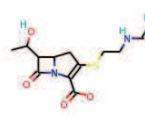
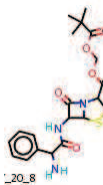
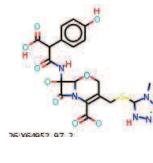
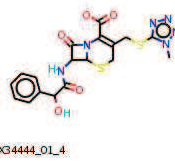
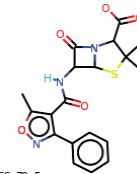
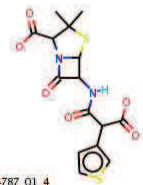

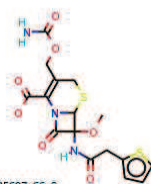
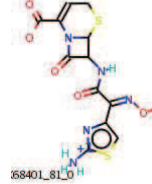
Nous avons, dans une autre phase, analysé la reconnaissance moléculaire de l'allergène par les IgE dans le cas des HSA Immédiates. Nous pensons que cette reconnaissance se fera de manière plane et impliquera un Fab en interaction spécifique avec une molécule d'amoxicilline portée par une protéine telle que l'albumine.

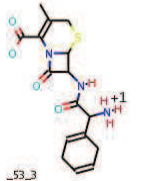
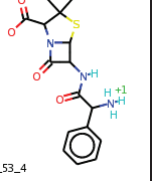
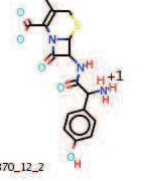
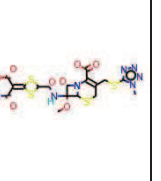
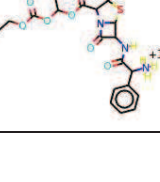
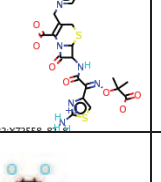
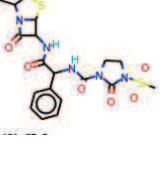
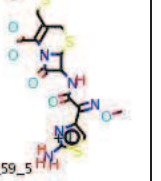
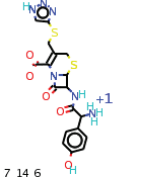
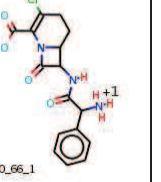
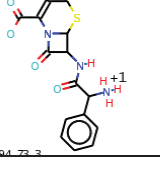
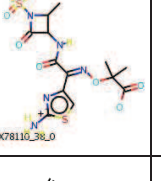
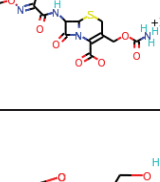
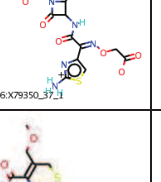
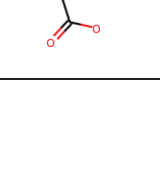
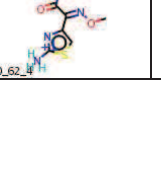
Enfin, concernant le versant retardé des phénomènes d'HSA, nous avons étudié plusieurs complexes mettant en jeu le CMH, le peptide hapténisé par l'amoxicilline et le TCR. Nous pensons que plusieurs mécanismes peuvent intervenir : l'hapténisation, le p-i concept et le new p-i concept.

Cette thèse a été menée en collaboration avec plusieurs équipes scientifiques qui nous ont apporté les données cliniques et les validations expérimentales de nos modèles. Les perspectives de ce travail visent à poursuivre les analyses concernant la reconnaissance de l'antibiotique par les IgE et les TCR. Nous attendons pour cela que des complexes entre amoxicilline et IgE ou TCR spécifiques soient résolus et déposés à la PDB. D'autre part, nous attendons de nombreux résultats de manipulations qui nous donneront le nombre et la position des molécules d'amoxicilline sur une protéine cible. Enfin, nous continuerons à améliorer la sensibilité des tests de diagnostic des hypersensibilités allergiques.

L'ensemble des travaux ici présentés pourra être étendu et adapté aux hypersensibilités allergiques liées à d'autres classes thérapeutiques.

## ANNEXE

N°CAS	Nom	Structure	Log P	N°CAS	Nom	Structure	Log P
153-61-7	cefalotine		-0,65	62587-73-9	cefsulodine		-1.23
15686-71-2	cefalexine		0,09	62893-19-0	cefoperazone		-0,58
21593-23-7	cefapirine		-0,98	63527-52-6	cefotaxime		-1,31
25953-19-9	cefazoline		-1,55	64221-86-9	imipeneme		-1,67
33817-20-8	pivampicilline		2,51	64952-97-2	moxalactam		-0,16
34444-01-4	cefamandole		-0,25	66-79-5	oxacilline		1,69
34787-01-4	ticarcilline		0,51	68373-14-8	Sulbactam		-0,03
35607-66-0	cefoxitine		-1,18	68401-81-0	ceftizoxime		-1,11

38821-53-3	cefradine		0,20	69-53-4	ampicilline		0,74
50370-12-2	cefadroxil		-0,57	69712-56-7	cefotetan		-2,35
50972-17-3	bacampicilline		2,09	72558-82-8	ceftazidime		-1,96
51481-65-3	mezlocilline		1,14	73384-59-5	ceftriaxone		-1,44
51627-14-6	cefatrizine		-1,15	76470-66-1	loracarbef		1,46
53994-73-3	cefaclor		0,29	78110-38-0	Aztreonam		-1,09
55268-75-2	cefuroxime		0,08	79350-37-1	cefixime		-1,54
58001-44-8	Acide clavulanique		-1,41	80210-62-4	cefepodoxime		-1,87



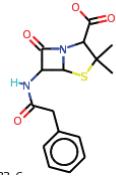
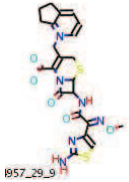
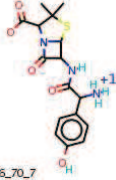
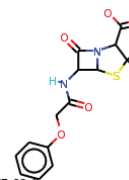
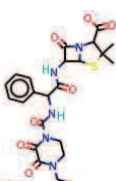
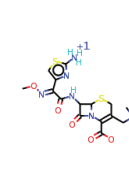
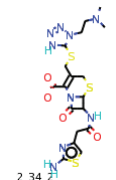
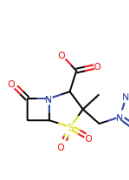
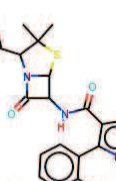
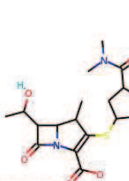
61-33-6	pénicilline G		1,39	84957-29-9	cefpirome		1,38
61336-70-7	amoxicilline		0,08	87-08-1	pénicilline V		1,59
61477-96-1	pipéracilline		1,34	88040-23-7	cefepime		-4,51
61622-34-2	cefotiam		-1,19	89786-04-9	tazobactam		-1,19
61-72-3	cloxacilline		2,16	96036-03-2	Meropenem		-1,33

Tableau 18 : Quarante deux structures de bêta-lactamines

En rouge sont représentées les pénicillines, en bleu les céphalosporines, en vert les carbapénèmes, en noir la monobactame, et en rose les inhibiteurs de bêta-lactamases

## **BIBLIOGRAPHIE**

Antunez C., Blanca-Lopez N., Torres M.J. et al., "Immediate allergic reactions to cephalosporins: evaluation of cross-reactivity with a panel of penicillins and cephalosporins", *J Allergy Clin Immunol*, **2006**, 117:404-10.

Antunez C., Martin E., Cornejo-Garcia J.A., et al., "Immediate hypersensitivity reactions to penicillins and other betalactams", *Curr Pharm De*, **2006**, 26:3327-33.

Apaydin R., Bilen N., Dökmeci S. et al., "Drug eruptions: a study including all inpatients and outpatients at a dermatology clinic of a university hospital", *J Eur Acad Dermatol Venereol*, **2000**, 14(6): 518-20

Baldo B.A., "Diagnosis of allergy to penicillins and cephalosporins – structural and immunochemical considerations", *ACI International*, **2000**, 12-5

Banerjee R., Das K., Ravishankar R., et al. "Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex", *J Mol Biol*, **1996**, 259(2):281-96

Bates P.A., Kelley L.A., MacCallum R.M., et al., "Enhancement of protein modeling by human intervention in applying the automatic programs 3DJIGSAW and 3DPSSM", *Proteins*, **2001**, 45 (Suppl. 5):39–46

Becke A.D., Salahub D., Zerner M., "In the challenge of d and f electrons", *ACS symp Series 394, American Chemical Society, Washington*, **1989**, 165

Becke, A. D.; Salahub D. ; Zerner, M., (1989), In the challenge of d and f electrons, *ACS symp. Series 394, American Chemical Society, Washington*, 165.

Benveniste J., "Hypersensibilité immédiate", l'Allergologie, 2eme edition, Charpin J. Flammarion, **1986**

- Bian H., Hammer J., « Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE », *Methods.*, **2004**, 34(4):468-75
- Bigby M., “Rates of cutaneous reactions to drugs”, *Arch Dermatol*, **2001**, 137 (6): 765-70
- Birnbaum J., Verloet D., “Allergie aux pénicillines”, *Rev Fr Allergol*, **1997**, 37 (1):29–35
- Bohacek R. S., McMartin C., "Definition and Display of Steric, Hydrophobic, and Hydrogen-Bonding Properties of Ligand Binding Sites in Proteins Using Lee and Richards Accessible Surfaces: Validation of a High-Resolution Tool for Drug Design", *J Med Chem*, **1992**, 35:1671-84
- Bohm H.J., “LUDI: Rule-based automatic design of new substituents for enzyme inhibitor leads”, *J. Comput. Aided Molec. Des.*, **1992**, 6 , 593-606
- Bohm H.J., Klebe G., “What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs?”, *Angew. Chem. Int. Ed. Engl.*, **1996**, 35, 2588-2614
- Bolland S., Pearse R.N. Kurosaki T. et al., “SHIP modulates immune receptor responses by regulating membrane association of Btk”, *Immunity*, **1998**, 8(4): 509-16
- Bordner A.J., Mittelman H.D., « MultiRTA: a simple yet reliable method for predicting peptide binding affinities for multiple class II MHC allotypes », *BMC Bioinformatics*. **2010**, 11:482
- Bousquet P.J., Kvedariene V., Co-Minh H.B., et al., “Clinical presentation and time course in hypersensitivity reactions to beta-lactams”, *Allergy*, **2007**, 62:872-876.
- Bousquet J., Breitenbach M., Dreborg S. et al., “Diagnosis of allergy and specific immunotherapy using recombinant allergens and epitopes” *Mol. Biol. Immunol. Allergens*, **1993**, 311-20.
- Bower M.J., Cohen F.E., Dunbrack R.L., “Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool”, *J Mol Biol*, **1997**, 267(5):1268-82
- Bradley J.S., Garau J., Lode H. et al. “Carbapenems in clinical practice: a guide to their use in serious infection”, *Int J Antimicrob Agents*, **1999**, 11(2):93-100

Britschgi M., Steiner U.C., Schmid S. et al., «T-cell involvement in drug-induced acute generalized exanthematous pustulosis», *J.Clin Invest*, **2001**, 107(11): 1433-41

Brooks., “CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations.”, *J Comp Chem*, 1983, 4:187-217

Brown J.R., “Structural origins of mammalian albumin”, *Fed. Proc.*, **1976**, 35: 2141-4

Brown J. R., Shockley P., “Serum albumin: structure and characterization of its ligand binding sites”, *Lipid-Protein Interact.*, **1982**, voll1, 25-68.

Bryskier A., «Histoire des antibiotiques», InBergogne-berezine E., Dellamonica P., « Antibiothérapie en pratique clinique », *Masson*, Paris, **1996**, 1-16

Case D.A., Pearlman D.A., Caldwell J.W., et al., « AMBER 9 », University of California, San Francisco, **1997**.

Castelli F.A., Houitte D., Munier G. et al., « Immunoprevalence of the CD41 T-cell response to HIV Tat and Vpr proteins is provided by clustered and disperse epitopes, respectively », *Eur J Immunol*, **2008**, 38: 2821–2831

Chan V.W., Lowell C.A., Defranco A.L., “Defective negative regulation of antigen receptor signaling in Lyn-deficient B lymphocytes”, *Curr Biol*, **1998**, 8(10):545-53

Chemelle J.A., Terreux R., Rozières A., Nicolas J.F. Abstract of congress, E.R.G.E.C.D., Lyon, **2006**

Chipot C., Kollman P.A., “Alternative approaches to potential of mean force calculations: free energy perturbation versus thermodynamic integration. Case study of some representative nonpolar interactions”, *J Comput Chem*, **1996**, 17:1112-31

Chothia C., Lesk A.M., “The relation between the divergence of sequence and structure in proteins”, *The EMBO Journal*, **1986**, 5(4):823

Cochlovius B, Stassar M, Christ O, et al. , “In vitro and in vivo induction of a Th cell response toward peptides of the melanoma-associated glycoprotein 100 protein selected by the TEPITOPE program », *J Immunol*, **2000**, 165(8):4731-41

Combet C., Jambon M., Deleage G. et al., "Geno3D: automatic comparative molecular modelling of protein", *Bioinformatics*, **2002**, 18(1):2134

Cone R.E., Li X., Sharafieh R. et al., « The suppression of delayed-type hypersensitivity by CD8<sup>+</sup> regulatory cells requires interferon-gamma », *Immunology* , **2007**, 120(1): 112-9

Cornell W.D., Cieplack P., Bayly C.I., Gould I.R., et al., "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules", *J Am Soc*, **1995**, 117:5179-5197

Cornell W.D., Cieplack P., Bayly C.I. et al., "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules", *J. Am. Soc.*, **117**, **1995**, 5179-5197

Clark R.D., Strizhev A., Leonard J.M., et al., "Consensus scoring for ligand/protein interactions", *J Mol Graph Mod.*, **2002**, 20:281

Da Silva E., Rousseau C., Zanella-Cleon I., "Mass Spectrometric Determination of Association Constants of Bovine Serum Albumin (BSA) with para-Sulfonato-Calix[n]arene Derivatives", *Journal of Inclusion Phenomena and Macrocyclic Chemistry* , **2006**, 54(1-2), 53-59

David B. « De la synthèse des IgE aux syndromes allergiques ». *Rev Prat*, **1996**, 46 :935-941

Demoly P., Hillaire-Buys D., Raison-Peyron N., et al. « Identifier et comprendre les allergies médicamenteuses », *Med Sci (Paris)*, **2003**, 19(3):327-36

Demoly P., Messaad D., Benahmed S. et al., "Immunoallergic reactions of drug origin: epidemiologic and clinical data", *Thérapie*, **2000**, 55(1), 13-9

Depil S., Moralès O., Castelli F.A. et al., « Determination of a HLA II promiscuous peptide cocktail as potential vaccine against EBV latency II malignancie », *J Immunother.*, **2007**, 30(2):215-26

Dewar M.S.J., Zoebish, E.G., Healy E. F.; Stewart, J.J.P., *J. Am. Chem. Soc.*, **1985**, 107:3902

Donowitz G.R. et Mandell G.L., « Beta-lactam antibiotics », *N Engl J Med* , **1988**, 318:419-26, 490-500

Dubois B.L., Chapat L., Goubier A. et al., “CD4+CD25+ T cells as key regulators of immune responses”, *Eur J Dermatol*, **2003**, 13(2): 111-6

Dubois B.L., Chapat L., Goubier A. et al., “Innate CD4+CD25+ regulatory T cells are required for oral tolerance and inhibition of CD8+ T cells mediating skin inflammation”, *Blood*, **2003**, 102(9):3295-301

Dugaiczyk A., Law S.W., Dennison O.E., “Nucleotide sequence and the encoded amino acids of human serum albumin mRNA“, *Proc Natl Acad Sci USA*, **1982**, 79:71-75

Duncia J.V., Chiu A.T., Carini D.J. et al., “The discovery of potent nonpeptide angiotensin II receptor antagonist: a new class of potent antihypertensives.”, *J. Med. Chem.*, **1990**, 33, 1312-1329

Fiser A., Sali A., “Modeller: generation and refinement of homologybased protein structure models”, *Methods Enzymol*, **2003**, 374:46191

Fish D.N., Singletary T.J., “Meropenem, a new carbapenem antibiotic”, *Pharmacotherapy*, **1997**, 17(4):644-69

Fock, V. Z., *Physik*, **1930**, 61:126.

Galmiche J.M., « La découverte de la pénicilline », *Hygiène et Médecine*, Édition Louis Pariente, **1999**, p. 453-454.

Galmiche JM, (**1999**) , « la Découverte de la pénicilline », *Hygiène et Médecine*, Édition Louis Pariente, **1999**, pages 453-454.

Geourjon C., Combet C., Blanchet C. et al., “Identification of related proteins with weak sequence identity using secondary structure information”, *Protein Sci*, **2001**, 10(4):788-97

Glen R.C., Martin G.R., Hill P.A. et al., “Computer-aided design and synthesis of 5-substituted tryptamines and their pharmacology at the 5-HT1D receptor: Discovery of compounds with potential anti-migraine properties”, *J Med Chem*, **1995**, 38:3566-80

Gomes E.R., Demoly P. "Epidemiology of hypersensitivity drug reaction" *Curr Opin Allergy Clin Immunol*, 2005, 5(4): 309-16

Gordon E., Mouz N., Duee E. et al., "The Crystal Structure of the Penicillin-binding Protein 2x from *Streptococcus pneumoniae* and its Acyl-enzyme Form: Implication in Drug Resistance", *Journal of Molecular Biology*, **2000**, 299(2):477-485.

Guglielmi L., Demoly P., « Polymorphismes génétiques et allergies médicamenteuses », *Rev Fr Allergol*, **2005**, 45 :214-217

Hahn M.A., "Receptor Surface Models. 1. Definition and Construction", *J Med Chem*, **1995**, 38:2080-90

Hahn M.A., Rogers D., "Receptor Surface Models. 2. Application to Quantitative Structure-Activity Relationships Studies", *J Med Chem*, **1995**, 38:2091-2102

Hansch C., *Accounts of Chem Res*, **1969**, 2:232

Hartree D.R., *Proc. Cambridge Phil. Soc.*, **1928**, 24:89

Hellinger W.C., Nelson S.B., "Carbapenems and monobactams: imipenem, meropenem, and aztreonam" *Mayo Clin Proc*, **1999**, 74:420-434

Hennecke J., Carfi A., Wiley D.C., "Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1", *EMBO J.*, **2000**, 19(21):5611-24

Hwang M.J., Ni X., Waldman M., Ewig, C.S., et al., "Derivation of Class II Force Fields. VI. Carbohydrate Compounds and Anomeric Effects", *Biopolymers*, **1998**, 45:435-468

Ivanciuc O., Ivanciuc T., Cabrol-Bass D. et al., "Comparison of Weighting Schemes for Molecular Graph Descriptors: Application in Quantitative Structure-Retention Relationship Models for Alkylphenols in Gas-Liquid Chromatography", *J Chem Inf Comp Sci*, **2000**, 40 (3): 732-43

Ivanciuc O., Ivanciuc T., Cabrol-Bass D. et al. "Evaluation in Quantitative Structure-Property Relationship Models of Structural Descriptors Derived from Information-Theory Operators", *J Chem Inf Comp Sci*, 2000, 40 (3): 631-43

Jambon M., « Un système bioinformatique de recherche de similitudes fonctionnelles dans les structures 3D de protéines », Thèse, Université Claude Bernard-Lyon 1, **2003**, 161 p.

Jambon M., Andrieu O., Combet C. et al., "The SuMo server: 3D search for protein functional sites", *Bioinformatics*, **2005**, 21(20):3929-30

Jambon M., Imbert A., Deléage G., et al., "A new bioinformatic approach to detect common 3D sites in protein structures", *Proteins*, **2003**, 52(2):137-45

Johansson S.G., Hourihane J.O., Bousquet J., et al. "A revised nomenclature for allergy. An EAACI position statement from the EAACI nomenclature task force", *Allergy*, **2001**, 56: 813-24.

Juchet A., Rance F., Bremont F. et al., « Exploration de l'allergie aux bêta-lactamines chez 45 enfants », *Rev Fr Allergol*, **1994**, 34 (5):369-375

Just N., Lamblin C., Perez T. et al., « Allergie aux antibiotiques au cours de la mucoviscidose », *Rev Fr Allergol*, **2002**, 42 : 454-463

Kashima A., Inoue Y., Sugio S. et al., "X-ray crystal structure of a dipeptide-chymotrypsin complex in an inhibitory interaction", *Eur J Biochem*, **1998**, 255(1):12-23

Koehl P., Delarue M., "Application of a self-consistent mean field theory to predict protein side-chains conformation and estimate their conformational entropy", *J Mol Biol*, **1994**, 239(2):249-75

Kollman P.A., "Free energy calculation Applications to chemical and biochemical phenomena", *Chem Rev*, **1993**, 93(7):2395-2417

Krieger E., Nielsen J.E., Spronk C.A.E.M. et al., "Fast empirical pKa prediction by Ewald summation", *J Mol Graph Model*, **2006**, 25(4):481-6



Kroemer R.T., Hecht P, "Replacement of steric 6-12 potential-derived interaction energies by atom-based indicator variables in CoMFA leads to models of higher consistency", *J Comput-Aided Mol. Design*, **1995**, 9:205-12

Kyachko E.S., Ludena E., "Density Functional Theory of many electron system", *Kluwer Academic Press*, Dordrech, **1990**

Lachmann P.J., "Classification of allergic reactions responsible for clinical hypersensitivity and disease, clinical aspects of immunology" Blackwell, 1975. Goldsby RA, Kindt TK, Osborne BA and Kuby J, *Immunology* 5<sup>th</sup> edition, WH Freeman and Compagny, **2003**.

Lafaye P., Lapresles C., "Fixation of penicilloyl groups to albumin and appearance of anti-penicilloyl antibodies in penicillin-treated patients", *J Clin Invest*, **1988**, 82:7-12

Landsteiner K., "The Specificity of Serological Reactions", *Dover Publications*, New York. **1962**

Lepoittevin J.P., "Metabolism versus chemical transformation or pro- versus prehapten?" *Contact Dermatitis*, **2006**, 54(2): 73-4

Levitt M., "Accurate modeling of protein conformation by automatic segment matching", *J Mol Biol*, **1992**, 226(2):507-33

Lin R.Y., "A persepective on penicillin allergy", *Arch Intern Med*, **1992**, 152: 930-937

Liyanage U.K., Moore TT, Joo H.G. et al., « Prevalence of regulatory Tcells is increased in peripheral blood and tumor microenvironment of patients with pancreas or breast adenocarcinoma", *J Immunol*, **2002**, 169(5): 2756-61

Lu L., Vollmer J., Moulon C. et al., « Components of the ligand for Ni<sup>++</sup> reactive human T cell clon ». *J Exp Med*, **2003**, 197(5): 567-74

MacKerell A.D., Bashford D., Bellott M., et al., "All-atom empirical potential for molecular modeling and dynamics studies of proteins", *J Phys Chem B*, **1998**, 102:3586-3616

MacKerell Jr., "An all-atom empirical energy function for the simulation of nucleic acids", *J Am Chem Soc*, **1995**, 117:11946-11975

Malissen B., Malissen M., “Functions of TCR and pre-TCR subunits: lessons from gene ablation”, *Curr Opin Immunol*, **1996**, 8(3):383-93

Maple J.R., Hwang M.J., Jalkanen, K.J., et al., “Derivation of Class II Force Fields: V. Quantum Force Field for Amides, Peptides, and Related Compounds”, *J Comp Chem*, **1998**, 19:430-458

Miranker A., Karplus M., “Functionality maps of binding sites: a multiple copy simultaneous search method”, *Proteins: Struct Func Gen*, 1991,11:29-34

Miyake S., Yamamura T., “NKT cells and autoimmune diseases: unraveling the complexity”, *Curr Top Microbiol Immunol*, **2007**, 314 : 251-67

Moller C., Plesset M.S., *Phys. Rev.*, **1934**, 46:418

Naisbitt D.J., Gordon S.F., Pirmohamed M., et al., « Immunological principles of adverse drugs reactions : the initiation and propagation of immune responses elicited by drug treatment », *Drug Saf*, **2000**, 23 (6) : 483-507

Nielsen M, Lund O, Buus S., et al., « MHC class II epitope predictive algorithms », *Immunology.*, **2010**, 130(3):319-28

Nigen S., Knowles S.R., Shear N.H., “Drug eruptions: approaching the diagnosis of drug-induced skin diseases”, *J Drugs Dermatol*, **2003**, 2(3): 278-99

Padovan E., Bauer T., Tonigio M.M. et al., «Penicilloyl peptides are recognized as Tcell antigenic determinants in penicillin allergy», *Eur J Immunol.*, **1997**, 27(6): 1303-7

Park K., Pirmohamed M., Kitteringham N.R., “Role of drug disposition in drug hypersensitivity: a chemical, molecular and clinical perspective”, *Chem Res Toxicol*, **1998**, 11(9): 969-88

Park M.A., Li J.T., “Diagnosis and management of penicillin allergy”, *Mayo Clin Proc*, **2005**, 80 (3):405-410.

Parr R.G., Yang W., “Density functional theory of atoms and molecules”, *Oxford University Press*, New York, **1989**

Pearlman D.A., Case D.A., Caldwell J.W., et al., "AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules", *Comp Phys Commun*, **1995**, 91:1-41.

Pearlman D.A., Kollman P.A., "The lag between the Hamiltonian and the system configuration in free energy perturbation calculations", *J Chem Phys*, **1989**, 91:7831-39

Peng Z., Ewig C.S., Hwang M.J., Waldman M., et al., "Derivation of Class II Force Fields. 4. Van der Waals Parameters of Alkali Metal Cations and Halide Anions", *J Phys Chem. A*, **1997**, 101: 7243-7252

Peng, Z., Ewig, C.S., Hwang, M.-J., Waldman, M., Hagler, A.T, (**1997**), Derivation of Class II Force Fields. 4. Van der Waals Parameters of Alkali Metal Cations and Halide Anions. *J. Phys. Chem. A* , 101, 7243-7252,.

Perdew J.P., Burke K., Wang Y., « Generalized gradient approximation for the exchange-correlation hole of a many-electron system », *Phys. Rev. B*, **1996**, 54 : 16533-39

Petrey D., Xiang Z., Tang C.L. et al., « Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modelling », *Proteins*, **2003**, 53 (Suppl. 6):430–5

Pichler W.J. « Pharmacological interaction of drugs with antigen-specific immune receptors : the p-i concept », *Curr Opin Allergy Clin Immunol*, **2002**, 2 (4) : 301-5

Pichler W.J., « Modes of presentation of chemical neoantigens to the immune system » *Toxicology*, **2002**, 181-182: 49-54

Pichler W.J., Beeler A., Keller M., Lerch M. et al., « Pharmacological interaction of drugs with immune receptors: the p-i concept ». *Allergol Int.*, **2006** , 55(1):17-25

Pichler W.J., Schnyder B., Zanni M.P. et al., «Role of T cells in drug allergies», *Allergy*, **1998**, 53(3): 225-32

Pirson F., « Prise en charge des allergies médicamenteuses », *Congrès de médecine générale, Bruxelles , Belgique*, **2005**, 124, 9, 234-244.

- Ponvert C., Le Bourgeois M., R. Scheinmann et al., « Les réactions allergiques aux médicaments anti-infectieux courants », *Rev Fr Allergol*, **1999**, 39 (6) :455-467
- Ponvert C., Mazeirat R., Scheinmann R., Paupe J. et al., « Exploration allergologique des réactions aux bêta-lactamines chez l'enfant : résultats d'une étude de 50 cas, comportant des tests cutanés lecture immédiate, des RAST et des tests de réintroduction par voie orale », *Rev Fr Allergol*, **1996**, 36 (2) :162-175
- Ponvert C., Scheinmann P., “Les réactions allergiques et pseudo-allergiques aux bêta-lactamines”, *Arch Ped*, **2003**, 10 (7):658–665
- Pouyane P., F. Haramburu et al., « Admissions to hospital caused by adverse drug reaction: cross sectional incidence study. French Pharmacovigilance Centres», *BMJ*, **2000**,320 (7241): 1036
- Reig Rincon de Arellano I., Villalon Garcia A.L., Cimarra Alvarez-Lovell M., et al., “Flare up to betalactams” *Allergol Immunopathol (Madr)*, **2005**, 33, 282-4
- Romano A., Quaratino D., Papa G., et al., “Aminopenicillin allergy”, *Arch Dis Child*, **1997**, 76:513-517.
- Romano. A., “Delayed hypersensitivity to aminopenicillins”, *Clin Exp Allergy*, **1998**, suppl 4:29-32
- Roncarolo M.G., Baccheta C., Bordignon C., et al., “Type 1 T regulatory cells”, *Immunol Rev*, **2001**, 182 :68-79.
- rotamers from a backbonedependent rotamer library: a new homology modeling tool”, *Journal of Molecular Biology*, **1997**, 267(5):126882
- Rozieres A., Hennino A., Rodet K., et al., “Detection and quantification of drug-specific T cells in penicillin allergy”, *Allergy*, **2009**, 64(4):534-42
- Rudolph A.H., Price E.V., “Penicillin reactions among patients in venereal diseases clinics: a national survey”, *JAMA*, **1973**, 223: 99

Sanz M.L., Garcia M.C., Caballero M.R., et al. “Basophil activation test in the diagnosis of allergy to medicines”, *Anales del sistema sanitario de Navarra*, **2003**, 26 Suppl 2 39-47

Sanz M.L., Gamboa P.M., Garcia-Aviles C. et al., “Drug hypersensitivities: which room for biological tests”, *Allerg Immunol (Paris)*. **2005**,37, 6, 230-5.

Sanz M.L., Garcia M.C., Caballero M.R., et al. “Basophil activation test in the diagnosis of allergy to medicines”, *Anales del sistema sanitario de Navarra* , **2003** , 26 Suppl 2 39-47

Schwede T., Kopp J., Guex N., et al., “SWISSMODEL: An automated protein Homologymodeling server”, *Nucleic Acids Res*, **2004**, 31:3381–5

Shindyalov I.N. et Bourne P.E., “Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Engineering”, *Protein Engineering*, **2008**, 11(9):73947

Singh S.B., Kollman P.A., “Calculating the absolute free energy of association of netropsin and DNA”, *J Am Chem Soc*, **1999**, 121:3267-71

Spratt B.G., “Resistance to antibiotics mediated by target alterations”, *Science*, 1994, 264:388-93

Stewart J.J.P., *J Comput Chem*, **1989**, 10:209

Stewart J.J.P., *J Comput Chem*, **1989**, 10:221

Sugio S., Kashima A., Mochizuki S. et al., “Crystal structure of human serum albumin at 2.5 Å resolution”, *Protein Engineering*, **1999**, 12(6):439-446

Suri-Payer E., Amar A.Z., Thomton A.M.et al., “CD4+CD25+ T cells inhibit both the induction and effector function of autoreactive T cells and represent a unique lineage of immunoregulatory cells”, *J Immunol*, **1998**, 160(3):1212-8

Sybyl X : Tripos internationnal, 1699 South Hanley Rd, St. Louis, Missouri, 63144 (USA)

Teraki Y., Shiohara T., “IFN-gamma-producing effector CD8+ T cells and IL-10-producing regulatory CD4+ T cells in fixed drug eruption”, *J Allergy Clin Immunol*, **2003**, 112(3): 609-15

Terreux R., Domard M., Vitton C. et al., "Interaction study between a copper II ion and chitosan by DFT calculation", *Biomacromolecules*, **2006**, 7, 31-37

Torres M.J., Mayorga C., Garcia J.J., Romano A, et al. "New aspects in betalactam recognition", *Clin Exp Allergy*, **1998**, 4, 25-8

Torres M.J., Padial A., Mayorga C., et al., "The diagnostic interpretation of basophil activation test in immediate allergic reactions to betalactams", *Clin Exp Allergy*, **2004**,34, 1768-75

Tripos Inc. St. Louis, Missouri (USA)

Van Gunsteren W. F., Berendsen H.J.C., "Algorithms for macromolecular dynamics and constraint dynamics", *Mo. Phys*, **1977**, 34:1311-27

Verlet L., "Computer experiments on classical fluids. Thermodynamical properties of lennard jones molecules", *Phys Rev*, **1967**, 159:98-103

Viola M., Quaratino D., Gaeta F., Valluzzi R.L. et al., "Allergic reactions to antibiotics, mainly betalactams: facts and controversies", *Allerg Immunol (Paris)*, **2005**, 37, 223-9

Wal J.M., Bories G., Mamas M., et al. , "Radioimmunoassay of penicilloyl groups in biological fluids", *FEBS Lett*, **1975**, 57(1):9-13

Weiner H.L. "Induction and mechanism of action of transforming growth factor-beta-secreting Th3 regulatory cells", *Immunol Rev*, **2001**, 182:207-14

Weintraub, « Ernst Boris Chain (1906-1979) and Penicillin », *Bull. Israel Chem. Soc.*, **2003**, 29, 13

Weiss M.E, Adkinson NF., "Immediate hypersensitivity reactions to penicillin and related antibiotics", *Clin Allergy*. **1988**, 18(6):515-40

Weltzien, H. U., and E. Padovan, "Molecular features of penicillin allergy", *J Invest Dermatol*, **1998**, 110(3):203-6.

Wise R., "The pharmacokinetics of the oral cephalosporins - A review.", *J Antimicrob Chemother*, **1990**, 26(SupplE):13-20

York D.M., Darden T.A., Pedersen L.G., “The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods”, *J Chem Phys*, **1993**, 99(10): 8345-48

York D.M., Wlodawer A., Pedersen L.G. et al., “Atomic-level accuracy in simulations of large protein crystals”, *Proc Natl Acad Sci*, **1994**, 91:8715-18

Yvon M., and J.M. Wal, “Identification of lysine residue 199 of human serum albumin as a binding site for benzylpenicilloyl groups”, *FEBS Lett*, **1988**, 239(2):237-40

Yvon M., P. Anglade, and J. M. Wal, “Binding of benzyl penicilloyl to human serum albumin. Evidence for a highly reactive region at the junction of domains 1 and 2 of the albumin molecule”, *FEBS Lett*, **1989**, 247(2):273-8

Yvon, M., P. Anglade, and J. M. Wal, “Identification of the binding sites of benzyl penicilloyl, the allergenic metabolite of penicillin, on the serum albumin molecule”, *FEBS Lett*, **1990**, 263(2):237-40

Zanecki M., Cours en ligne, Université Angers, **2007**

## **Etude par modélisation moléculaire de l'effet allergène des antibiotiques de la famille des $\beta$ -lactamines, tant sur le plan immédiat que retardé**

---

Les hypersensibilités allergiques médicamenteuses sont des pathologies mettant en jeu le système immunitaire et induites par la prise de médicaments. Notre travail s'est décomposé en quatre étapes successives : 1- Nous avons classé les  $\beta$ -lactamines en fonction de leurs champs moléculaires, et obtenons un dendrogramme de 4 familles, validé par les données cliniques. Nous avons également réalisé une étude de 3D-QSAR visant à connaître les parties du médicament impliquées dans la pathologie, et à prédire l'allergénicité des  $\beta$ -lactamines. 2- Partant de l'hypothèse que les  $\beta$ -lactamines sont des haptènes, nous avons étudié leur réactivité vis-à-vis d'acides aminés de type lysine et sérine. Nous avons ensuite réalisé des expériences de « docking » afin de définir les interactions entre le médicament et l'albumine sérique humaine. Nous concluons que les sites des lysines 190 et 212 sont les plus adaptés pour la fixation covalente de la drogue et avons validé cette analyse par des méthodes mixtes QM/MM. Enfin, grâce à notre logiciel SuMo, nous avons déterminé d'autres protéines candidates pour l'hapténisation. 3- S'agissant des HyperSensibilités Allergiques Immédiates, nous avons modélisé les différents partenaires que sont les IgE, la  $\beta$ -lactamine portée ou non par une protéine. Nous avons envisagé plusieurs modes de reconnaissance. D'autre part, nous avons analysé les modifications de la protéine, induites par la fixation de la drogue. 4- Concernant les HSA retardées, nous avons émis plusieurs scénarios de reconnaissance de la  $\beta$ -lactamine par le TCR. Nous avons modélisé différents complexes impliquant le TCR, le peptide hapténisé par le médicament, un ion éventuel, ainsi que le CMH, et les soumettons à des dynamiques moléculaires afin d'en étudier la pertinence. D'autre part, nous avons déterminé plusieurs peptides, issus des protéines d'hapténisation et susceptibles de présenter le médicament au TCR, via le CMH. L'ensemble des résultats obtenus est ou sera validé par des expériences *in vitro* et *in vivo*.

---

### **Molecular modeling study of immediate and delayed drugs hypersensitivities : $\beta$ -lactams antibiotics allergenicity**

---

Drug hypersensitivity is an immune-mediated reaction to a drug. Our work was divided into four stages: 1 - We have classified  $\beta$ -lactam antibiotics based on their molecular fields, and obtained a dendrogram of 4 families, validated by clinical data. We also conducted a 3D-QSAR study to determine what parts of the drug are involved in the pathology and to predict the allergenicity of  $\beta$ -lactams. 2 - Under the assumption that  $\beta$ -lactam antibiotics are haptens, we studied their reactivity in comparison with lysine and serine. We then conducted "docking" experiments to define the interactions between the drug and human serum albumin. We conclude that lysine 190 and 212 are the most suitable sites for the covalent binding of the drugs. We validate this analysis by mixed QM / MM methods. Finally, thanks to our software SuMo, we have found other candidate proteins for haptening. 3 - Regarding immediate hypersensitivity reactions, we modeled the IgE, the  $\beta$ -lactam and the haptened protein. We considered several modes of recognition. Secondly, we analyzed the structural changes of the protein induced by the binding of the drug. 4 - Concerning delayed reactions, we considered different scenarios for the recognition of  $\beta$ -lactam by the TCR. We modeled complexes involving the TCR, the peptide haptened by the  $\beta$ -lactam, a possible ion, and the MHC. We investigated them with molecular dynamics to study their relevance. On the other hand, we have identified many peptides derived from haptening proteins and able to present the drug to the TCR through the MHC. The validity of the obtained results is or will be confirmed using experiments *in vitro* and *in vivo*

---

DISCIPLINE : Modélisation moléculaire

---

MOTS-CLES : Hypersensibilité allergique,  $\beta$ -lactamine, réactivité, mécanique quantique / moléculaire, albumine, IgE, docking, dynamique moléculaire, TCR, CMH, SuMo.

---

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE : Institut de Biologie et Chimie des Protéines UMR No 5086 – CNRS / UCBL