



HAL
open science

Modélisation moléculaire de l'acétylation de la quercétine par des lipases : étude des interactions enzyme-substrat

Christelle Bidouil

► **To cite this version:**

Christelle Bidouil. Modélisation moléculaire de l'acétylation de la quercétine par des lipases : étude des interactions enzyme-substrat. Alimentation et Nutrition. Université de Lorraine, 2012. Français. NNT : 2012LORR0263 . tel-01749457

HAL Id: tel-01749457

<https://hal.univ-lorraine.fr/tel-01749457v1>

Submitted on 29 Mar 2018

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

UNIVERSITÉ DE LORRAINE
École Nationale Supérieure d'Agronomie et des Industries Alimentaires
École Doctorale Ressources, Procédés, Produits et Environnement
Laboratoire d'Ingénierie des Biomolécules

THÈSE

présentée à l'UL par

Christelle BIDOUIL

en vue de l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ DE LORRAINE
Spécialité Procédés Biotechnologiques et Alimentaires

Sujet

**Modélisation Moléculaire de l'Acétylation de la Quercétine par des
Lipases : Étude des Interactions Enzyme-Substrat**

Soutenue publiquement le 13 novembre 2012

Membres du jury

Président :	Mr. Mohamed GHOUL	Professeur, Université de Lorraine, Nancy
Rapporteurs :	Mr. Yves BOULARD Mr. Vinh TRAN	Ingénieur de recherche, CEA Saclay, Paris Professeur, Université de Nantes, Nantes
Examineurs :	Mr. Jean-Marc ENGASSER Mme. Marianne GRABER Mme. Catherine HUMEAU	Professeur, Université de Lorraine, Nancy Professeur, Université de La Rochelle, La Rochelle M. d. C., Université de Lorraine, Nancy
Invité :	Mme. Latifa CHEBIL	M. d. C., Université de Lorraine, Nancy

«Il y a, au cours de l'existence, des rencontres imprévues et singulières, où bien des faits se produisent, en apparence anodins, et qui auront été les dons du destin.»
Hector Bianciotti

Remerciements

Ce travail de recherche a été réalisé au Laboratoire d'Ingénierie des Biomolécules de l'École Nationale Supérieure d'Agronomie et d'Industries Alimentaires sous la direction de Monsieur Jean-Marc Engasser, Professeur à l'ENSAIA, Monsieur Mohamed Ghoul, Professeur à l'ENSAIA et Mme Catherine Humeau, Maître de conférence à l'ENSAIA. Je les remercie pour m'avoir accueillie dans leur équipe et pour m'avoir fait confiance depuis mon arrivée au laboratoire.

Je tiens à exprimer ma sincère gratitude à M. Yves Boulard, Ingénieur de recherche au CEA Saclay, M. Vinh Tran, Professeur à l'Université de Nantes pour m'avoir fait l'honneur d'avoir évalué ce travail en tant que rapporteur. Je remercie également Mme Marianne Graber, Professeur à l'Université de La Rochelle, pour m'avoir examiné et jugé ce travail de thèse.

Je voudrais remercier tout particulièrement Catherine Humeau et Latifa Chebil pour leur encadrement scientifique, leur disponibilité, leur sympathie, leur encouragement durant ces trois années de thèse.

Je tiens à remercier M. Bernard Maigret, directeur de recherche émérite au LORIA-CNRS (Université Henri Poincaré, Nancy) pour ses conseils techniques tout le long de ce travail.

Je tiens à remercier Cédric Paris pour sa contribution à la réalisation des analyses LC-MS.

Ces remerciements seraient incomplets si n'apparaissait pas Nidal, Nabila, Leila, Florent, Charlotte, Jennifer, Citlalli, Marie-Céleste, Gregory, Florentin qui sont tous passés ou sont toujours au sein du laboratoire. Un merci tout particulier à Hugues pour son humour quotidien.

Ma famille a suivi avec attention mon parcours au fur et à mesure de ma scolarité. J'espère qu'ils seront fiers de moi.

Liste d'abréviations

Enzyme :

APE1547	Estérase de <i>Aeropyrum pernix</i> K1
CALB	Lipase B de <i>Candida antarctica</i>
MML	Lipase de <i>Mucor meihei</i>
PCL	Lipase de <i>Pseudomonas cepacia</i>

Acides aminés :

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Acide aspartique ou aspartate
Cys, C	Cystéine
Gln, Q	Glutamine
Glu, E	Acide glutamique ou glutamate
Gly, G	Glycine
His, H	Histidine
Ile, I	Isoleucine
Leu, L	Leucine
Lys, K	Lysine
Met, M	Méthionine
Phe, P	Phénylalanine
Pro, P	Proline
Ser, S	Sérine
Thr, T	Thréonine
Trp, W	Tryptophane
Tyr, Y	Tyrosine
Val, V	Valine

Autres abréviations :

Ac ₂ O	Anhydride d'acétate
AcONa	Acétate de sodium
BuOH	Butanol
CL	Chromatographie Liquide
CLHP	Chromatographie Liquide Haute Performance
DM	Dynamique Moléculaire
DMF	Diméthylformamide
H	Hydrogène
HEE	Ethyl-héxyl-phosphonate
kDa	Kilodalton
LTQ	Linear Trap Quadripole
MM	Modélisation Moléculaire
N	Nombre de particules dans le système
NAC	Near Attack Complex, complexe d'attaque
ns	nanoseconde (10 ⁻⁹ s)
OH	Groupement hydroxyle
PB	Poisson Boltzmann
PDB	<i>Protein Data Bank</i>
pNP	p-nitrophényle
ps	picoseconde (10 ⁻¹² s.)
QCT	quercétine
RMSD	<i>Root Mean Square Deviation</i>
s	seconde
SASA	Surface accessible au solvant
SM	Spectromètre de masse
t-BME	ter-Butyl méthyle éther

Sommaire général

Sommaire général	8
Liste des figures	12
Liste des tableaux.....	15
INTRODUCTION GENERALE	17
CHAPITRE I. ÉTUDE BIBLIOGRAPHIQUE.....	21
1. <i>Les principales approches de la modélisation moléculaire</i>	22
1.1. Simulation moléculaire classique.....	22
1.1.1. La mécanique moléculaire.....	22
1.1.2. Le champ de forces	22
1.1.3. Dynamique moléculaire	24
1.2. Docking moléculaire.....	24
1.2.1. Algorithmes de docking.....	24
1.2.2. Les fonctions de score.....	25
1.3. Énergie libre de liaison.....	27
1.3.1. Aspect thermodynamique	27
1.3.2. La méthode MM-PBSA.....	29
2. <i>La quercétine</i>	29
2.1. Généralités	29
2.2. Propriétés biologiques de la quercétine	31
2.2.1. Propriétés anti-oxydantes	31
2.2.2. Propriétés anti-inflammatoires.....	32
2.2.3. Propriétés anti-cancérigènes	33
2.2.4. Autres propriétés	33
2.3. Relation structure-activité de la quercétine.....	33
2.3.1. Propriétés anti-oxydantes	33
2.3.2. Chélation des métaux	34
2.4. Propriétés physico-chimiques de la quercétine.....	34
2.4.1. Acidité des groupements hydroxyles de la quercétine	34
2.4.2. Solubilité	35
2.5. Études des interactions de la quercétine avec des enzymes par modélisation moléculaire	36
2.5.1. Interactions entre les flavonoïdes et les acides aminés	37
2.5.2. Les modes d'interactions entre la quercétine et diverses protéines	37
2.6. L'acylation de la quercétine.....	39
2.6.1. Acylation par voie chimique	39
2.6.2. Acylation par voie enzymatique.....	41
2.6.3. Acylation par voie chimio-enzymatique.....	41
3. <i>Les lipases</i>	42
3.1. Introduction	42
3.2. Structure des lipases et mécanisme catalytique	44
3.2.1. Structure générale des lipases	44
3.2.2. Le mécanisme enzymatique des lipases	45
3.3. La lipase B de <i>Candida antarctica</i>	46
3.3.1. Structure de la CALB.....	46
3.3.2. Activation interfaciale.....	48
3.4. La lipase de <i>Pseudomonas cepacia</i>	49
3.5. La spécificité et la sélectivité de substrat des lipases	52
3.5.1. La spécificité de substrat	52
3.5.2. La sélectivité de substrat	53
4. <i>Le design de protéines</i>	57
4.1. Design rationnel de la lipase B de <i>Candida antarctica</i>	57
4.1.1. Impact des mutations de la CALB sur sa spécificité de réaction	58
4.1.2. Impact des mutations de la CALB sur son énantiométrie.....	59
4.1.3. Impact des mutations de la CALB sur son activité hydrolytique.....	60
4.1.4. Impact des mutations de la CALB sur sa thermostabilité	61
5. <i>Références</i>	63
CHAPITRE II. MATERIELS ET METHODES	71
A. PARTIE EXPERIMENTALE.....	72
1. <i>Matériels</i>	72

1.1. Descriptif du biocatalyseur	72
1.2. Réactifs et solvants	72
2. Mise en œuvre des réactions d'acylation enzymatique	73
3. Méthodes analytiques	73
3.1. Analyse des milieux de synthèse par chromatographie liquide haute performance	73
3.1.1. Matériels et protocole de séparation	73
3.1.2. Calcul du pourcentage de conversion des substrats	74
3.2. Analyse des milieux de synthèse par spectrométrie de masse	74
B. MODÉLISATION MOLÉCULAIRE	76
1. Ressources informatiques	76
1.1. Ordinateurs	76
1.2. Logiciels de simulation moléculaire	76
2. Préparation des structures	76
2.1. Structures des ligands	76
2.2. Structures des lipases	77
2.2.1. Structures de la lipase B de <i>Candida antarctica</i> (CALB)	77
2.2.2. Structures de la lipase de <i>Pseudomonas cepacia</i> (PCL)	79
3. Simulations de dynamique moléculaire	79
3.1. Paramètres des simulations	79
3.1.1. La température	79
3.1.2. Le pas d'intégration	80
3.1.3. L'algorithme SHAKE	80
3.1.4. Les ensembles thermodynamiques	80
3.1.5. Distance de troncation et méthode 'Particle Mesh Ewald'	80
3.1.6. Conditions périodiques frontières	81
3.2. Conditions des simulations de dynamique moléculaire	82
3.3. Protocole des simulations de dynamique moléculaire	83
3.3.1. Minimisation des systèmes	83
3.3.2. Équilibration des systèmes	84
3.3.3. Production des trajectoires de dynamique moléculaire	84
3.3.4. Analyse des résultats	84
4. Simulations de docking	86
4.1. Considérations générales	86
4.2. Définition du site actif	86
4.2.1. Principe	86
4.2.2. Définition du site actif des lipases CALB et PCL	87
4.3. Génération des complexes flavonoïde-lipase	88
4.3.1. Principe	88
4.3.2. Docking des flavonoïdes quercétine, isoquercitrine et des dérivés de la quercétine	90
4.4. Classement des complexes issus du docking	90
4.5. Calcul de l'énergie libre de liaison	93
4.5.1. Impact de la modification du substrat sur l'affinité de liaison	93
4.5.2. Impact de mutations dans CALB sur l'affinité de liaison	93
4.6. Validation des performances de LigandFit	94
5. Résumés schématiques des protocoles de modélisation moléculaire	95
5.1. Étude de la spécificité de la lipase B de <i>Candida antarctica</i> et de la lipase de <i>Pseudomonas cepacia</i> dans l'acétylation de la quercétine	95
5.2. Ingénierie de la lipase B de <i>Candida antarctica</i>	96
5.3. Étude des interactions CALB/quercétine par modification structurale du substrat	97
6. Références	98
CHAPITRE III. ÉTUDE EXPERIMENTALE DE L'ACÉTYLATION DE LA QUERCÉTINE EN PRÉSENCE DE LA LIPASE B DE CANDIDA ANTARCTICA	99
1. Introduction	100
2. Article	101
1. Introduction	102
2. Materials and methods	104
2.1. Reagents	104
2.2. Flavonoid ester synthesis procedure	104
2.3. HPLC - MS analysis	105
3. Results	105
3.1. Acetylation of quercetin in the absence of lipase	105
3.2. Acetylation of quercetin in the presence of CALB lipase	106

4. Discussion	108
5. References	109
<i>3. Contribution de l'article</i>	<i>111</i>
CHAPITRE IV. ÉTUDE PAR MODELISATION MOLECULAIRE DE LA SPECIFICITE DE LA LIPASE B DE <i>CANDIDA ANTARCTICA</i> ET DE LA LIPASE DE <i>PSEUDOMONAS CEPACIA</i> DANS LA REACTION D'ACETYLTATION DE LA QUERCETINE	113
1. Introduction	114
2. Article	116
1. INTRODUCTION	117
2. MODELLING METHODOLOGY	118
2.1. Computational resources	118
2.2. Preparation of the systems	118
2.3. Docking simulations	120
2.4. Molecular dynamics simulations	120
3. RESULTS	121
3.1. Quercetin acetylation catalyzed by <i>Pseudomonas cepacia</i> lipase (PCL)	121
3.2. Quercetin acetylation catalyzed by <i>Candida antarctica</i> lipase B (CALB)	125
4. DISCUSSION	129
5. References	132
<i>3. Contribution de l'article</i>	<i>135</i>
<i>4. Remarque complémentaire : les interactions CH₃-π</i>	<i>137</i>
CHAPITRE V. ÉTUDE PAR MODELISATION MOLECULAIRE DE L'INFLUENCE DE LA MUTATION DES RESIDUS HYDROPHOBES ILE189 ET ILE285 SUR LES MODES D'INTERACTIONS DE LA QUERCETINE DANS LE SITE ACTIF DE LA LIPASE B DE <i>CANDIDA ANTARCTICA</i>	139
1. Introduction	140
2. Article	142
1. Introduction	143
2. Computational methods	144
2.1. Construction of the <i>in silico</i> enzyme variants	144
2.2. MD calculations	144
2.3. Docking calculations	145
2.4. MM/PBSA Calculations	146
3. Results	146
3.1. Dynamics of wild-type CALB and its variants	146
3.2. Docking results	149
3.3. Dynamics of enzyme-substrates complexes	150
4. Discussion	156
5. References	159
<i>3. Contribution de l'article</i>	<i>162</i>
<i>4. Étude complémentaire: étude structural des variants double Ile189Ala-Ile285Ala, Ile189Ala-Ile285Val, Ile189Val-Ile285Ala, Ile189Val-Ile285Val de la lipase B de Candida antarctica</i>	<i>164</i>
4.1. Stabilité structurale des mutants doubles	164
4.2. Énergie d'interaction	168
<i>5. Étude complémentaire : étude structurale du variant Asp134Ala et des interactions avec la quercétine</i>	<i>170</i>
5.1. Étude structurale du variant Asp134Ala	170
5.2. Interaction de la quercétine avec le variant Asp134Ala	173
CHAPITRE VI. ÉTUDE THEORIQUE DES INTERACTIONS ENTRE LA QUERCETINE ET SES DERIVES, ET LE SITE ACTIF DE LA LIPASE B DE <i>CANDIDA ANTARCTICA</i>	175
1. Introduction	176
2. Article	179
1. Introduction	180
2. Computational Methods	181
2.1. - Preparation of structures	181
2.2. Molecular Docking	182
2.3. Binding free energy calculations	182
3. Results	183

3.1. Dynamic behaviour of CALB and acetyl-CALB targets	183
3.2. Molecular Docking	186
3.3. Binding free energy calculations	189
4. Discussion	194
5. References	196
<i>3. Contribution de l'article.....</i>	<i>198</i>
CONCLUSION GENERALE – PERSPECTIVES	201
ANNEXE.....	211

Liste des figures

CHAPITRE I

FIGURE I.1. REPRESENTATION SCHEMATIQUE DE L'EQUATION SUIVANTE $RAQ + LAQ \rightleftharpoons RLAQ$	28
FIGURE I.2. CYCLE THERMODYNAMIQUE POUR LE CALCUL DE L'ENERGIE LIBRE DE LIAISON ENTRE UN RECEPTEUR (R) ET UN LIGAND (L).	28
FIGURE I.3. NOMENCLATURE DES CYCLES ET NUMEROTATION SYSTEMATIQUE DES HYDROXYLES DE LA QUERCETINE.	30
FIGURE I.4. MECANISMES DE L'ACTIVITE ANTI-OXYDANTE DE LA QUERCETINE.	32
FIGURE I.5. REACTION DE TRANSFERT D'ATOME D'HYDROGENE ENTRE LA QUERCETINE ET UN RADICAL LIBRE, R.	34
FIGURE I.6. LA DISTRIBUTION DE PROBABILITE DE LA DENSITE DE SOLVANTS AUTOUR DE LA QUERCETINE : (A) L'EAU, (B) L'ACETONITRILE, (C) L'ACETONE, (D) LE <i>TERT</i> -AMYL ALCOOL, (E) LE CHLOROFORME [58].....	36
FIGURE I.7. SYNTHESES CHIMIQUES DU 3,3',4',5,7-PENTAACYL QUERCETINE D'APRES [72].....	40
FIGURE I.8. LES DIFFERENTES REACTIONS CATALYSEES PAR LES LIPASES.	43
FIGURE I.9. REPRESENTATION SCHEMATIQUE DU REPLIEMENT DES α/β -HYDROLASES.	45
FIGURE I.10. MECANISME CATALYTIQUE DES LIPASES.	45
FIGURE I.11. STRUCTURE TRIDIMENSIONNELLE DE LA CALB (PDB ID: 1LBS).	46
FIGURE I.12. TOPOLOGIE DE LA STRUCTURE SECONDAIRE DE LA CALB.	47
FIGURE I.13. REPRESENTATION EN SURFACE DE LA CAVITE DE LA LIPASE B DE <i>CANDIDA ANTARCTICA</i>	47
FIGURE I.14. TOPOLOGIE DE LA STRUCTURE SECONDAIRE DE LA LIPASE DE <i>PSEUDOMONAS CEPACIA</i>	50
FIGURE I.15. (A) STRUCTURE TRIDIMENSIONNELLE DE LA PCL (PDB ID: 3LIP). (B) LA TRIADE CATALYTIQUE DE LA PCL.	51
FIGURE I.16. REPRESENTATION EN SURFACE DE LA CAVITE CATALYTIQUE DE LA LIPASE DE <i>PSEUDOMONAS CEPACIA</i>	51
FIGURE I.17. SPECIFICITE DES LIPASES CALB ET PCL ENVERS LE FLAVONOÏDE AGLYCONE QUERCETINE.....	52
FIGURE I.18. REACTION D'ACYLATION D'UNE FONCTION ALCOOL ET D'UNE FONCTION THIOL PAR LA LIPASE B DE <i>CANDIDA ANTARCTICA</i>	53
FIGURE I.19. REACTION DE TRANSACYLATION SUR UN MELANGE ENANTIOMERIQUE DE 1-PHENYLETHANOL CATALYSEE PAR CALB.	53

CHAPITRE II

FIGURE II.1. ANALYSE CL-SM D'UNE SOLUTION TEMOIN DE QUERCETINE DANS L'ACETONITRILE. (A) CHROMATOGRAMME UV A 400 NM (B) SPECTRE DE MASSE DU PIC PRESENTANT UN TEMPS DE RETENTION DE 13,50 MIN.	75
FIGURE II.2. STRUCTURE CHIMIQUE, NUMEROTATION SYSTEMATIQUE ET DENOMINATION DES LIGANDS ETUDIES.....	77
FIGURE II.3. REPRESENTATION GRAPHIQUE DU POTENTIEL DE LE NNARD-JONES 6-12.	81
FIGURE II.4. POTENTIEL DE VAN DER WAALS EN PRESENCE (TRAIT DISCONTINU) ET EN ABSENCE (TRAIT CONTINU) DE LA FONCTION 'SWITCHING'.	81
FIGURE II.5. REPRESENTATION EN DEUX DIMENSIONS DES CONDITIONS PERIODIQUES FRONTIERES.	82
FIGURE II.6. REPRESENTATION SCHEMATIQUE DU PROTOCOLE DES SIMULATIONS DE DYNAMIQUE MOLECULAIRE.....	83
FIGURE II.7. CRITERES POUR ETABLIR LA PRESENCE D'UNE LIAISON HYDROGENE ENTRE UN DONNEUR (X) D'HYDROGENE ET UN ACCEPTEUR (A) D'HYDROGENE.	85
FIGURE II.8. REPRESENTATION SCHEMATIQUE DE LA GRILLE ETENDUE A TRAVERS LA PROTEINE.	87
FIGURE II.9. REPRESENTATION SCHEMATIQUE DES QUATRE ORIENTATIONS DU LIGAND DANS LE SITE ACTIF DU RECEPTEUR.	88
FIGURE II.10. DENDROGRAMME DES POSES ISSUES DU DOCKING DE LA QUERCETINE DANS LA CALB ACETYLEE.	92
FIGURE II.11. SUPERPOSITION DE LA STRUCTURE CRISTALLOGRAPHIQUE 1LBS ET DU MEILLEUR COMPLEXE OBTENU PAR DOCKING DE L'INHIBITEUR HEE DANS LA CAVITE DE CALB (CODE PDB : 1LBS), EN UTILISANT LIGANDFIT.	94
FIGURE II.12. REPRESENTATION SCHEMATIQUE DU PROTOCOLE APPLIQUE POUR L'ETUDE DE LA SPECIFICITE DES LIPASES CALB ET PCL DANS LA REACTION D'ACETYLATION DE LA QUERCETINE.	95
FIGURE II.13. REPRESENTATION SCHEMATIQUE DU PROTOCOLE APPLIQUE POUR L'ETUDE STRUCTURALE DES VARIANTS SIMPLES DE CALB.	96
FIGURE II.14. REPRESENTATION SCHEMATIQUE DU PROTOCOLE SUIVI POUR L'ETUDE STRUCTURALE DES VARIANTS DOUBLES DE CALB.	96
FIGURE II.15. REPRESENTATION SCHEMATIQUE DU PROTOCOLE APPLIQUE POUR L'ETUDE DES INTERACTIONS CALB/QUERCETINE, PAR MODIFICATION STRUCTURALE DU SUBSTRAT.	97

CHAPITRE III

FIGURE III.1. CONVENTIONAL NOMENCLATURE OF QUERCETIN STRUCTURE	102
FIGURE III.2. REGIOSELECTIVITY OF QUERCETIN ACYLATION BY CHEMICAL, ENZYMATIC AND CHEMO-ENZYMATIC APPROACH USED.	103
FIGURE III.3. MS SPECTRUM (B, C, D) CORRESPONDING TO UV CHROMATOGRAM AT 400 NM (A) AFTER 72 HOURS REACTION FOR PRODUCTS QUERCETIN IN THE ABSENCE OF LIPASE CALB.	106
FIGURE III.4. MS SPECTRUM (B, C AND D) CORRESPONDING TO UV CHROMATOGRAM AT 400NM (A) AFTER 72 HOURS REACTION FOR PRODUCTS QUERCETIN IN THE PRESENCE OF CALB.....	107
FIGURE III.5. CHROMATOGRAM CORRESPONDING TO THE MASSES OF TRIACETATE QUERCETIN (A, B) $m/z [M+H]^+ = 429$ AND $m/z [M+NA]^+ = 451$, AND OF TETRAACETATE QUERCETIN (C, D) $m/z [M+H]^+ = 471$ AND $m/z [M+NA]^+ = 493$ IN THE PRESENCE OF CALB (A, C) AND IN THE ABSENCE OF CALB (B, D) FOR SAMPLES UNDILUTED.....	108

CHAPITRE IV

FIGURE IV.1. SCHEMATISATION DES CRITERES STRUCTURAUX DEVANT ETRE RESPECTES POUR QUE LES COMPLEXES ENZYME/SUBSTRATS SOIENT CONSIDERES COMME PRODUCTIFS.	115
FIGURE IV.2. CHEMICAL STRUCTURE OF QUERCETIN, SHOWING THE SYSTEMATIC NUMBERING OF THE CARBON ATOMS, THE NOMENCLATURE OF THE RINGS AND THE DIHEDRAL ANGLE τ , DEFINED BY THE ATOMS O1-C2-C1'-C2'.....	117
FIGURE IV.3. SUPERIMPOSITION OF REPRESENTATIVE SNAPSHOTS OF THE TRAJECTORY OF THE PAQ50 COMPLEX.....	123
FIGURE IV.4. SUPERIMPOSITION OF REPRESENTATIVE SNAPSHOTS OF THE PBQ21 COMPLEX (FRONTAL VIEW).	124
FIGURE IV.5. FRONTAL VIEW OF THE CAQ2 COMPLEX.	127
FIGURE IV.6. SUPERIMPOSITION OF REPRESENTATIVE SNAPSHOTS OF THE CBQ18 COMPLEX WHEN THE CARBONYL OF THE SERINE-BOUND ACETATE POINTED TOWARDS (YELLOW) AND AWAY (GREEN) THE OXYANION HOLE	129
FIGURE IV.7. (A) LORSQUE LA LIAISON C-H DU METHANE EST PERPENDICULAIRE AU CYCLE BENZENIQUE, L'ANGLE EST PAR CONVENTION FIXE A 0° TANDIS QUE L'ANGLE EST DE 90° QUAND LA LIAISON C-H EST DANS LE PLAN DU CYCLE AROMATIQUE ; (B) SURFACE D'ENERGIE POTENTIELLE POUR LES INTERACTIONS ENTRE LE METHANE ET LE BENZENE.	137
FIGURE IV.8. REPRESENTATION DU SYSTEME METHANE - NAPHTALENE. L'INTERACTION CH - π EST MAXIMISEE LORSQUE LE CARBONE DE LA MOLECULE DE METHANE SE POSITIONNE AU CENTRE DE LA MOLECULE DE NAPHTALENE.....	138

CHAPITRE V

FIGURE V.1. (A) SCHEMATISATION SIMPLIFIEE DE L'ENCOMBREMENT STERIQUE OCCASIONNE PAR LES CHAINES LATERALES DES RESIDUS ILE189 ET ILE285 DANS LA CAVITE DE LA CALB. (B) SCHEMATISATION DU POSITIONNEMENT IDEAL DE LA QUERCETINE POUR L'ATTAQUE NUCLEOPHILE DE L'ACYLE-ENZYME.	140
FIGURE V.2. REPRESENTATION SCHEMATISEE DES MUTATIONS EFFECTUEES SUR LES RESIDUS ILE189 ET ILE285.	141
FIGURE V.3. B-FACTOR VALUES CALCULATED FROM MD TRAJECTORIES OF CALB WILD TYPE AND VARIANTS.	148
FIGURE V.4. VOLUME OF CATALYTIC POCKET OVER THE 20 NS SIMULATION FOR CALB WILD-TYPE (GREEN) AND FOR VARIANTS ILE189ALA (BLUE) AND ILE285ALA (RED).....	148
FIGURE V.5. SNAPSHOT FROM CALB MOLECULAR DYNAMICS TRAJECTORY SHOWING CATALYTIC DISTANCES (A). OPTIMIZED STRUCTURE OF ACYL-ENZYME SYSTEM (B).	149
FIGURE V.6. EVOLUTION OF $C\alpha$ ATOMS RMSD WITH TIME FOR COMPLEXES INVOLVING QUERCETIN AND CALB WILD TYPE OR CALB VARIANTS (A AND C). EVOLUTION OF QUERCETIN RMSD IN THE SAME SYSTEMS (B AND D).	151
FIGURE V.7. ATOMIC FLUCTUATIONS (B-FACTOR) OF RESIDUES, BASED ON $C\alpha$ ATOMS AND CALCULATED OVER 10 NS MOLECULAR DYNAMICS TRAJECTORIES	152
FIGURE V.8. (A) MAJOR RESIDUES ESTABLISHING HYDROGEN BOND INTERACTIONS WITH QUERCETIN HYDROXYL GROUPS. (B) HYDROGEN BOND NETWORK STABILIZING QUERCETIN WITHIN ILE189ALA MUTANT (ILE189ALA-QA COMPLEX).	154
FIGURE V.9. SUIVI DU RMSD DES $C\alpha$ AU COURS DES TRAJECTOIRES DE DM (20 NS) POUR LES QUATRES VARIANTES DOUBLES DE LA CALB.	164
FIGURE V.10. B-FACTOR DES ATOMES $C\alpha$ CALCULE SUR LES 18 DERNIERES NANOSECONDES DE TRAJECTOIRE POUR LES DOUBLES VARIANTES DE LA CALB	166
FIGURE V.11. SUPERPOSITION DE LA FRAME INITIALE (BLEU) ET DE LA FRAME FINALE (JAUNE) POUR LE SYSTEME ILE189VAL-ILE285VAL	166
FIGURE V.12. REPRESENTATION EN BOITE A MOUSTACHE DE L'ENERGIE D'INTERACTION NON LIEE CALCULEE ENTRE LE RESIDU ILE189 ET L'HELICE $\alpha 10$ POUR LA CALB SAUVAGE ET SES VARIANTES SIMPLES ET DOUBLES PENDANT LA PHASE DE PRODUCTION DE LA TRAJECTOIRE DE DM.	168

FIGURE V.13. REPRESENTATION EN BOITE A MOUSTACHE DE L'ENERGIE D'INTERACTION DE VAN DER WAALS ENTRE LE RESIDU ILE189 ET L'HELICE α 10 POUR LA CALB SAUVAGE ET SES VARIANTS SIMPLES ET DOUBLES.	169
FIGURE V.14. REPRESENTATION EN BOITE A MOUSTACHE DE L'ENERGIE D'INTERACTION ELECTROSTATIQUE ENTRE LE RESIDU ILE189 ET L'HELICE α 10 POUR LA CALB ET SES VARIANTS SIMPLES ET DOUBLES.	169
FIGURE V.15. VARIANT ASP134ALA DE LA LIPASE B DE <i>CANDIDA ANTARCTICA</i>	170
FIGURE V.16. SUIVI DU RMSD DU C α AU COURS DU TEMPS POUR LE VARIANT ASP134ALA DE LA CALB.	171
FIGURE V.17. B-FACTOR EN FONCTION DES RESIDUS DU VARIANT ASP134ALA DE LA CALB (ROSE) ET DE LA CALB SAUVAGE (BLEU).	171
FIGURE V.18. SUIVI DU VOLUME DE LA CAVITE CATALYTIQUE DU VARIANT ASP134ALA AU COURS DE LA TRAJECTOIRE DE DM.	172
FIGURE V.19. IMAGES MONTRANT LES DEUX ORIENTATIONS MAJORITAIRES DE LA QUERCETINE, CA (A) ET CCB (B), OBSERVEES A L'ISSUE DE LA PROCEDURE DE DOCKING DANS LA CAVITE CATALYTIQUE DU VARIANT ASP134ALA.	173

CHAPITRE VI

FIGURE VI.1. PRESENTATION DES DIFFERENTS SYSTEMES ETUDIES.	178
FIGURE VI.2. CHEMICAL STRUCTURES AND NOMENCLATURES OF THE LIGANDS UNDER STUDY	182
FIGURE VI.3. 2D-RMSD PLOTS CALCULATED FROM MOLECULAR DYNAMICS TRAJECTORIES OF CALB (A) AND ACETYL CALB (B).	184
FIGURE VI.4. DISTANCES (Å) WITHIN CATALYTIC TRIAD OF CALB (A) AND ACETYL CALB (B) THROUGHOUT MOLECULAR DYNAMICS SIMULATIONS.	184
FIGURE VI.5. DISTANCES (Å) BETWEEN THE SERINE-BOUND ACETATE AND THE OXYANION HOLE RESIDUES THROUGH OUT MOLECULAR DYNAMICS SIMULATIONS OF ACETYL-CALB SYSTEM.	185
FIGURE VI.6. SCHEMATIC REPRESENTATION OF THE CRITERIA RETAINED FOR THE SELECTION OF THE FRAMES SUBJECTED TO FURTHER MOLECULAR DOCKING.	185
FIGURE VI.7. MODELS OF COMPLEXES SHOWING THE SUBSTRATE 7 3',4' - METHYL QUERCETIN IN A CA ORIENTATION WITHIN ACETYL-CALB TARGET (A) AND IN A CB ORIENTATION WITHIN CALB TARGET (B).	188
FIGURE VI.8. MODELS OF ACETYL-CALB (A) AND CALB (B) TARGETS.	188
FIGURE VI.9. EXAMPLE OF POSE RESULTING FROM THE DOCKING OF PENTAACETYL-QUERCETIN AGAINST CALB AND SHOWING A MODEL OF COMPLEX INVOLVED IN ALCOHOLYSIS REACTION.	189
FIGURE VI.10. DISTRIBUTION OF BINDING FREE ENERGIES OF ACETYL-CALB/QUERCETIN AND ACETYL-CALB/ISQUERCITRIN COMPLEXES GENERATED BY DOCKING.	190
FIGURE VI.11. DISTRIBUTION OF BINDING FREE ENERGIES OF COMPLEXES INVOLVING METHYL QUERCETIN DERIVATIVES WITHIN ACETYL-CALB OR CALB CATALYTIC POCKET.	192
FIGURE VI.12. DISTRIBUTION OF BINDING FREE ENERGIES OF COMPLEXES ASSOCIATING ACETYL QUERCETIN DERIVATIVES AND CALB CATALYTIC POCKET.	193

Liste des tableaux

CHAPITRE I

TABLEAU I.1. LES DIFFERENTS TERMES ENERGETIQUES D'UN CHAMP DE FORCES CLASSIQUE.....	23
TABLEAU I.2. CONCENTRATION EN QUERCETINE DANS DIVERS FRUITS ET LEGUMES, ET DANS DIVERSES BOISSONS.....	30
TABLEAU I.3. VALEURS DE PKA DETERMINEES EXPERIMENTALEMENT OU THEORIQUEMENT ET ATTRIBUTION AU DIFFERENTS GROUPEMENTS HYDROXYLES DE LA QUERCETINE.....	35
TABLEAU I.4. SOLUBILITE DE LA QUERCETINE DANS DIFFERENTS SOLVANTS.....	36
TABLEAU I.5. QUELQUES EXEMPLES DE TYPES D'INTERACTIONS REPERTORIEES DANS LA LITTERATURE ENTRE LA QUERCETINE ET DIVERSES PROTEINES.....	38
TABLEAU I.6. CONDITIONS EXPERIMENTALES DE LA SYNTHESE CHIMIQUE DES DERIVES ACYLES DE LA QUERCETINE.....	40
TABLEAU I.7. CONDITIONS OPERATOIRES ET PRODUITS IDENTIFIES POUR L'ACYLATION ENZYMATIQUE DE LA QUERCETINE (QCT) AVEC LES BIOCATALYSEURS PCL, CALB ET APE1547.....	41
TABLEAU I.8. ALCOOLYSE DE LA QUERCETINE PENTAACYLEE PAR DIFFERENTS BIOCATALYSEURS.....	41
TABLEAU I.9. DIVERSES APPLICATIONS DES LIPASES DANS DIFFERENTS SECTEURS INDUSTRIELS.....	43
TABLEAU I.10. QUELQUES TRAVAUX TRAITANT DE LA REGIOSELECTIVITE DE LA CALB DANS L'ACYLATION DE COMPOSES POLY-HYDROXYLES.....	55
TABLEAU I.11. MUTANTS DE LA LIPASE B DE <i>CANDIDA ANTARCTICA</i> IMPLIQUES DANS L'AMELIORATION DE DIVERSES REACTIONS.....	58
TABLEAU I.12. MUTATIONS DE LA CALB AFFECTANT LA SYNTHESE ENANTIOSELECTIVE DE DIFFERENTS SUBSTRATS.....	59
TABLEAU I.13. IMPACT DE MUTATIONS DE LA CALB SUR L'HYDROLYSE ENANTIOSELECTIVE DE DIFFERENTS SUBSTRATS.....	60
TABLEAU I.14. ACTIVITE HYDROLYTIQUE DE MUTANTS DE LA CALB.....	60
TABLEAU I.15. MUTATIONS DE LA CALB QUI AGISSENT SUR LA THERMOSTABILITE DE LA LIPASE.....	61

CHAPITRE II

TABLEAU II.1. DESCRIPTIF DES REACTIFS ET SOLVANTS UTILISES.....	72
TABLEAU II.2. GRADIENT D'ELUTION UTILISE POUR LA SEPARATION DE LA QUERCETINE ET DE SES ESTERS.....	74
TABLEAU II.3. DETAIL DES DIFFERENTES ETAPES DE PREPARATION DES STRUCTURES INITIALES DE LA CALB.....	78
TABLEAU II.4. DETAIL DES DIFFERENTES ETAPES DE PREPARATION DES STRUCTURES DE LA PCL.....	79
TABLEAU II.5. SPECIFICATION DES PARAMETRES UTILISES LORS DES SIMULATIONS DE DOCKING.....	90
TABLEAU II.6. LES TYPES DE FONCTIONS DE SCORE IMPLEMENTES DANS <i>DISCOVERY STUDIO 2.5</i>	91
TABLEAU II.7. PARAMETRES POUR LE CALCUL DE L'ENERGIE LIBRE DE LIAISON.....	94

CHAPITRE IV

TABLE IV.1. AVERAGE DISTANCES BETWEEN THE QUERCETIN HYDROXYL GROUPS AND THE CATALYTIC RESIDUE His286 AND THE SERINE-BOUND ACETATE.....	125
TABLE IV.2. AVERAGE DISTANCES (Å) BETWEEN THE CARBONYL OXYGEN OF THE SERINE-BOUND ACETATE AND THE OXYANION HOLE RESIDUES THROUGHOUT THE 10 NS MOLECULAR DYNAMICS TRAJECTORIES FOR BOTH CA COMPLEXES.....	126
TABLE IV.3. DISTANCES (Å) BETWEEN THE CENTER OF MASS OF QUERCETIN RINGS AND THE CARBON ATOMS OF THE SIDE CHAINS METHYL GROUP OF Ile189 AND Ile285.....	127

CHAPITRE V

TABLE V.1. DATA RELATIVE TO THE STRUCTURE AND THE FLEXIBILITY OF CALB WILD TYPE AND ITS VARIANTS THROUGHOUT MOLECULAR DYNAMICS TRAJECTORIES.....	147
TABLE V.2. AVERAGE VALUE AND STANDARD DEVIATION OF DISTANCES AND ANGLES BETWEEN CATALYTIC RESIDUES FOR CALB AND ITS VARIANTS.....	149
TABLE V.3. MAJOR CONFORMATIONAL FAMILIES OBTAINED FROM DOCKING FOR CALB WILD TYPE AND ITS VARIANTS. SELECTED COMPLEXES ARE GIVEN IN THE RIGHT SECTION.....	150
TABLE V.4. OCCURRENCE RATIO AS WELL AS AVERAGE VALUE AND STANDARD DEVIATION OF DISTANCE AND ANGLE OF H-BOND INTERACTIONS BETWEEN CALB RESIDUES AND QUERCETIN HYDROXYL GROUPS THROUGHOUT MOLECULAR DYNAMICS TRAJECTORIES.....	153

TABLE V.5. PERCENTAGE OF FRAMES EXHIBITING A GOOD ORIENTATION OF THE SERINE-BOUND ACETATE TOWARDS THE OXYANION HOLE AND A CORRECT CONFORMATION OF THE CATALYTIC TRIAD.....	155
TABLE V.6. AVERAGE AND STANDARD DEVIATION DISTANCES BETWEEN QUERCETIN HYDROXYL GROUPS AND BOTH THE CATALYTIC RESIDUE HIS224 AND THE SERINE-BOUND ACETATE.....	155
TABLE V.7. AVERAGE VALUES AND STANDARD DEVIATION OF BINDING FREE ENERGIES AND INDIVIDUAL ENERGY TERMS. ALL ENERGIES WERE EXPRESSED IN KCAL.MOL ⁻¹	156
TABLEAU V.1. DIFFERENTS PARAMETRES CONTROLES AU COURS DES TRAJECTOIRES DE DM DES DOUBLES MUTANTS DE LA CALB.	165
TABLEAU V.2. VALEURS MOYENNES ET ECART-TYPES DU VOLUME DE LA POCHE DU SITE ACTIF (Å ³) POUR LES MUTANTS DOUBLES ET LES MUTANTS SIMPLES DE LA CALB.	167
TABLEAU V.3. DISTANCES ET ANGLES ENTRE LES RESIDUS CONSTITUANT LA TRIADE CATALYTIQUE AU COURS DES 10 NS DE LA PHASE DE PRODUCTION DE LA TRAJECTOIRE DE DM.....	172

CHAPITRE VI

TABLE VI.1. DISTANCES (Å) AND ANGLES (°) MEASURED IN CALB CATALYTIC CAVITY.....	186
TABLE VI.2. DISTANCES (Å) AND ANGLES (°) MEASURED IN ACETYL-CALB CATALYTIC CAVITY.....	186
TABLE VI.3. AVERAGE AND STANDARD DEVIATION OF BINDING FREE ENERGY CONTRIBUTIONS FOR COMPLEXES INVOLVING QUERCETIN AND ISOQUERCETIN.	191
TABLE VI.4. AVERAGE AND STANDARD DEVIATIONS OF BINDING FREE ENERGY CONTRIBUTIONS FOR COMPLEXES INVOLVING EITHER ACETYL-CALB OR CALB TARGET AND METHYL QUERCETIN DERIVATIVES.	192
TABLE VI.5. AVERAGE AND STANDARD DEVIATION OF BINDING FREE ENERGY CONTRIBUTIONS FOR ACETYL QUERCETIN/CALB COMPLEXES.	194

Introduction Générale

Le nombre de procédés industriels faisant appel à la biocatalyse pour la synthèse organique d'intermédiaires est devenu important. L'utilisation des enzymes dans ces réactions s'est développée grâce à leurs caractéristiques en terme de spécificité, de sélectivité, de « turnover », de stabilité thermique et de résistance aux solvants. Ainsi, cette sélectivité, incluant l'énantiosélectivité et la régiosélectivité vis-à-vis d'un substrat ou d'une classe de substrats donnée, en fait un outil de choix pour l'accès à de nouveaux produits ayant des structures moléculaires difficilement accessibles par les voies « classiques » de synthèse organique. Par ailleurs, ces procédés enzymatiques s'effectuent aussi bien en milieu aqueux qu'en milieu organique, ce qui permet de répondre aux problèmes de solubilité dans l'eau de nombreuses molécules organiques. En outre, elles s'opèrent dans des conditions plus douces que leurs analogues chimiques : température moins élevée, pression atmosphérique, pH moins extrême et absence de catalyseurs inorganiques toxiques.

Le Laboratoire d'Ingénierie de Biomolécules s'intéresse depuis plusieurs années au développement et à l'application de procédés enzymatiques exploitant des lipases pour produire des dérivés de flavonoïdes acylés régiosélectivement. En milieu organique, les lipases catalysent la formation de liaisons esters. Les flavonoïdes sont des composés naturels présents dans une large variété de fruits et de légumes consommés quotidiennement par l'être humain. Certains de ces composés présentent des activités biologiques potentiellement bénéfiques, dont les plus connues sont la prévention de la propagation d'espèces radicalaires néfastes aux cellules ainsi que les pouvoirs anti-oxydant, anti-inflammatoire et anti-cancérigène. L'éventail de propriétés biologiques des flavonoïdes a été corrélé à des éléments structuraux, notamment le nombre et la position des groupements hydroxyles. Toutefois, la structure amphiphile entraîne une miscibilité réduite des flavonoïdes, que ce soit en milieu hydrophile ou lipophile, ce qui constitue un obstacle pour la valorisation technologique de leurs propriétés. Une solution pour obtenir des dérivés avec un caractère plus lipophile est le greffage d'une chaîne acyle qui permet de bénéficier de la bioactivité des molécules natives dans des matrices hydrophobes. L'utilisation de biocatalyseurs dans ces réactions d'acylation est un compromis entre l'amélioration de la solubilité, de la stabilité, de la biodisponibilité et le maintien des propriétés d'intérêt des flavonoïdes. En effet, l'acylation sur des groupements hydroxyles à l'origine des effets biologiques recherchés peut être évitée grâce à la régiosélectivité des enzymes.

Des travaux antérieurs ont été menés au Laboratoire sur l'acylation des flavonoïdes glycosylés, isoquercitrine et rutine. Parmi les enzymes capables de catalyser l'acylation de ces flavonoïdes, la lipase B de *Candida antarctica* (CALB) s'est révélée hautement régiosélective. Une étude de modélisation moléculaire a ensuite été entreprise pour expliquer la régiosélectivité de cette lipase envers l'isoquercitrine et la rutine dans la réaction d'acétylation. Ainsi, à la suite d'une procédure de docking/scoring, une définition de complexe enzyme-substrat dit « productif » a été établie en se basant sur le mécanisme catalytique des lipases et la littérature. En outre, cette étude a démontré que le positionnement et l'orientation des substrats dans la cavité du site actif de l'enzyme

sont directement corrélés à la régiosélectivité de la réaction. L'application de ce protocole de modélisation moléculaire a permis de rationaliser les résultats observés expérimentalement.

D'autre part, notre équipe a également étudié la réaction d'acylation enzymatique de flavonoïdes aglycones dont la quercétine. Dans ce cas, les résultats expérimentaux obtenus lors de l'acylation de ce flavonoïde montre qu'avec la lipase B de *Candida antarctica*, il n'y a pas d'activité d'acylation tandis qu'avec la lipase de *Pseudomonas cepacia* (PCL) des produits acétylés de la quercétine ont été détectés. Toutefois, des résultats contradictoires ont été récemment publiés concernant l'activité catalytique de la CALB envers le substrat quercétine. Une équipe a obtenu des dérivés acétylés de la quercétine en présence de ce biocatalyseur et d'un fort excès d'acétate de vinyle [1].

Dans le cadre de ce travail de thèse, nous nous sommes focalisés sur la réaction d'acétylation du flavonoïde aglycone quercétine en présence du biocatalyseur CALB. Nous tenterons d'apporter des réponses aux questions suivantes :

- (1) La lipase B de *Candida antarctica* (CALB) catalyse-t-elle ou non l'acétylation de la quercétine avec l'acétate de vinyle comme donneur d'acyle ?
- (2) Comment les modes d'interactions enzyme-substrat peuvent-ils expliquer la spécificité de la lipase B de *Candida antarctica* et de la lipase de *Pseudomonas cepacia* dans la réaction d'acétylation enzymatique du flavonoïde quercétine ? Peut-on expliquer au niveau moléculaire les différences d'activité d'acétylation entre la CALB et la PCL ? Quels sont les résidus étroitement impliqués dans le positionnement de la quercétine dans la CALB ?
- (3) Quelle est l'influence de la mutation des résidus Ile189 et Ile285 sur la structure de la CALB et, sur le positionnement des substrats acétate et quercétine dans son site actif ?
- (4) Comment la méthylation et l'acétylation de la quercétine influencent-elles son orientation dans la cavité de la CALB native et/ou acétylée et, son affinité pour le site actif ? Comment l'acétate lié de façon covalente à la sérine catalytique intervient-il sur le positionnement de la quercétine et de ses dérivés ainsi que sur l'affinité de liaison ?

Dans le **premier chapitre**, nous commençons par une description des approches de modélisation moléculaire utilisées dans le cadre de ce travail. Il s'ensuit un récapitulatif des propriétés biologiques et physico-chimiques de la quercétine ainsi que des types d'interactions de ce flavonoïde avec diverses protéines. Les différentes méthodes d'acylation chimiques, enzymatiques et chimio-enzymatique de ce flavonoïde aglycone répertoriées dans la littérature sont détaillées. Ensuite, les structures des lipases CALB et PCL sont décrites. Enfin, l'ingénierie de la CALB est abordée.

Le **deuxième chapitre** contient d'une part, le protocole expérimental et les méthodes d'analyse et d'autre part, une description du mode opératoire suivi pour les simulations de docking et de dynamique moléculaire. Les différents paramètres de modélisation sont également présentés.

Dans le **troisième chapitre**, le potentiel catalytique de la lipase B de *Candida antarctica* dans la réaction d'acétylation de la quercétine est clarifié grâce à une expérimentation. Cette partie est rédigée sous forme de projet de publication.

Dans le **quatrième chapitre**, nous nous focalisons sur la compréhension au niveau moléculaire de la spécificité des lipases CALB et PCL envers le flavonoïde aglycone quercétine. Pour ce faire, une combinaison de simulations de docking et de dynamique moléculaire est appliquée pour étudier les orientations et les différences de stabilisation des substrats (acétate et quercétine) dans les cavités des deux enzymes. Cette partie fait l'objet d'une publication parue dans le *Journal of Biotechnology* (v.156, p. 203-210, 2011).

Dans le **cinquième chapitre**, l'impact de mutations simples de deux résidus hydrophobes du site actif de la CALB sur la structure de celle-ci et, sur le positionnement et l'orientation de la quercétine dans la cavité de l'enzyme est investigué. Dans un premier temps, une étude structurale des variants simples est conduite par simulation de dynamique moléculaire (DM). Ensuite, la méthodologie de docking et de DM du quatrième chapitre est reprise pour déterminer l'influence de la substitution des deux résidus sur les complexes enzyme-substrat. Cette partie des résultats est rédigée sous forme de projet de publication. En outre, une étude structurale complémentaire par DM est effectuée sur les mutations doubles des deux résidus. Enfin, le positionnement et l'orientation de la quercétine sont évalués par docking dans un variant additionnel.

Dans le **sixième chapitre**, nous examinons le positionnement et l'orientation de la quercétine ainsi que de ses dérivés méthylés et acétylés dans les cavités de la CALB native et acétylée. En outre, l'affinité de liaison de ces différents substrats pour les sites actifs de la CALB native et acétylée est estimée. Cette dernière partie est rédigée sous forme de projet de publication.

Introduction Générale

1. Kyriakou, E., et al., *Unexpected enzyme-catalyzed regioselective acylation of flavonoid aglycones and rapid product screening*. *Organic & Biomolecular Chemistry*, 2012. **10**(9): p. 1739-1742.

Chapitre I. Étude Bibliographique

1. Les principales approches de la modélisation moléculaire

La modélisation moléculaire (MM) fait généralement référence à une collection de méthodes utilisées pour étudier le comportement de systèmes de molécules. Un des principaux atouts de ces méthodes est qu'il fournisse des informations structurales et dynamiques avec un niveau de détail qui est difficile, voire impossible, d'obtenir par des techniques expérimentales. La modélisation moléculaire est bien adaptée pour l'interprétation à l'échelle moléculaire de phénomènes fondamentaux d'interactions physiques, et elle est donc un outil utile pour comprendre le comportement des protéines en présence d'un substrat.

Cette section donne une description de trois méthodes de modélisation moléculaire, à savoir les méthodes de simulation moléculaire classique, de docking moléculaire et de calcul d'énergie libre de liaison, qui ont été utilisées au cours de ce travail.

1.1. Simulation moléculaire classique

1.1.1. La mécanique moléculaire

La mécanique moléculaire est une méthode d'analyse conformationnelle basée sur quatre simplifications pour la description du système modélisé :

- (i) Les atomes sont considérés comme des noyaux réduits à des masses ponctuelles dotés d'un rayon et d'une charge électrique ponctuelle (généralement dérivés de calculs quantiques ou de détermination expérimentales)
- (ii) Les liaisons covalentes entre les atomes constitutifs d'une molécule sont modélisées comme des ressorts présentant une distance d'équilibre (déterminés soit par des calculs de mécanique quantique soit par des méthodes expérimentales)
- (iii) Les atomes non liés peuvent interagir par l'intermédiaire d'interactions électrostatiques et de van der Waals
- (iv) Les paramètres de petites molécules sont transférables à des molécules plus grandes.

Ces simplifications permettent d'exprimer l'énergie potentielle comme une fonction des coordonnées nucléaires d'un système moléculaire à l'aide de paramètres géométriques. La forme fonctionnelle de l'énergie est désignée sous l'appellation « champ de forces ».


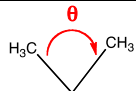
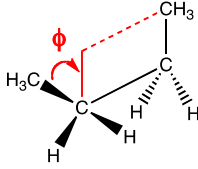
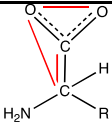
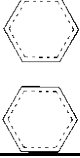
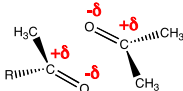
1.1.2. Le champ de forces

Dans les simulations moléculaires classiques, un modèle empirique dénommé *champ de forces* est employé pour décrire les interactions entre les atomes. Dans le cadre de simulations d'atomes, chaque atome dans le système étudié est représenté par sa position dans l'espace tridimensionnel. Le champ de forces est une expression de l'énergie configurationnelle totale du système en termes de coordonnées d'atomes. Sa forme fonctionnelle est représentée comme suit (Eq. 1.1) [1]:

$$U_{\text{configurationnelle}} = U_{\text{liaison}} + U_{\text{angle}} + U_{\text{dièdre}} + U_{\text{impropre}} + U_{\text{vdW}} + U_{\text{elec}} \quad (1.1)$$

où les termes désignent respectivement l'énergie configurationnelle (ou plus communément appelée énergie potentielle) résultant de l'étirement des liaisons chimiques, de la flexion des angles, de la torsion des angles dièdres, de la variation de l'angle impropre (écart par rapport à la planéité des liaisons sp²), des interactions de Lennard-Jones et électrostatiques. Les quatre premiers termes sont considérés comme des termes intramoléculaires (également appelés termes liés) car ils représentent les interactions entre des atomes appartenant à la même molécule et s'assure que leur géométrie moléculaire est maintenue. Les deux autres termes sont désignés comme des termes intermoléculaires ou non liés qui représentent les interactions entre des atomes de molécules différentes et les interactions entre atomes appartenant à la même molécule mais séparés par au moins quatre liaisons. L'expression des différents termes du champ de forces est détaillée et illustrée dans le Tableau I.1.

Tableau I.1. Les différents termes énergétiques d'un champ de forces classique.

Terme énergétique	Illustration	Équation
Longueur des liaisons		$U_{liaison} = \sum_{liaisons} k_r (r - r_0)^2$
Angle des liaisons		$U_{angle} = \sum_{angles} k_\theta (\theta - \theta_0)^2$
Interactions intramoléculaires	Angles dièdres 	$U_{torsion} = \sum_{dièdres} k_\phi [1 + \cos(n\phi - \delta)]$
	Angles impropres 	$U_{impropres} = \sum_{impropres} k_\omega (\omega - \omega_0)^2$
Interactions intermoléculaires	van der Waals (vdW) 	$U_{vdW} = \sum_{i,j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$
	Électrostatique 	$U_{elec} = \sum_{i,j} \frac{q_i q_j}{\epsilon r_{ij}}$

Plusieurs champs de forces ont été développés pour les simulations des biomolécules. Des exemples typiques sont AMBER (en anglais, *Assisted Model Building and Energy Refinement*) [2], CHARMM (en anglais, *Chemistry at Harvard Molecular mechanics*) [3], GROMACS (en anglais, *Groningen Machine for Chemical Simulations*) [4] et OPLS (en anglais, *Optimized Potential for Liquid Simulations*) [5]. Ces champs de forces diffèrent dans leur forme fonctionnelle ainsi que dans les paramètres utilisés pour modéliser les interactions.

Dans ce travail, nous avons employé le champ de forces CHARMM. Il a été développé ces trois dernières décennies plus particulièrement pour des molécules d'intérêt biologique comprenant les protéines, les peptides, les lipides, les acides nucléiques, les carbohydrates ainsi que de petites

molécules [6]. La forme générale de la fonction d'énergie potentielle de CHARMM est décrite plus en détail dans l'annexe A.

1.1.3. Dynamique moléculaire

Lors de simulations classiques de dynamique moléculaire (DM), le système se propage de façon déterministe par intégration numérique des équations du mouvement de Newton. Celles-ci sont élégamment exprimées sous la forme hamiltonienne [7]

$$\begin{aligned}\frac{dr_i}{dt} &= \frac{p_i}{m_i}, i=1, \dots, N \\ \frac{dp_i}{dt} &= -\nabla_{r_i} U_{total}(r_1, \dots, r_M), i=1, \dots, N\end{aligned}\quad (1.2)$$

où t désigne le temps ; r_i , p_i et m_i désignent respectivement la position, le moment et la masse de l'atome i . ∇_{r_i} est le gradient par rapport à la position de l'atome i , et N correspond au nombre d'atomes dans le système. Par intégration numérique de l'équation 1.2, on obtient une trajectoire d'un système à partir de laquelle les configurations peuvent être échantillonnées pour l'analyse. La méthode standard de résolution numérique des équations du mouvement est une méthode par différences finies. L'idée générale de l'algorithme d'intégration numérique est la suivante : à partir des coordonnées spatiales et des vitesses initiales au temps t , les positions et les vitesses sont calculées au temps $t + \Delta t$ (où Δt est le pas d'intégration). En répétant itérativement la procédure, il en résulte une *trajectoire* qui spécifie comment les positions et les vitesses des particules varient avec le temps dans le système. Parmi les algorithmes intégrateurs, les algorithmes de Verlet et de Verlet-Leapfrog sont les plus souvent implémentés dans les logiciels. Une description des paramètres de mise en œuvre d'une dynamique moléculaire est détaillée dans le chapitre II.

1.2. Docking moléculaire

Le docking moléculaire ou amarrage moléculaire est une méthode de calcul qui prédit la structure de complexe intermoléculaire formé entre un récepteur et un ligand. Un récepteur est généralement une protéine ou un acide nucléique (ADN ou ARN) et le ligand est soit une petite molécule soit une autre protéine. La plupart des algorithmes de docking sont capables de générer un grand nombre de structures et de ce fait, ils requièrent un moyen de classification afin d'identifier les structures d'intérêt. Le docking, qui fait référence au docking protéine-ligand, comporte deux éléments principaux, l'algorithme de docking qui positionne le ligand dans le récepteur et, les fonctions de score qui évaluent et classifient les structures plausibles des complexes intermoléculaires.

1.2.1. Algorithmes de docking

Les algorithmes de docking ont été développés pour rechercher rapidement et efficacement les modes d'associations protéine-ligand les plus favorables. Cependant, un système biologique constitué d'un ligand et d'un récepteur protéique comporte un nombre important de degrés de liberté à

explorer (les degrés de liberté conformationnelle de chaque molécule et ceux d'une molécule par rapport à l'autre). De ce fait, les protocoles de docking ont du adopter des stratégies pour réduire la dépendance du temps de calcul à la taille du système. Les algorithmes de docking peuvent être classés en deux grands groupes selon la flexibilité du ligand et/ou du récepteur: le docking de corps rigides (en anglais, *rigid-body docking*) et le docking flexible. [8]

- Les **méthodes de docking de corps rigides** sont des algorithmes simples qui considèrent le ligand et la protéine comme deux corps rigides. Cela limite la spécificité et la précision des résultats étant donné qu'elles tiennent compte uniquement de la complémentarité géométrique des deux molécules [9].
- Les **méthodes de docking flexibles** explorent la flexibilité du ligand ou du récepteur voire les deux molécules simultanément avec un temps de calcul supérieur aux méthodes précédentes.
 - Les procédures examinant la flexibilité du ligand tout en considérant la protéine comme un corps rigide sont les plus répandues [10]. Les algorithmes de modélisation moléculaire, de DM, de forme, systématique et stochastique sont employés pour générer les conformères du ligand et les positionner dans le récepteur. Plusieurs articles de revue décrivant le principe, les points forts et les limitations de ces méthodes sont disponibles [11-14]. Toutefois, cette classification doit être considérée avec précaution, puisqu'un grand nombre d'algorithmes combinent plus d'une méthode pour la génération et l'échantillonnage de conformères du ligand. Dans la plupart des cas, l'utilisation d'algorithmes considérant la protéine comme un corps rigide mène à des résultats intéressants, principalement quand la protéine a une flexibilité limitée. En effet, dans de tel cas, la structure cristallographique peut être considérée comme représentative de son état dans un environnement naturel, ce qui accroît les probabilités de simuler correctement l'association des ligands [12, 15, 16].
 - Les procédures tenant compte de la flexibilité de la protéine partiellement ou totalement sont moins souvent utilisées en raison d'un gain de précision inférieur à l'augmentation du temps de calcul [9, 17, 18]. Toutefois, certaines macromolécules subissent des changements conformationnels ou des ajustements induits lors de la liaison du ligand afin de maximiser les interactions favorables. Dans certains cas, l'incorporation de la flexibilité du récepteur conduit à des modes d'associations plus réalistes que les procédures précédentes [9].

1.2.2. Les fonctions de score

Dans la dernière étape d'un protocole de docking, les poses issues de l'échantillonnage doivent être évaluées en termes d'énergie d'interaction afin de récupérer les complexes susceptibles de reproduire au mieux le mode d'association réel. Cette étape s'effectue par une estimation quantitative de l'énergie libre de liaison. Les associations entre ligands et récepteurs sont une combinaison d'effets enthalpiques et entropiques. Cela comprend les interactions électrostatiques incluant les liaisons hydrogène, l'attraction et la répulsion coulombienne, les interactions de van der Waals, les

effets de solvation et de désolvation. La formation du complexe est favorable si la variation d'énergie libre de liaison est négative ($\Delta G_{\text{liaison}} < 0$). Cependant, l'évaluation de l'énergie libre des complexes est une tâche coûteuse d'un point de vue informatique, ce qui limite son utilisation en routine. De ce fait, $\Delta G_{\text{liaison}}$ est estimé en utilisant des méthodes simplifiées désignées comme *fonctions de score*. Les fonctions de score ne sont pas censées fournir une estimation précise de l'énergie libre de liaison mais leurs objectifs est de caractériser le plus précisément les poses du ligand de façon à ce qu'elles ressemblent aux modes de liaison expérimentaux.

Les fonctions de score ont été classifiées en trois catégories principales: les fonctions de score basées sur le champ de forces (en anglais, *force-field based*), les fonctions de score empiriques et les fonctions de score basées sur la connaissance (en anglais, *knowledge based*).

- Les **fonctions de score basées sur le champ de forces** ont été développées en s'appuyant sur les énergies d'interactions d'un champ de forces. Une expression complète fournit l'énergie globale du système c'est-à-dire les interactions liées et non liées. La limitation majeure de ces fonctions de score est la négligence de l'effet de désolvation et de la perte d'entropie configurationnelle ou leur introduction dans le score final par des heuristiques. [19, 20]
- Les **fonctions de score empiriques** estiment l'affinité de liaison d'un complexe sur la base de la somme d'un ensemble de termes d'énergies indépendants et pondérés :

$$\Delta G = \sum_i W_i \Delta G_i \quad (1.3)$$

où ΔG_i représente les différents termes d'énergie comme l'énergies de vdW, électrostatique, liaison hydrogène, désolvation, entropie, hydrophobicité, etc. Les coefficients W_i correspondants sont déterminés en ajustant les données de l'affinité de liaison d'un ensemble de complexes protéine-ligand dont les structures tridimensionnelles sont connues. Ces fonctions de score sont calculées rapidement en raison de la simplicité des termes d'énergie. La principale limitation de ce type de méthode est l'efficacité de leur pouvoir prédictif en présence de complexes récepteur-ligand radicalement différents de ceux employés dans l'ensemble de référence. [21, 22]

- Les **fonctions de score basées sur la connaissance** emploient des potentiels d'énergie dérivés de l'analyse statistique de l'information structurelle intégrée dans les structures protéiques déterminées expérimentalement. Les potentiels interagissant entre une paire d'atomes sont obtenus à partir de leur fréquence relative dans une base de données par rapport à un état de référence. Dans les études protéine-ligand, les potentiels sont calculés par la relation inverse de Boltzmann :

$$w(r) = -k_B T \ln[g(r)]$$

avec $g(r) = \rho(r) / \rho^*(r)$ (1.4)

où k_B est la constante de Boltzmann, T est la température absolue du système, $\rho(r)$ est la densité en nombre des paires d'atomes protéine-ligand à la distance r , et $\rho^*(r)$ est la densité de paires d'atomes dans un état de référence où les interactions interatomiques sont nuls.

[23-25]. Cette méthode offre un bon compromis entre précision et rapidité. En effet, les potentiels sont extraits à partir de structures plutôt que de tenter de reproduire les affinités connues par ajustement. De plus, la base d'information structurale est vaste et diversifiée permettant d'obtenir des fonctions de score robustes et relativement insensibles à l'ensemble de base. [22]

En dépit d'un grand nombre de fonctions de score qui ont été développées, aucune n'est parfaite en termes de précision et d'applicabilité générale. Chaque fonction de score a ses avantages et ses limites. Pour tirer parti des avantages et équilibrer les déficiences des différentes fonctions de score, la technique de consensus de score a été mise en place pour améliorer la probabilité de trouver des solutions correctes en combinant les diverses fonctions de score [26]. L'étape critique dans le consensus de score est la conception d'une stratégie appropriée de consensus des scores individuels de telle sorte que les poses correctes peuvent être discriminées [27, 28].

1.3. Énergie libre de liaison

Le calcul des interactions énergétiques entre un récepteur et un ligand est d'un intérêt fondamental. Pour des applications telles que la conception de protéines ou de principes actifs la sélection de nouveaux composés ou encore la détermination de constantes d'inhibition, la connaissance de l'énergie libre de liaison est précieuse. Diverses méthodes ont été développées pour estimer cette énergie. Ces techniques varient selon la précision et le coût en terme de calcul. En effet, les méthodes fournissant des valeurs rigoureuses et précises telles que la perturbation de l'énergie libre (en anglais, *free energy perturbation*, FEP) et l'intégration thermodynamique (en anglais, *thermodynamic integration*, TI) sont onéreuses en temps de calcul. Néanmoins, plusieurs méthodes alternatives induisant les méthodes MM-PBSA (en anglais, *molecular mechanics Poisson-Boltzmann surface area*) ou LIE (en anglais, *linear interaction energy*) ont été développées pour estimer de façon plus rapide et plus pratique l'énergie libre de liaison. Plus particulièrement, l'approche MM-PBSA est un compromis entre précision et temps de calcul.

L'aspect thermodynamique de l'énergie libre de liaison a été décrit succinctement dans cette section. Finalement, la méthode MM-PBSA employée dans le cadre de ce travail a été abordée.

1.3.1. Aspect thermodynamique

L'association non covalente et réversible d'un récepteur (R) et d'un ligand (L) pour former un complexe récepteur-ligand (RL), dans un milieu aqueux, peut être représentée comme dans la Figure I.1.

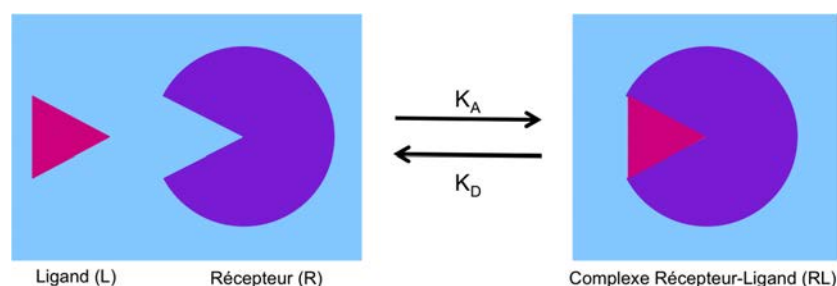


Figure I.1. Représentation schématique de l'équation suivante $R_{aq} + L_{aq} \rightleftharpoons RL_{aq}$. K_A est la constante d'association du récepteur (R) et de son ligand (L) et K_D est la constante de dissociation du complexe récepteur-ligand (RL).

A l'équilibre, cette réaction est déterminée par l'énergie libre de liaison ΔG^0 (Éq. I.5). Cette quantité est liée aux constantes d'association (K_A) et de dissociation (K_D) qui peuvent être déterminées expérimentalement (Éq. I.6). Elle contient une composante enthalpique (ΔH^0) et entropique ($T\Delta S^0$) où T est la température absolue (Éq. I.5). ΔH est le reflet de la variation de l'énergie interne incluant les liaisons, les interactions, les déformations, etc. ΔS est la variation d'entropie qui est une mesure du désordre.

$$\Delta G^0 = -RT \ln K_A = \Delta H^0 - T\Delta S^0 \quad (I.5)$$

$$K_A = K_D^{-1} = \frac{[RL]}{[R][L]} \quad (I.6)$$

Le calcul de l'énergie libre de liaison se fait en se basant sur le cycle thermodynamique donné dans la Figure I.2. D'après ce schéma, l'énergie libre de liaison $\Delta G_{liaison,Solv}^0$ peut être calculée par l'équation I.7.

$$\Delta G_{liaison,Solv}^0 = \Delta G_{liaison,Gaz}^0 + \Delta G_{Solv,RL}^0 - (\Delta G_{Solv,L}^0 - \Delta G_{Solv,R}^0) \quad (I.7)$$

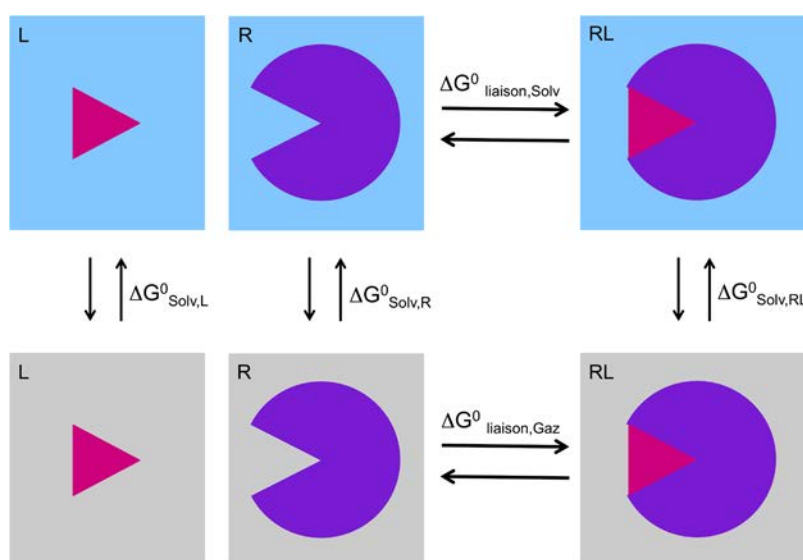


Figure I.2. Cycle thermodynamique pour le calcul de l'énergie libre de liaison entre un récepteur (R) et un ligand (L). Les composantes $\Delta G_{Solv,L}^0$, $\Delta G_{Solv,R}^0$ et $\Delta G_{Solv,RL}^0$ correspondent à l'énergie libre de solvation respectivement pour le ligand, le récepteur et le complexe.

1.3.2. La méthode MM-PBSA

La méthode MM-PBSA estime l'énergie libre d'association de complexes moléculaires [29]. Cette approche est basée sur une analyse des trajectoires de dynamique moléculaire en utilisant une approche combinée de mécanique moléculaire et de modèle de solvant continu pour estimer l'affinité de liaison des ligands. Elle évalue l'énergie libre selon l'équation 1.7 décrite dans le cycle thermodynamique de la Figure 1.2. Le terme $\Delta G_{liaison,Gaz}^0$ est l'énergie d'interaction entre le ligand et le récepteur en phase gazeuse. Elle est déterminée à partir de l'énergie de la mécanique moléculaire incluant les termes intra- et inter-moléculaires d'un champ de forces classique (Éq 1.1). $\Delta G_{Solv,L}^0$, $\Delta G_{Solv,R}^0$, $\Delta G_{Solv,RL}^0$ sont les contributions à l'énergie libre de solvation du ligand, du récepteur et du complexe récepteur-ligand. Elles comportent une composante électrostatique résolue numériquement par l'équation de Poisson-Boltzmann (PB) (ΔG_{PB}) et, une composante hydrophobique qui tient compte de la fraction non polaire du solvant (ΔG_{SA}) c'est-à-dire que :

$$\Delta G_{Solv} = \Delta G_{PBSA} = \Delta G_{PB} + \Delta G_{SA} \quad (1.8)$$

Ces différents termes sont obtenus en effectuant la moyenne sur un échantillon d'images instantanées (en anglais, *snapshot*) issus d'une trajectoire de dynamique moléculaire du système.

La méthode MM-PBSA a été utilisée initialement pour étudier la stabilité de fragments d'ADN et d'ARN et s'est répandue cette dernière décennie pour estimer les énergies libres de liaison de ligand [29-31]. Cette méthode peut être utilisée dans des études de DM de complexes protéine-ligand où l'énergie libre est déterminée à partir d'un échantillonnage d'images instantanées issues d'une trajectoire [31-33] mais également dans des applications de criblages virtuels ou de docking pour affiner la classification des poses [30, 34, 35]. Dans ce cas, la technique MM-PBSA, qui prend en compte l'énergie de solvation, est plus rigoureuse que les fonctions de score basées sur la connaissance ou empiriques. Plusieurs études ont montré qu'elle est efficace pour identifier les modes de liaisons correctes et classer les inhibiteurs de cibles spécifiques [30, 34, 36].

2. La quercétine

2.1. Généralités

La quercétine [2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-1-benzopyran-4-one ou 3,5,7,3',4'-pentahydroxyflavone] est un flavonoïde polyphénolique appartenant à la classe des flavonols. La structure chimique de la quercétine comporte un noyau commun flavone composé de deux cycles benzéniques A et B reliés entre eux par un pyrane hétérocycle C. La Figure 1.3 présente la nomenclature des cycles et la numérotation des hydroxyles de ce flavonoïde aglycone utilisées dans le cadre de ce travail.

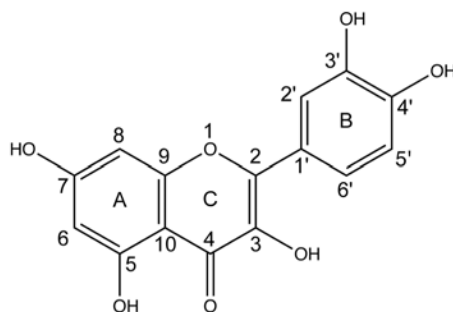


Figure I.3. Nomenclature des cycles et numérotation systématique des hydroxyles de la quercétine.

La quercétine est l'un des composés les plus représentatifs de la famille des flavonoïdes et elle est consommée quotidiennement par l'être humain vu sa distribution étendue dans l'alimentation. Ce flavonol se trouve principalement sous forme glycosylée dans une grande variété de fruits, de légumes et de boissons. Dans les plantes, différents types de sucres incluant le glucose, le galactose, le rhamnose, le xylose ou le rutinose, sont liés aux groupements hydroxyles de la molécule aglycone par des liaisons glycosidiques. Les principales cibles de la glycosylation sont les groupements hydroxyles sur les positions 3, 7 et 4' [37]. Le *Tableau I.2* reprend une liste non exhaustive de la concentration en quercétine dans divers aliments. Plusieurs études ont parfois été effectuées pour un même aliment ce qui correspond aux différentes valeurs reprises dans le tableau. Il a été montré que ces variations de teneur en quercétine dépendent de facteurs extrinsèques tels que des facteurs géographiques, climatiques mais également des ressources dont les plantes disposent, du degré de maturation de la plante, de la fraction du fruit étudié ou encore de la préparation des aliments susceptibles de modifier les concentrations analysées.

Tableau I.2. Concentration en quercétine dans divers fruits et légumes, et dans diverses boissons [38].

Fruits, légumes et boissons	Concentration en quercétine*
Abricot	2.5 ; 2.6 ; 5.3
Canneberge	149
Ciboulette	10.4 ; 30
Fèves	2 ; 134
Jus de tomate	1.3 ; 1.1-1.3
Myrtille	10.5-16
Oignon	28-49 ; 18-63 ; 41
Pomme	2-7 ; 2 ; 10-26
Raisin noir	1.5 ; 3.7
Sureau	10.5-24
Thé noir	1.7-2.5
Thé vert	1.4-2.3 ; 0.11
Vin rouge	0.4-1.6 ; 0-0.2

*La concentration est exprimée en mg/100 g dans les fruits et légumes et en mg/100 ml dans les boissons

Les principales propriétés biologiques, les propriétés physico-chimiques ainsi que la relation structure-activité de la quercétine sont décrites dans les sections suivantes.

2.2. Propriétés biologiques de la quercétine

Au cours des dernières années, les recherches concernant la bioactivité de la quercétine se sont intensifiées dans les domaines académiques et industriels s'intéressant aux secteurs des cosmétiques, des produits agro-alimentaires et pharmaceutiques. En effet, les propriétés anti-oxydantes de ce flavonoïde sont à l'origine de plusieurs fonctions biologiques attribuées à la quercétine. En raison du nombre important d'études sur la quercétine, seules les propriétés les plus citées dans la littérature sont brièvement décrites dans les sections suivantes.

2.2.1. Propriétés anti-oxydantes

D'après Halliwell [39], un anti-oxydant est une substance qui lorsqu'elle est présente en faible concentration, comparée à celle d'un substrat oxydable (c'est-à-dire tout type de molécules rencontrées *in vivo*), empêche ou réduit considérablement l'oxydation de ce substrat. Il a été démontré dans de nombreuses études *in vitro* que la quercétine a un puissant pouvoir anti-oxydant comparé à ses analogues [37, 40-42].

Les bases moléculaires de l'activité anti-oxydante aussi bien pour ce composé aglycone que pour les flavonoïdes en général peuvent se produire selon trois mécanismes distincts [43] :

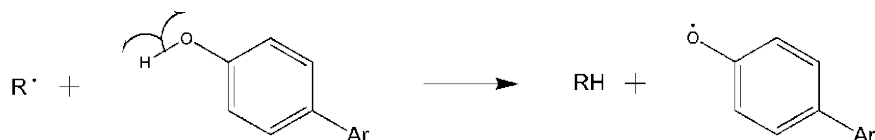
(i) La désactivation des radicaux libres selon les deux mécanismes suivants :

Les radicaux libres (R^\cdot) sont des molécules réactives dues à la présence d'un ou plusieurs électron(s) non apparié(s). Les radicaux libres sont inactivés soit par un transfert d'un atome d'hydrogène (1) soit par un transfert d'électron (2) ce qui mène à la formation de radicaux libres stables ArO^\cdot et $ArOH^{+\cdot}$ (Figure 1.4). Leurs structures aromatiques permettent une délocalisation de l'électron à travers la molécule résultant en une stabilisation de la forme radicalaire [42].

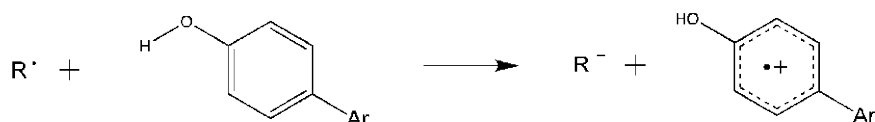


Il a été démontré que la quercétine est un puissant anti-oxydant *in vitro* [42] et notamment un puissant désactivateur d'espèce réactive de l'oxygène (en anglais, *reactive oxygen species*, ROS). Les espèces réactives de l'oxygène importantes physiologiquement comprennent les radicaux de l'oxygène (radical hydroxyle, HO^\cdot ; superoxyde, $O_2^{\cdot-}$) et les dérivés non radicalaires de l'oxygène (peroxyde d'hydrogène, H_2O_2 ; ozone, O_3 ; acide hypochloreux, $HOCl$) qui peuvent facilement générer des radicaux libres.

1. Transfert d'un atome d'hydrogène



2. Transfert d'un électron



3. Chélation des métaux

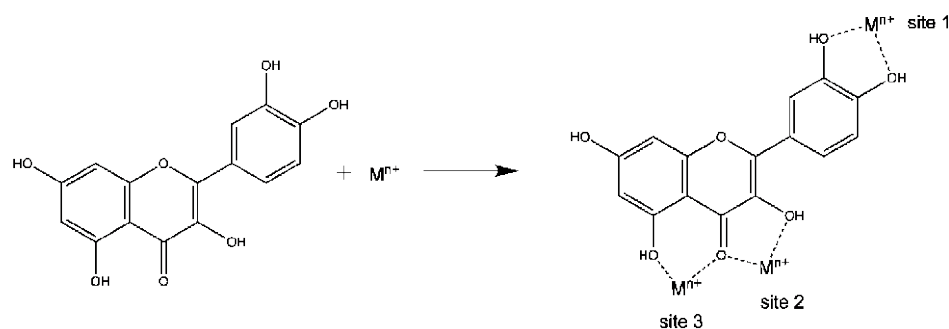


Figure I.4. Mécanismes de l'activité anti-oxydante de la quercétine.

(ii) La chélation de métaux libres

Les flavonoïdes peuvent chélater des ions de métaux de transition (Fe^{2+} , Fe^{3+} , Cu^{2+} , Zn^{2+} , Al^{3+} , Cr^{3+} , Pb^{2+} , Co^{2+}) conduisant à des composés complexés stables (Figure I.4). Les métaux chélatés sont alors dans l'incapacité de prendre part à des réactions génératrices de radicaux libres. La quercétine avec ses trois sites de chélation est un chélateur de métaux très efficace. Plus spécifiquement, la complexation de la quercétine avec des métaux tels que du cuivre (Cu^{3+}) ou du fer (Fe^{2+}) enrayer la production de radicaux hydroxyles formés lors de réactions de Fenton. Brièvement, la réaction de Fenton entraîne l'accumulation de radicaux libres dans les tissus nerveux et ceux-ci sont impliqués dans le développement de maladies telles que la maladie d'Alzheimer ou de la chorée d'Huntington. [44, 45].

2.2.2. Propriétés anti-inflammatoires

La quercétine est connue pour posséder de fortes capacités anti-inflammatoires. Une inflammation peut être reliée à un stress oxydatif. Le stress oxydatif se définit comme étant un déséquilibre entre la production de substances réactives et la protection contre celles-ci [46]. Ce déséquilibre résulte d'un accroissement des dommages oxydatifs causés soit par une surproduction de radicaux libres et d'autres espèces réactives soit par une déficience du système de défense antioxydant endogène [46, 47]. La conséquence est l'apparition de dégâts souvent irréversibles pour les cellules [45]. Les espèces réactives ne sont pas seulement impliquées dans la survenue du stress oxydatif, mais également dans la promotion de processus inflammatoires. Par conséquent, la

désactivation des substances réactives nocives pour l'organisme ne serait pas seulement de prévenir la survenue du stress oxydatif, mais aussi aider à atténuer l'inflammation [41].

2.2.3. Propriétés anti-cancérigènes

Le potentiel anti-cancérigène de la quercétine a été attribué à différents mécanismes incluant l'activité anti-oxydante, l'inhibition d'enzymes activant des substances cancérigènes, la modification des voies de transduction de signaux et les interactions avec des récepteurs et autres protéines [48, 49].

2.2.4. Autres propriétés

En plus des propriétés déjà précitées, il a également été démontré *in vitro* que la quercétine possède des propriétés anti-fibrotique, anti-coagulante, anti-bactérienne, anti-athérogène, anti-hypertenseur et anti-proliférative [37]. Ce flavonoïde aglycone a également une activité inhibitrice vis-à-vis de certaines enzymes notamment des oxydoréductases qui intègrent des espèces radicalaires au cours de leurs cycles catalytiques (peroxydase, lipoxigénase, cyclo-oxygénases, monooxygénases, xanthine-oxydases, protéine-kinases et etc.). Dans l'ensemble, ces études indiquent que la quercétine peut exercer ses bienfaits pour la santé en modulant des effets néfastes pour l'organisme.

2.3. Relation structure-activité de la quercétine

Des études de la relation structure-activité montrent que les propriétés biologiques de la quercétine sont corrélées à sa structure chimique.

2.3.1. Propriétés anti-oxydantes

Les propriétés anti-oxydantes de la quercétine peuvent être assignées à la présence des éléments structuraux suivants [37, 43, 45, 50] :

- L'occurrence de plusieurs groupes hydroxyles fixés au noyau aromatique
- Le fragment catéchol avec un arrangement des hydroxyles en conformation ortho-dihydroxy (OH_3 et OH_4) : la déprotonation de l'oxygène en position 4' conduit à un anion stable dans lequel l'établissement d'une liaison hydrogène est le principal effet stabilisateur. La réaction est schématisée dans la Figure 1.5.
- La double liaison $\text{C}_2=\text{C}_3$ dans le cycle C conjuguée avec le groupe 4-céto : ce motif a une grande importance dans la délocalisation des électrons non appariés et dans les effets de résonance
- Le groupement fonctionnel 4-céto dont les paires d'électrons libres forment des liaisons hydrogène intramoléculaires avec les groupements hydroxyles aux positions 3 et 5

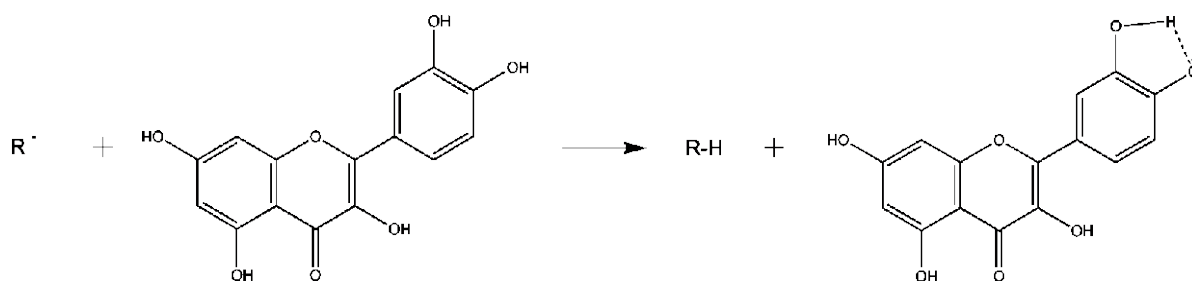


Figure 1.5. Réaction de transfert d'atome d'hydrogène entre la quercétine et un radical libre, R. La déprotonation de l'hydroxyle en position 4' mène à une forme radicalaire de la quercétine qui est stabilisée par la formation d'une liaison H avec le groupement OH adjacent.

2.3.2. Chélation des métaux

La quercétine possède trois sites possibles de chélation qui sont indiqués dans la Figure 1.4 [51, 52] :

- La fraction catéchol comportant deux groupements hydroxyles aux positions 3' et 4'
- Le groupement hydroxyle en C_3 et le 4-céto
- Le groupement hydroxyle en C_5 et le 4-céto

Les sites préférentiellement chélatés dépendent de l'ion métallique. Il a été démontré que le groupe catéchol présente un plus grand pouvoir de complexation envers le plomb (Pb^{2+}) tandis que l'aluminium (Al^{3+}) se coordonne préférentiellement à la partie 3-OH/4-céto. Par ailleurs, le complexe quercétine/ Cr^{3+} a une capacité anti-oxydante supérieure à la quercétine seule car la chélation accentue le transfert d'électron.

2.4. Propriétés physico-chimiques de la quercétine

2.4.1. Acidité des groupements hydroxyles de la quercétine

Les flavonoïdes déprotonés sont de meilleurs donneurs d'électrons et désactivateurs de radicaux que les molécules neutres. L'exploration expérimentale et théorique de l'acidité des groupements hydroxyles contribue à la compréhension de la relation structure-activité. Des valeurs expérimentales et théoriques de pK_a sont données dans le Tableau 1.3 [43, 50, 53, 54].

Les données expérimentales disponibles pour le pK_a de la quercétine permettent de définir des intervalles: pK_{a1} [5.5-5.7] ; pK_{a2} [6.9-7.7] ; pK_{a3} [8.0-8.7] ; pK_{a4} [9.3-9.9] et pK_{a5} [11.0-11.1]. Ces variations peuvent être attribuées à une instabilité des produits à l'air et à une contamination par des produits d'oxydation. Cependant, les méthodes de titrations spectrophotométrique et potentiométrique ne distinguent pas l'acidité d'un groupement OH particulier. L'attribution des valeurs de pK_a aux groupements OH, donnée dans le Tableau 1.3, est effectuée par analogie avec des flavonoïdes monohydroxylés ou d'autres systèmes modèles. Des divergences sont constatées pour l'affectation du groupement 3-OH qui est soit le plus acide soit le moins acide selon l'étude. D'après ces différentes études, les conclusions peuvent être synthétisées comme suit :

- La formation de liaisons hydrogène intramoléculaires engendre une diminution de l'acidité des groupements OH donneurs de liaisons H (5-OH et 3-OH)

- Le groupement le moins acide de la quercétine est le 5-OH car la génération de son anion entraîne la disparition d'une liaison hydrogène intramoléculaire, ce qui introduit une répulsion électronique entre la charge négative de l'oxygène déprotoné et la paire d'électron libre de l'oxygène du carbonyle du cycle C.
- L'acidité du groupe 3',4'-dihydroxy n'est pas plus forte que l'acidité du 7-OH de la quercétine.

Tableau I.3. Valeurs de pKa déterminées expérimentalement ou théoriquement et attribution au différents groupements hydroxyles de la quercétine.

Méthodes		pKa ₁	pKa ₂	pKa ₃	pKa ₄	pKa ₅	Réf.
Expérimentale	Potentiométrie	7.7	8.77	9.81	-	-	[55]
		5.54	6.95	8.21	9.77	-	[54]
		catéchol	catéchol	3-OH		-	
	Spectrophotométrie	5.7	7.1	8.0	9.9	11.0	[54]
		-	7.7	8.7	9.8	-	[56]
		-	3-OH	7-OH	4'-OH	-	
-		-	8.45 ± 0.1	9.31 ± 0.2	11.12 ± 0.2	[50]	
	-	-	7-OH	4'-OH	3-OH		
Théorique	B3LYP/6-311++G** ^a	-9.8	-8.5	-7.4	-7.4	-5.6	[43]
		7-OH	3'-OH	4'-OH	3-OH	5-OH	

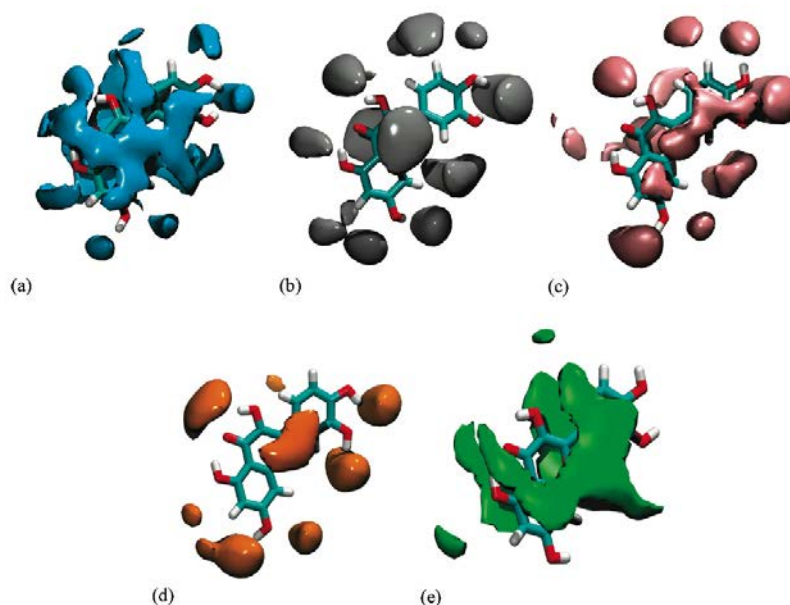
^a les valeurs ont été calculées à partir de l'énergie libre totale selon la relation : $pK_a = \Delta G / \ln(10)RT$ où R est la constante de gaz et T la température absolue. Elles sont exprimées en kcal/mol.

2.4.2. Solubilité

La solubilité de la quercétine dans différents milieux est généralement faible. Dans les solvants organiques d'usage courant en bioprocédé, leurs solubilités sont significativement accrues par rapport à l'eau mais restent encore très limitées. A titre d'exemple, les valeurs de solubilité obtenues expérimentalement dans l'eau, le *tert*-amyl alcool, l'acétone et l'acétonitrile sont données dans le *Tableau I.4*. Elles varient entre 0.03 mM dans l'eau et 96 mM dans le propan-2-ol. D'autre part, le coefficient de partage eau-octanol de la quercétine ($\log P = 1.82$) indique qu'elle est moyennement hydrophobe [57]. D'après la Figure I.6, qui montre la carte de répartition du solvant autour de la quercétine, il apparaît que sa solubilité est améliorée dans l'acétone due aux interactions hydrophobes plus prononcées, et dans le *tert*-amyl alcool grâce aux interactions hydrophiles [58]. Les auteurs ont conclu et vérifié qu'un accroissement de la solubilité serait possible dans un solvant comportant une fraction hydrophile et une fraction hydrophobe tel que le propan-2-ol (*Tableau I.4*). Malgré que la quercétine expose un comportement amphipathique dû aux cycles aromatiques formant la partie hydrophobe et les groupements hydroxyles constituant la partie polaire, elle manifeste une solubilité et une stabilité incomplète dans les phases lipophile et hydrophile.

Tableau I.4. Solubilité de la quercétine dans différents solvants [58, 59].

Solvant	Température (°C)	Solubilité (mM)
Eau	50	0.03
Acétonitrile	50	5.4
<i>tert</i> -amyl alcool	50	67
Acétone	50	80
Propan-2-ol	50	96

Figure I.6. La distribution de probabilité de la densité de solvants autour de la quercétine : (a) l'eau, (b) l'acétonitrile, (c) l'acétone, (d) le *tert*-amyl alcool, (e) le chloroforme [58]

2.5. Études des interactions de la quercétine avec des enzymes par modélisation moléculaire

La quercétine est une molécule fonctionnalisée qui possède dans sa structure des cycles aromatiques, des fonctions hydroxyles (-OH) et une fonction carbonyle (-C=O). Grâce à ces éléments structuraux, cette molécule peut établir différents types d'interactions avec un large éventail d'autres molécules. En particulier, les interactions que peuvent établir ce flavonoïde avec certaines protéines expliquent, sur un plan moléculaire, une partie des activités biologiques de ce polyphénol. Pour la compréhension de ces activités, un élément-clé est la connaissance des interactions susceptibles de se former entre les acides aminés du site actif des protéines et la quercétine, que celle-ci soit inhibitrice ou substrat.

Dans cette section, les interactions moléculaires susceptibles de s'établir entre les flavonoïdes en général et les acides aminés sont présentées. Ensuite, quelques études récentes sont synthétisées pour évaluer les modes d'interactions répertoriés pour la quercétine.

2.5.1. Interactions entre les flavonoïdes et les acides aminés

Codomiu-Hernandez et al. [60] ont appliqué des calculs quantiques semi-empiriques (AM1 et PM3) pour caractériser les interactions individuelles entre deux flavonoïdes modèles (la catéchine et la robinétinédine) et les 20 acides aminés des protéines. Ces auteurs ont proposé un ordre théorique d'interaction (TAO, de l'anglais *theoretical affinity order*) en analysant les valeurs d'énergie, de distances et d'angles d'interaction par la méthode de l'hypersurface des minima multiples (MMH) :

Lys, Glu, Asp > His, Arg > Trp, Leu, Gln, Asn, Phe > les autres acides aminés

Dans une étude postérieure [61], ce TAO a été confronté à des données expérimentales issues d'une part, de la PDB pour identifier les modes d'interaction dans plusieurs complexes protéine-flavonoïde ; et d'autre part, de la littérature pour collecter dans les réactions d'inhibition d'enzymes par les flavonoïdes, les acides aminés qui établissent des interactions directes avec les flavonoïdes. Ainsi, la capacité des flavonoïdes à former des complexes avec des protéines a été corrélée à la présence d'acides aminés hydrophiles dans le site actif des protéines.

Ces résultats sont en accord avec les informations plus générales relatives aux interactions intermoléculaires disponibles dans la littérature : des liaisons hydrogène ou d'autres interactions attractives de nature électrostatique sont susceptibles d'être formées entre les hydroxyles et carbonyles des flavonoïdes et les acides aminés hydrophiles Asp, Arg, Asn, Glu, Lys, Gln, Ser, Thr [62, 63]. Des interactions aromatiques de type σ - π , π - π , OH- π , NH- π peuvent s'établir entre les cycles aromatiques présents dans les molécules de flavonoïdes et les acides aminés à chaîne latérale aromatique (Phe, Trp, Tyr) [64]. Par ailleurs, les acides aminés aliphatiques hydrophobes peuvent éventuellement contribuer à la stabilisation des flavonoïdes, en formant des interactions plus faibles de type CH- π avec les centres aromatiques [62, 64].

2.5.2. Les modes d'interactions entre la quercétine et diverses protéines

L'évaluation des modes d'interaction entre la quercétine et les protéines est aisément accessible par modélisation moléculaire. La littérature comporte un grand nombre d'étude portant sur l'action inhibitrice de la quercétine vis-à-vis de protéines impliquées dans certaines pathologies. Cette démarche a pour but principal d'identifier les interactions intermoléculaires stabilisant les complexes protéine-quercétine. L'importance de ces interactions est souvent corrélée aux propriétés inhibitrices de la quercétine déterminées *in vitro*. Sans être exhaustif, le *Tableau 1.5* présente une synthèse de quelques-unes de ces études.

Tableau I.5. Quelques exemples de types d'interactions répertoriés dans la littérature entre la quercétine et diverses protéines.

Protéines	Activités biologiques de la quercétine	Outils de modélisation moléculaire Interactions protéine – quercétine identifiées	Réf.
Inhibition de la Métalloprotéinase matricielle (MMP)	Neuro-protection contre les accidents vasculaires cérébraux	<ul style="list-style-type: none"> • Docking (AutoDock 4.0) • Interaction avec MMP-2 <ul style="list-style-type: none"> ○ Liaisons H : 7-OH...Ala₂₁₇, 7-OH...Ala₂₂₀, 3'-OH...Ala₁₆₅, 4'-OH...Leu₁₆₄, 4'-OH...Ala₁₆₅ • Interaction avec MMP-9 <ul style="list-style-type: none"> ○ Liaisons H : 7-OH...Pro₄₁₅, 7-OH...Arg₄₂₄, 3'-OH...Pro₄₂₁ 	[65]
Inhibition de l'arginase de <i>Leishmania amazonensis</i>	Effet leishmanicide et réduction de la charge parasitaire dans les cellules infectées	<ul style="list-style-type: none"> • Docking (DOCK) • Liaisons H : 3'-OH...Asp₁₄₁, 4'-OH...Asp₁₄₁ 	[66]
Inhibition de la Glycolate oxydase (GOX)	Modulation de la synthèse excessive d'oxalate de calcium	<ul style="list-style-type: none"> • Docking (Glide 5.5.211) • Interactions avec sGOX (GOX de l'épinard) <ul style="list-style-type: none"> ○ Liaisons H : 5-OH...Tyr₂₄, 3-OH...Tyr₁₂₉ • Interactions avec hGOX (GOX de l'êtré humain) <ul style="list-style-type: none"> ○ Liaisons H : 3'-OH...Tyr₂₆, 4'-OH...His₂₆₀ ○ Interactions hydrophobiques : CH-π entre cycle B et Arg₁₆₇, π-π entre cycle A et Trp₁₁₀, π-π entre cycle C et Trp₁₁₀ 	[67]
Inhibition de la phospholipase A2	Réduction des médiateurs lipidiques pro-inflammatoires	<ul style="list-style-type: none"> • Docking (Gold 4.0) • Liaisons H : 3-OH...Cys₄₅, 3-OH...His₄₈ • Interaction ionique : 3'-OH...Ca²⁺ • Interaction hydrophobe avec Phe₅ 	[68]
Inhibition de la synthèse de métalloprotéinase matricielle (MMP)	Réduction du risque de développement et de la progression de l'athérosclérose	<ul style="list-style-type: none"> • Docking (AutoDock 4.0) et simulations de DM (champ de forces Amber ff99SB) • Interactions dans le premier mode de liaison <ul style="list-style-type: none"> ○ Liaisons H : 3-OH...Met₄₂₂, 4'-OH...Ala₄₁₇, 5-OH...Leu₂₈₈ ; 5-OH...Ala₁₈₉ • Interactions hydrophobiques : CH/π entre cycles AC Leu₁₈₈ et Val₃₉₈ • Interactions dans le second mode de liaison <ul style="list-style-type: none"> ○ Liaisons H : 3-OH...Pro₄₂₁ ; 3'-OH...Leu₁₈₈, 4'-OH...Glu₄₀₂ • Interactions hydrophobiques : CH-π entre cycle B et Leu₁₈₇, Leu₁₈₈ et Val₃₉₈, π-π entre cycles AC et Tyr₃₉₃ et Tyr₄₂₃ 	[69]
Inhibition de la 11 β -hydroxystéroïde dehydrogenase de type 1	Activité anti-diabétique par régulation de l'homéostasie du glucose	<ul style="list-style-type: none"> • Docking (AutoDock 4.0) • Liaisons H : 7-OH...Gly₂₁₆ • Interactions hydrophobiques : π-π entre cycle AC et cycle du cofacteur 	[70]

L'analyse de ces études montre que les liaisons hydrogène entre les groupements hydroxyles de la quercétine et les résidus polaires sont prédominantes. De plus, des interactions hydrophobes de type CH- π entre les cycles aromatiques de la quercétine et les chaînes aliphatiques des acides aminés hydrophobes sont également nombreuses, contrairement aux interactions de type π - π entre deux centres aromatiques qui sont plus sporadiques. Ces différentes interactions intermoléculaires sont des éléments stabilisateurs majeurs des complexes protéine-quercétine.

A travers cette étude bibliographique des propriétés physico-chimiques et biologiques de la quercétine, il s'avère que cette molécule possède des activités intéressantes mais aussi une photo-instabilité [71], une biodisponibilité restreinte [41] et est pratiquement insoluble que ce soit en milieu aqueux ou lipophile [57, 59]. Ces caractéristiques engendrent des désagréments pour son incorporation et sa stabilisation dans des formulations alimentaires, pharmaceutiques et cosmétiques. Afin de bénéficier de sa bioactivité dans des matrices hydrophobes, des stratégies de modifications structurales ont été développées pour fournir des formes lipophiles de cette molécule. Ces modifications doivent être contrôlées pour conserver le potentiel des propriétés biologiques de la quercétine. Dans le cadre de ce travail, seule la technique d'acylation de la quercétine est détaillée.

2.6. L'acylation de la quercétine

Différentes stratégies d'acylation de la quercétine ont été développées ces dernières années. L'acylation consiste à greffer une ou plusieurs chaîne(s) acyle(s) sur les fonctions hydroxyles de la molécule de quercétine. Cette réaction s'effectue par voie chimique, enzymatique ou chimio-enzymatique, pour former des esters du flavonoïde aglycone. Les données de la littérature concernant chacune de ces méthodes d'acylation sont détaillées dans les points suivants.

2.6.1. Acylation par voie chimique

Plusieurs procédés d'acylation de la quercétine via des méthodes chimiques sont recensés dans la littérature [72, 73]. La Figure I.7 schématise trois approches décrites par Picq et al. [72] pour la synthèse du 3,3',4',5,7-pentaacyl quercétine. Dans la première méthode, la quercétine est traitée avec des halogénures d'alkyle ou des sulfates d'alkyle dans de l'acétone séché avec du carbonate de potassium (K_2CO_3). Dans la deuxième méthode, la quercétine est mise en présence d'halogénures d'alkyle dans de l'eau avec de l'hydroxyde de potassium (KOH). Dans la dernière approche, l'acylation de la quercétine s'effectue avec des anhydrides d'acide dans la pyridine. En outre, le *Tableau I.6* reprend les différents dérivés acylés de la quercétine ainsi que les conditions expérimentales nécessaires à leur synthèse.

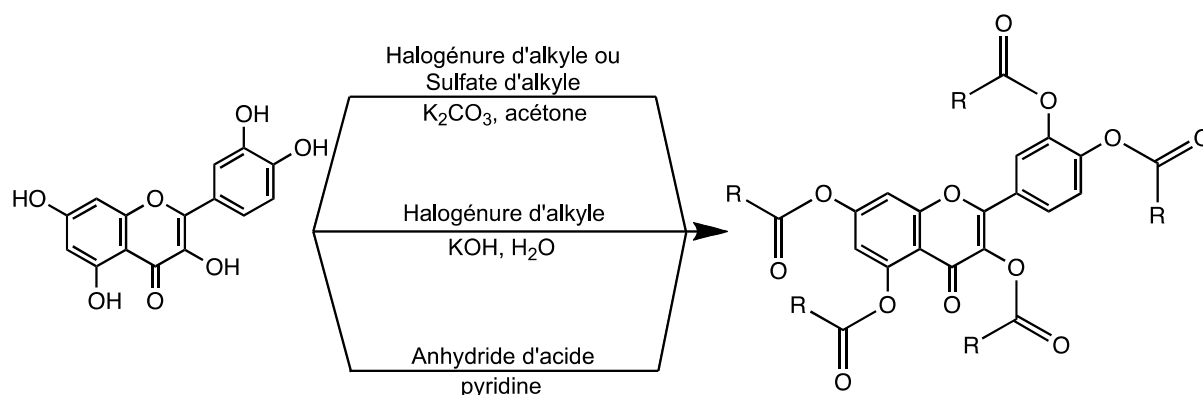


Figure I.7. Synthèses chimiques du 3,3',4',5,7-pentaacyl quercétine d'après [72].

Tableau I.6. Conditions expérimentales de la synthèse chimique des dérivés acylés de la quercétine.

Conditions expérimentales							Produits identifiés	Réf.
[Quercétine] (QCT, g.L ⁻¹)	Agent acylant (AA)	Ratio molaire [AA/QCT]	Catalyseur	Solvant	Température (°C)			
6,7	Sulfate de dialkyle	12	K_2CO_3	Acétone	T de reflux	3,5,7,3',4'-penta alkyle de QCT	[72]	
	Halogénure d'alkyle						[72]	
40	Iodure d'alkyle	-	-	Solution aqueuse de KOH (2 M)	80		[72]	
25	Chlorure d'acide laurique	10	K_2CO_3	Toluène	T de reflux	3,5,7,3',4'-penta laurate de QCT	[73]	
50			-	Pyridine	100		[73]	
25	Chlorure d'acide palmitique	10	K_2CO_3	Toluène	T de reflux	3,5,7,3',4'-penta palmitate de QCT	[73]	
50			-	Pyridine	100		[73]	
50	Anhydride d'acide butyrique	10	-	Pyridine	100	3,5,7,3',4'-penta butyrate de QCT	[73]	
1000	Anhydride acétique	30	-	Pyridine	T de reflux	3,5,7,3',4'-penta acétate de QCT	[72]	
1119		10,7					[74]	
67		20					[75]	
67		5	-	CH_2Cl_2 Pyridine	T ambiante	3,7,3',4'-tétra acétate de QCT	[75]	
338		1,2	Acétate de sodium	DMF	-10	3-mono acétate de QCT	[76]	
					0	3,3',4'-tri acétate de QCT	[77]	

Le nombre d'hydroxyles acylés dépend du solvant employé, de la température de la réaction et de la présence éventuelle d'un catalyseur. En effet, l'utilisation de toluène ou de pyridine, en présence ou non de catalyseur à différentes températures conduisent à des produits penta-estérifiés de la quercétine. En revanche, des solvants comme le DMF ou le dichlorométhane (CH_2Cl_2) permettent d'obtenir des mono-, tri- et tétra-ester de la quercétine. Et une modification de la température en présence de DMF mène à deux produits de degré d'estérification différent (mono- et tri-ester).

2.6.2. Acylation par voie enzymatique

Les travaux relatant la synthèse enzymatique de dérivés acylés de la quercétine sont limités malgré les nombreuses enzymes qui ont été testées dans le cadre de cette réaction. Néanmoins, la littérature mentionne trois enzymes capables d'accomplir l'acylation de ce flavonoïde aglycone. Le détail des conditions opératoires et les produits identifiés selon le biocatalyseur utilisé sont synthétisés dans le *Tableau I.7.*

Tableau I.7. Conditions opératoires et produits identifiés pour l'acylation enzymatique de la quercétine (QCT) avec les biocatalyseurs PCL, CALB et APE1547.

Conditions expérimentales							Produits identifiés	Réf.
Biocatalyseur (BC)	[BC] (g.L ⁻¹)	[QCT] (g.L ⁻¹)	Agent acylant (AA)	Ratio molaire [AA/QCT]	Solvant	Température (°C)		
PSL-C II Lipase de <i>Pseudomonas cepacia</i>	50	10	Acétate de vinyle	40	Acétone	50	4'-acétate QCT; 3',4'-acétate QCT; 7,3',4'-acétate QCT	[78]
Novozym 435 Lipase B de <i>Candida antarctica</i>	60	5		330	CH ₃ CN Acétone 2-méthyl-2-butanol	50 ou 60	3'-acétate QCT; 4'-acétate QCT	[79]
APE1547 Estérase de <i>Aeropyrum pernix</i> K1	10	3		25	CH ₃ CN	60	4'-acétate QCT 3',4'-acétate QCT 7,3',4'-acétate QCT	[80]

La difficulté de conduire cette réaction se traduit d'une part, par le nombre restreint d'enzymes ayant une spécificité envers la quercétine et d'autre part, par la nécessité d'utiliser des donneurs d'acyles activés de type ester vinylique. Par ailleurs, le degré de régiosélectivité est faible en présence des biocatalyseurs PCL et APE1547 (formation de mono-, di- et tri- ester).

2.6.3. Acylation par voie chimio-enzymatique

Les procédés d'acylation via des méthodes chimio-enzymatiques combinent séquentiellement (i) la synthèse chimique du 3,5,7,3',4'-pentaacyl quercétine issus du traitement de la quercétine avec de l'anhydride d'acide dans de la pyridine [81, 82] ; (ii) et l'alcoolyse du substrat penta-substitué est réalisée en utilisant la régiosélectivité des enzymes envers ce flavonoïde aglycone acylé (*Tableau I.8.*)

Tableau I.8. Alcoolyse de la quercétine pentaacylée par différents biocatalyseurs.

Conditions expérimentales						Produits identifiés	Réf.
Biocatalyseur (BC)	[BC] (g.L ⁻¹)	[3,5,7,3',4' - pentaacyl Q] (g.L ⁻¹)	Solvant	Température (°C)	Temps de réaction		
Amano PS Lipase de <i>Pseudomonas cepacia</i>	20	6	THF	42°C	24h	3,5,3',4'-tétra acétate QCT ; 3,5,3'- tri acétate QCT ; 3,3',4'-tri acétate QCT ; 3-mono acétate QCT	[81]

PSL Lipase de <i>Pseudomonas</i> <i>cepacia</i>	100	25	THF/1- butanol	45°C	36	3-monoacyl QCT	[83- 85]
Novozym 435 Lipase B de <i>Candida antarctica</i>	40	6	t-BME/n- butanol	45°C	24	3,5,7-tri acétate QCT	[86, 87]
Lipozyme® IM Lipase de <i>Mucor</i> <i>raii</i>	n.r.*	n.r.*	t-BME/n- butanol (10 equiv.)	45°C	12-96	3-monoacyl QCT	[88]
	40	20	t-BME/n- butanol		24	3,5,3',4'-tétra acétate QCT	[86, 87]
					48	3-mono acétate QCT	[87]
	10	12	t-BME		120	3-monoacyl QCT	[89]

*n.r. : non renseigné

La diversité de dérivés acylés de la quercétine obtenue lors de l'hydrolyse enzymatique dépend du biocatalyseur ainsi que du solvant et du temps de réaction. La spécificité des enzymes envers le dérivé penta-acétylé de la quercétine conduit à une régiosélectivité supérieure à celle obtenue pour les voies chimique et enzymatique.

Dans ce travail de thèse, nous nous sommes focalisés sur la réaction d'acylation du flavonoïde aglycone quercétine catalysée par des lipases pour différentes raisons. Premièrement, les lipases sont des biocatalyseurs adaptés pour les réactions d'estérification et de transestérification. Deuxièmement, des travaux antérieurs ont montré que les lipases présentent une forte régiosélectivité envers les formes glycosylées de la quercétine, l'isoquercitrine et la rutine. Troisièmement, le mécanisme catalytique de ces enzymes est bien connu. Finalement, les lipases ont de nombreux avantages qui sont détaillés dans la section suivante.

3. Les lipases

3.1. Introduction

Les lipases ou triacylglycéril acylhydrolases (EC 3.1.1.3) sont des enzymes largement exploitées en biotechnologie [90]. Leur rôle physiologique est de catalyser l'hydrolyse des triglycérides en diglycérides, monoglycérides, acides gras et glycérols [91]. Ces enzymes ont également la capacité de catalyser des réactions de synthèse telles que l'estérification, la transestérification, l'interestérification. La Figure 1.8 montre la liste des réactions catalysées par les lipases. Selon le solvant employé, la réaction d'hydrolyse ou de synthèse d'esters sera menée. Le vaste potentiel biotechnologique des lipases provient de leur gamme étendue de substrats acceptés, de leur spécificité de substrats, de leur stabilité dans les milieux aqueux et organiques et de leur pouvoir catalytique de biotransformations chémo-, régio- et énatio-sélectives [90, 92-94]. La polyvalence réactionnelle des lipases et la diversité de leurs propriétés catalytiques leur confèrent une place de choix pour des applications industrielles et académiques [95]. Le *Tableau 1.9* synthétise plusieurs applications spécifiques des lipases dans différents secteurs industriels. En outre, ces enzymes

présentent une régiospécificité, une spécificité de substrat et une stéréospécificité qui sont mises à profit dans de nombreuses applications.

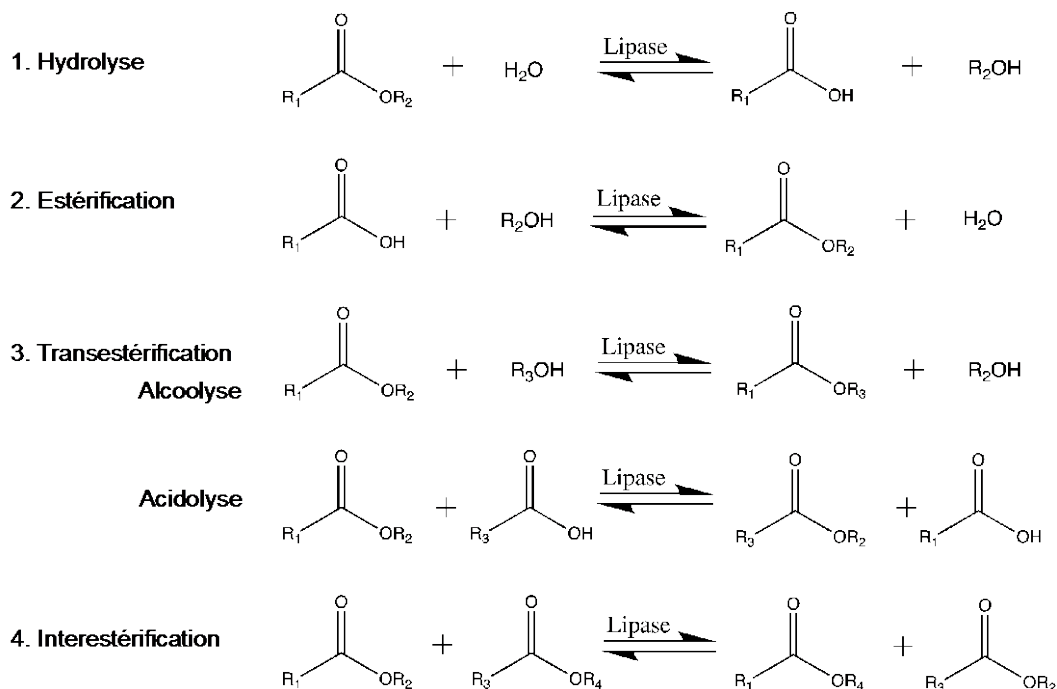


Figure I.8. Les différentes réactions catalysées par les lipases.

Tableau I.9. Diverses applications des lipases dans différents secteurs industriels.

Industrie	Produit/application	Réf.
Détergent	Élimine les taches de lipides	[96]
Laiterie	Développement des saveurs dans le lait, le fromage et le beurre	[97, 98]
Boulangerie	Amélioration et contrôle du brunissement non enzymatique, augmentation du volume du pain, amélioration de la structure de la mie	[92]
Brasserie et alimentation	Amélioration de la saveur et de la qualité dans des boissons, de la viande et des poissons	[92]
Cuir	Produits en cuir	[92]
Papier	Amélioration de la qualité du papier	
Huile naturelle	Beurre de cacao	[99]
Saveurs et fragrance	Synthèse d'esters aromatiques naturels	[100]
Tensioactif	Utilisation de monoglycérides comme tensioactifs	[90]
Pharmaceutique	Production de médicaments chiraux, résolution de mélange racémique	[93, 101]
Carburant	Biodiesel	[102]
Tensioactifs	Production d'esters mono-acyle de sucres comme tensioactif	[103]

Les lipases sont des enzymes ubiquitaires, retrouvées dans les organismes de la flore microbienne (comprenant les bactéries et les champignons) ainsi que dans les règnes végétal et animal. Elles sont préparées soit par extraction à partir d'animaux ou de tissus végétaux, ou par culture de micro-organismes. En général, les lipases disponibles dans le commerce sont produites

avec des rendements élevés par des organismes microbiens. Par exemple, la lipase B de *Candida antarctica* (CALB), la lipase de *Burkholderia cepacia* (anciennement *Pseudomonas cepacia*, PCL), la lipase de *Rhizomucor meihei* (RML) et la lipase de *Candida rugosa* (CRL) font partie des lipases les plus usitées dans les procédés de bioconversion. Par ailleurs, l'étape clé de la conception d'un procédé de bioconversion, catalysé par des lipases, est la sélection d'une enzyme possédant une activité catalytique élevée avec une sélectivité souhaitée. Depuis une décennie, l'amélioration de l'activité enzymatique et/ou de la sélectivité s'est développée en concevant de nouvelles lipases par mutation ou par évolution des enzymes existantes. Pour une sélection rationnelle ou d'ingénierie du biocatalyseur, une compréhension des bases moléculaires du mécanisme des lipases est nécessaire. En effet, la connaissance des résidus d'acides aminés de la lipase interagissant avec les groupements fonctionnels des substrats aide à déterminer l'activité de biotransformation et de la sélectivité.

Dans la littérature, un nombre croissant de publications a abordé le mécanisme de base de la catalyse par la lipase. Beaucoup d'entre elles impliquent des techniques de modélisation moléculaire pour étudier les processus responsables de la sélectivité des lipases qui peut être rationalisée en se basant sur les interactions des acides aminés des lipases avec les substrats de la réaction lors de la liaison du substrat dans la cavité catalytique, ou encore à une étape ultérieure de la transformation des substrats.

3.2. Structure des lipases et mécanisme catalytique

3.2.1. Structure générale des lipases

Les lipases ont une structure tertiaire qualifiée de repliement α/β -hydrolase [104]. Ce repliement consiste en un cœur de huit brins β (1 à 8) principalement parallèles qui est entouré de chaque côté par des hélices α (A à F). Le feuillet β central présente une torsion hélicoïdale de telle sorte que la surface du feuillet couvre environ la moitié d'un cylindre et les brins β ultrapériphériques sont à un angle de 90° l'un de l'autre. Les hélices localisées sur le côté convexe du feuillet β sont les hélices B, C, D et E tandis que les hélices A et F sont situées sur le côté concave. La connectivité des brins et des hélices des α/β -hydrolases est décrite dans la Figure 1.9.

Autour du feuillet central, plusieurs acides aminés sont responsables des propriétés catalytiques de l'enzyme ainsi que de la spécificité de substrat. L'ordre des acides aminés catalytiquement actifs dans la séquence primaire des lipases est le nucléophile (Ser), l'acide carboxylique (Glu, Asp), et l'histidine (His) qui constitue la triade catalytique. L'histidine est située sur une boucle entre le brin $\beta 8$ et l'hélice αF , l'acide est généralement localisé dans une boucle à l'extrémité du brin $\beta 7$ et, la sérine est incluse dans un pentapeptide hautement conservé (G-X-S-X-G). Cette dernière se trouve dans un virage étroit entre le brin $\beta 5$ et l'hélice αC qui correspond à un motif structural appelé coude nucléophile. Deux résidus complémentaires à proximité de la sérine catalytique participant à la stabilisation du substrat forment le trou oxyanionique. Le domaine de l'hélice αD forme le couvercle de

quelques lipases qui a un rôle dans la définition de la forme, de la régulation de l'accessibilité du substrat au site de fixation des enzymes.

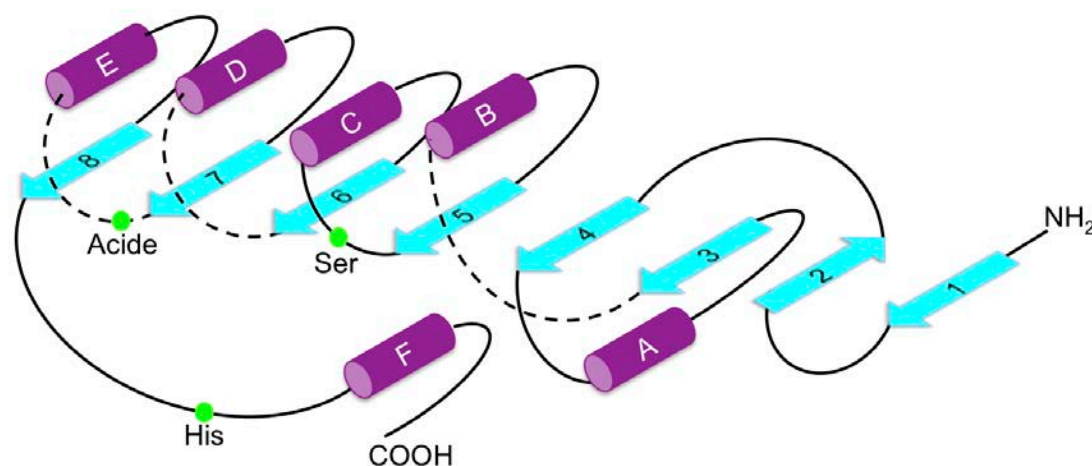


Figure I.9. Représentation schématique du repliement des α/β -hydrolases [91, 104]. Les hélices α sont représentées par des cylindres, les brins β par des flèches et les acides aminés catalytiques sont identifiés par des points verts (Ser, Sérine ; Acide ; His, Histidine).

3.2.2. Le mécanisme enzymatique des lipases

La réaction catalytique des lipases caractérisée par un mécanisme appelé bi-bi ping-pong est schématisée dans la Figure I.10. Il comprend la formation de deux intermédiaires tétraédriques et d'un acyl-enzyme.

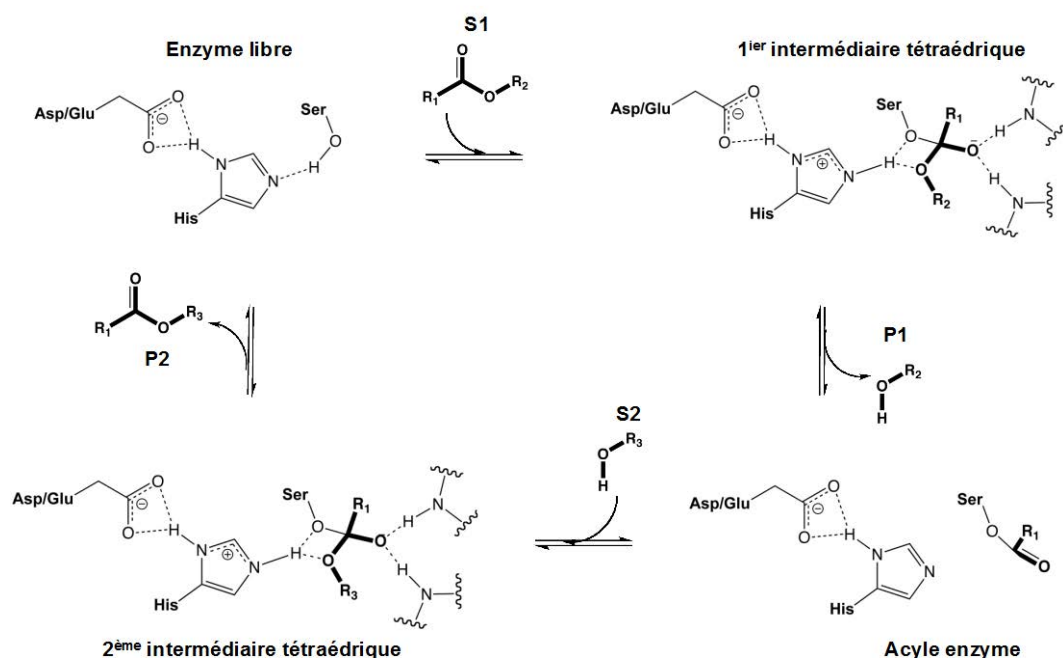


Figure I.10. Mécanisme catalytique des lipases. Si $R_2 = H$ et $R_3 =$ chaîne carbonée, la réaction est une estérification ; si $R_2 =$ chaîne carbonée et $R_3 = H$, la réaction est une hydrolyse, si $R_2 =$ chaîne carbonée et $R_3 =$ chaîne carbonée, la réaction est une transestérification.

3.3. La lipase B de *Candida antarctica*

3.3.1. Structure de la CALB

La lipase B de *Candida antarctica* (CALB) est une enzyme largement utilisée dans des applications industrielles [106]. Elle est notamment exploitée comme catalyseur pour l'estérification ou la transestérification d'acides gras dans la production de biodiesel [107] ou de glycérols monoacyl [108]. Elle est connue pour rester active dans des solvants organiques aussi bien polaires que non polaires ainsi que dans des milieux pauvres en eau.

La CALB est composée de 317 acides aminés et a un poids moléculaire de 33 kDa. Sa structure tridimensionnelle a été résolue par cristallographie et diffraction aux rayons X en 1994 par Uppenberg [109]. Sa structure centrale comprend sept brins β dont le premier brin est anti-parallèle aux six derniers. Ces brins β sont en alternance avec dix hélices α et quinze segments de boucle. Les dix résidus localisés à l'extrémité C-terminale forment deux brins de feuillet β ($\beta 8$, $\beta 9$) qui adoptent une conformation en épingle à cheveux (en anglais, *β -hairpin*). De plus, les six cystéines présentes dans la structure établissent les ponts disulfures suivants : Cys22-Cys64, Cys216-Cys258 et Cys293-Cys311. La Figure I.11 montre la structure tridimensionnelle de la CALB et la Figure I.12 présente la topologie de la structure secondaire.

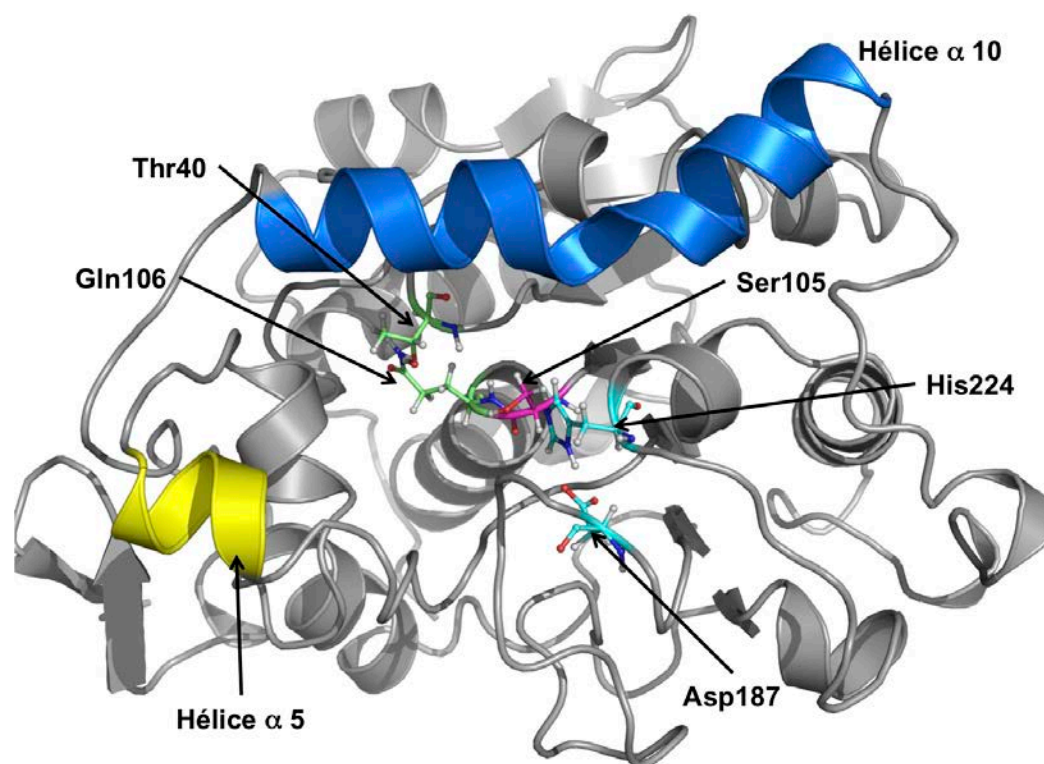


Figure I.11. Structure tridimensionnelle de la CALB (PDB ID : 1LBS). Les résidus du trou oxyanionique (Thr40 et Gln106) sont représentés en stick vert. Les résidus de la triade catalytique sont indiqués en stick cyan (His224 et Asp187) et en stick rose (Ser105). Les deux hélices α (hélice $\alpha 5$ et hélice $\alpha 10$) bordant l'entrée du site actif sont également mise en évidence.

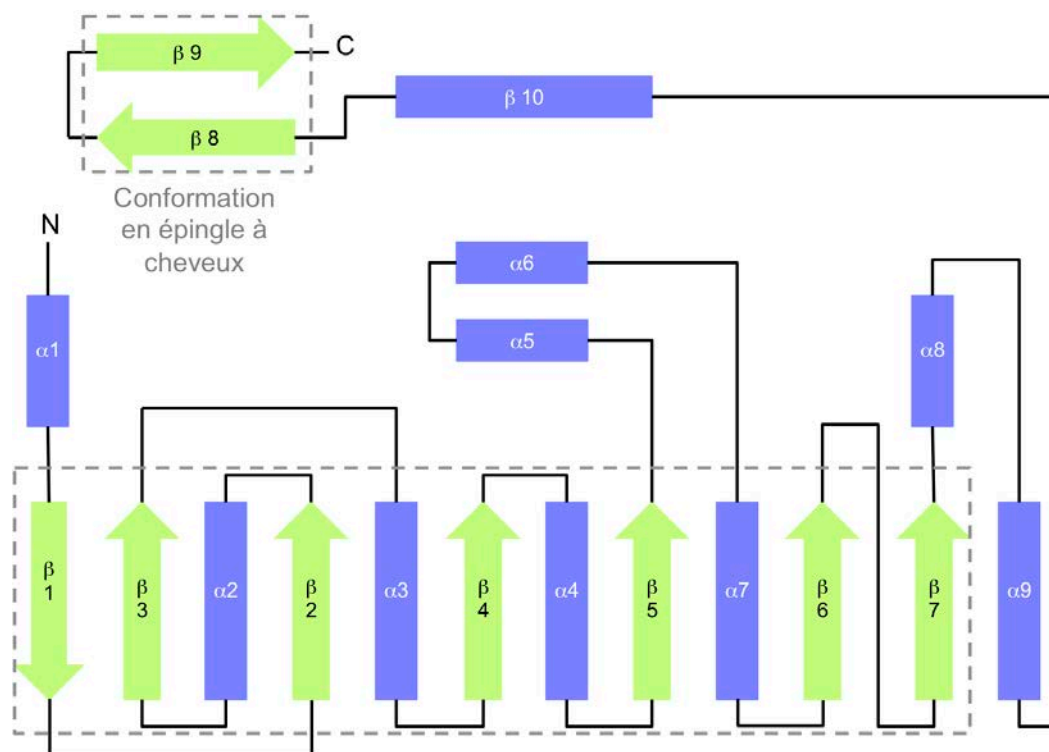


Figure I.12. Topologie de la structure secondaire de la CALB [109].

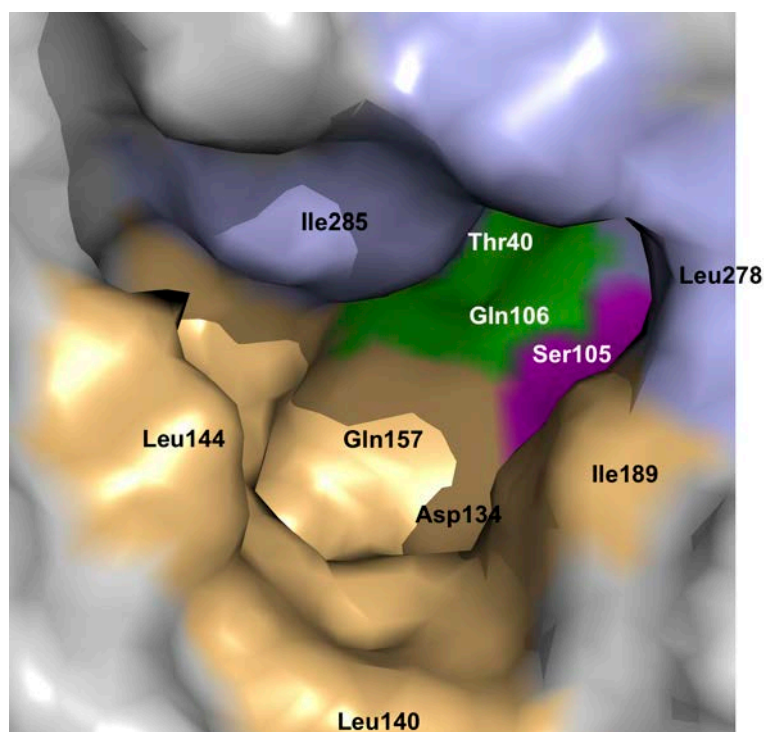


Figure I.13. Représentation en surface de la cavité de la lipase B de *Candida antarctica*. La poche du site actif peut être partitionnée en deux : le côté où se positionne la chaîne acyle (en orange) est bordé par les acides aminés Leu140, Leu144 et Ile189, et le côté de la chaîne alcool est délimité par les résidus Leu278 et Ile285 (en violet). La sérine catalytique localisée au fond de la poche est colorée en rose et l'histidine catalytique est située en dessous du résidu Ile189. Le trou oxyanionique incluant la Thr40 et le Gln106 est indiqué en vert. Les résidus hydrophiles (Gln157, Asp134 et Thr40) sont également localisés dans la cavité de la CALB.

La triade catalytique de cette lipase comprend les résidus Ser105-His224-Asp187 (Figure I.11). Le trou oxyanionique est un arrangement spatial de trois donneurs de liaisons hydrogène, l'une provient de la chaîne latérale de la Thr40 et les deux autres sont fournis par les amides du squelette peptidique des résidus Thr40 et Gln106. La sérine (Ser105) catalytiquement active est située au fond d'une poche étroite et profonde dont l'ouverture est bordée par une région hydrophobe [110]. L'intérieur de la poche est essentiellement recouvert d'acides aminés hydrophobes excepté une région au voisinage de la sérine catalytique qui comprend trois acides aminés hydrophiles. Les facteurs déterminant la sélectivité de substrat sont les restrictions physiques et la nature hydrophobe de la poche du site actif. En effet, la CALB a un espace limité comparé à d'autres lipases et par conséquent, elle présente un haut degré de sélectivité. En outre, la poche du site actif est composée de deux canaux, l'un hébergeant le fragment acyle du substrat et l'autre hébergeant la fraction alcool du substrat comme montré dans la Figure I.13.

3.3.2. Activation interfaciale

Plusieurs lipases possèdent un couvercle (en anglais, *lid*) qui est typiquement constitué d'une hélice α connectée au reste de la protéine par l'intermédiaire de deux segments de boucle. Cela permet au couvercle de s'ouvrir ou de se fermer par le biais d'un mouvement de type charnière. Le site actif est accessible dans la conformation ouverte mais bloqué par le couvercle dans la conformation fermée. Généralement, la lipase prend la conformation fermée dans les milieux aqueux. L'ouverture du couvercle se produit lorsque la protéine est située à l'interface entre une phase hydrophile et hydrophobe, ce phénomène est désigné comme l'activation interfaciale [111]. Pour CALB, aucune conformation fermée n'a été observée dans les études cristallographiques. Initialement, les hélices $\alpha 5$ et $\alpha 10$ ont été proposées comme candidates pour le rôle de couvercle car elles sont situées au bord de la poche du site actif (Figure I.11). Plus spécifiquement, l'hélice $\alpha 5$ qui dans deux structures cristallines (1TCC et 1LBS) est désordonnée c'est-à-dire que sa conformation n'est pas bien définie [109, 110]. Son rôle en tant que couvercle a toutefois été écarté car aucun des résultats expérimentaux ne semblaient soutenir cette hypothèse. Martinelle et al. (1995) [94] ont signalé que la CALB n'affiche pas d'activation interfaciale lors de l'hydrolyse du p-nitrophényl acétate en solution aqueuse. La vitesse de la réaction en fonction de la concentration en substrat suit la cinétique classique de Michaelis-Menten. Cette relation a également été observée à la concentration micellaire critique où les molécules de substrat sont trop peu nombreuses pour former des agrégats. Ceci a été comparé à la même réaction catalysée par la lipase de *Humigicola lanuginosa* qui possède un couvercle. Cette enzyme a été inactive en dessous de la concentration micellaire critique mais elle a été activée lorsque des concentrations plus importantes de substrat ont été atteintes. Récemment, Skjot et al. (2009) [112] ont réalisé une étude mutationnelle de la CALB expérimentale et par DM en milieux aqueux afin de déterminer le rôle de l'hélice $\alpha 5$. Dans l'étude de mutation expérimentale, la séquence de la région autour de l'hélice $\alpha 5$ a été échangée contre des séquences de protéines homologues. Les mutants avec des couvercles plus ou moins larges ont eu un impact significatif sur l'activité enzymatique et l'énantiosélectivité. Dans les simulations, il a été observé que cette hélice s'est déroulée et la boucle résultante s'est déplacée vers l'hélice $\alpha 10$, bloquant ainsi le

site actif. A la lumière de ces résultats, les auteurs ont suggéré que l'hélice $\alpha 5$ fonctionne comme un couvercle pour la CALB.

Trodler et Pleiss (2008) [113] ont effectué des simulations de DM de la CALB dans l'eau et dans cinq solvants organiques. De même, Branco et al. (2009) [114] ont réalisé des DM de la CALB dans un mélange gazeux d'eau et d'argon. Dans ces deux études, la région $\alpha 5$ a été signalé très flexible mais aucun déroulement de cette hélice n'a été rapporté. La flexibilité plutôt élevée de l'hélice $\alpha 10$ a également été mentionnée mais dans une moindre mesure que l'hélice $\alpha 5$. Etant donné que le support expérimental est dispersif, il est difficile d'établir si la CALB possède un couvercle. La dynamique des hélices $\alpha 5$ et $\alpha 10$ semble néanmoins d'une importance capitale pour les propriétés catalytiques.

3.4. La lipase de *Pseudomonas cepacia*

La lipase de *Pseudomonas cepacia* (PCL), renommée *Burkholderia cepacia*, est utilisée en synthèse organique et en hydrolyse pour son énantiosélectivité [90, 115]. En synthèse organique, PCL est fortement utilisée pour la résolution cinétique de mélange racémique d'alcool secondaire [116].

Sa chaîne polypeptidique est constituée de 320 acides aminés et elle a un poids moléculaire de 33 kDa. La résolution de la structure tridimensionnelle en conformation ouverte et en absence d'inhibiteur a été effectuée par cristallographie aux rayons X en 1997 par Kim [115]. Le feuillet β central est constitué de 6 brins parallèles qui sont flanqués de six hélices α . Les deux premiers brins β dans le repliement général des α/β -hydrolases sont absents dans cette lipase. De ce fait, le brin $\beta 1$ dans la topologie de la structure secondaire, représentée dans la Figure I.14, correspond au brin $\beta 3$ dans le repliement général des α/β -hydrolases. Par ailleurs, PCL possède un brin complémentaire ($\beta 6^*$) qui est aligné sur le sixième brin ($\beta 6$) du feuillet β central mais dans une direction opposée. Une autre particularité dans la structure de cette lipase est la petitesse de l'hélice $\alpha 10$. Un pont disulfure, entre Cys190 et Cys270, relie l'hélice $\alpha 10$ et la boucle du côté N-terminal du brin $\beta 6^*$.

La triade catalytique comporte les résidus Ser87, His286 et Asp264. Les acides aminés du trou oxyanionique, stabilisant l'état de transition via des liaisons hydrogène, sont formés par les amides du squelette peptidique de Leu17 et de Gln88. Contrairement à la CALB, la PCL a un couvercle défini par les hélices $\alpha 4$ et $\alpha 5$ (Figure I.15). Le déplacement de ces hélices vers l'hélice $\alpha 9$ conduit à la fermeture du site actif [117, 118] et donc, à l'inactivité de cette enzyme. Ce couvercle subit des réarrangements conformationnels qui permet à la lipase de commuter entre les états inactif (couvercle fermé) et actif (couvercle ouvert). La conformation fermée prédomine dans les milieux aqueux. Ceci est relié à la baisse de l'activité lipolytique de la plupart des lipases dans l'eau. La conformation ouverte prévaut dans les milieux organiques et aux interfaces eau/lipides dans lesquels les lipases sont très actives.

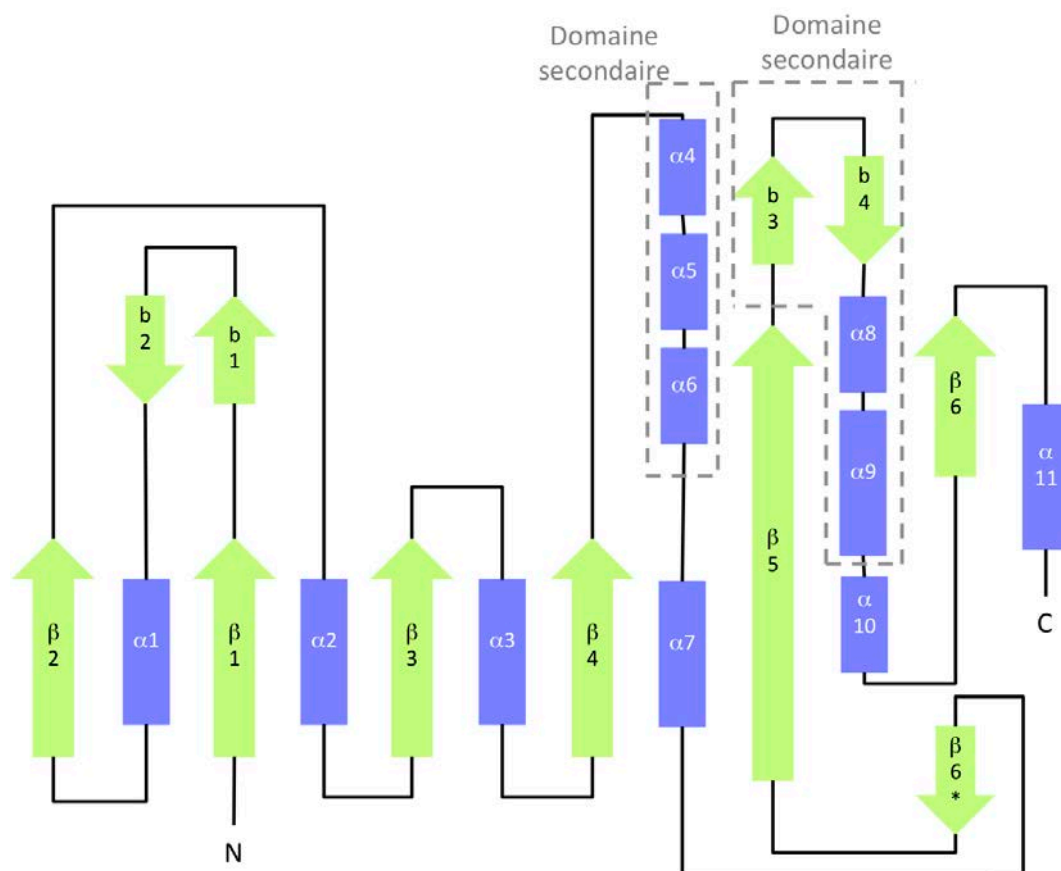


Figure I.14. Topologie de la structure secondaire de la lipase de *Pseudomonas cepacia*. Elle possède un domaine principal ($\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$, $\beta 5$, $\beta 6$, $\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 7$, $\alpha 10$ et $\alpha 11$) et deux domaines secondaires qui sont entourés en traits discontinus.

La cavité catalytique de la PCL a trois poches [119] comme montré à la Figure I.16.

- Une poche hydrophobique (HA) large et peu profonde qui est formée par les chaînes latérales des résidus Pro113, Phe119, Leu164, Leu167, Leu248, Val266, Val267 et Leu287 et le squelette peptidique de l'histidine catalytique (His286). En raison de sa taille et de son hydrophobicité, cette poche est appropriée pour accueillir de grands substituants hydrophobes.
- Une poche hydrophile/hydrophobe (HH) moyennement hydrophobe qui est composée des chaînes latérales des résidus Gly16, Thr18, Tyr29, His86, Leu287, Ile290, Gln292 et Leu293. Elle lie préférentiellement des chaînes alkyles de taille moyenne. Les chaînes latérales s'ajustent à la forme des substituants de taille moyenne.
- Une poche hydrophobe (HB) localisée le long de l'axe principal de la cavité de l'enzyme. Elle est constituée des chaînes latérales des résidus Ala247 et Thr251.

Le site actif est au fond de la poche de l'enzyme avec un chemin d'accès qui présente un étranglement marqué par les acides aminés Leu17 et Val266 (Figure I.16).

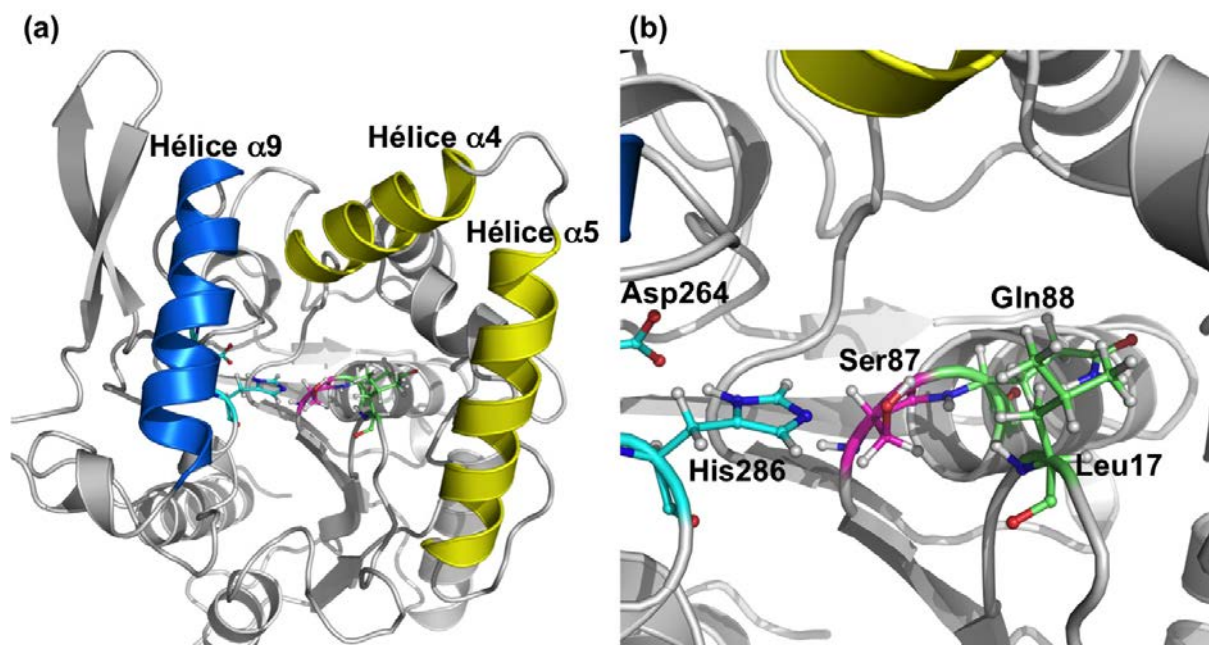


Figure I.15. (a) Structure tridimensionnelle de la PCL (PDB ID: 3LIP) avec les trois hélices participant au couvercle colorées en jaune et bleu. (b) La triade catalytique de la PCL est formée par les résidus Ser87 (rose), Asp264 (cyan), His286 (cyan) et, les résidus Leu17 (vert) et Gln88 (vert) constituent le trou oxyanionique de cette lipase.

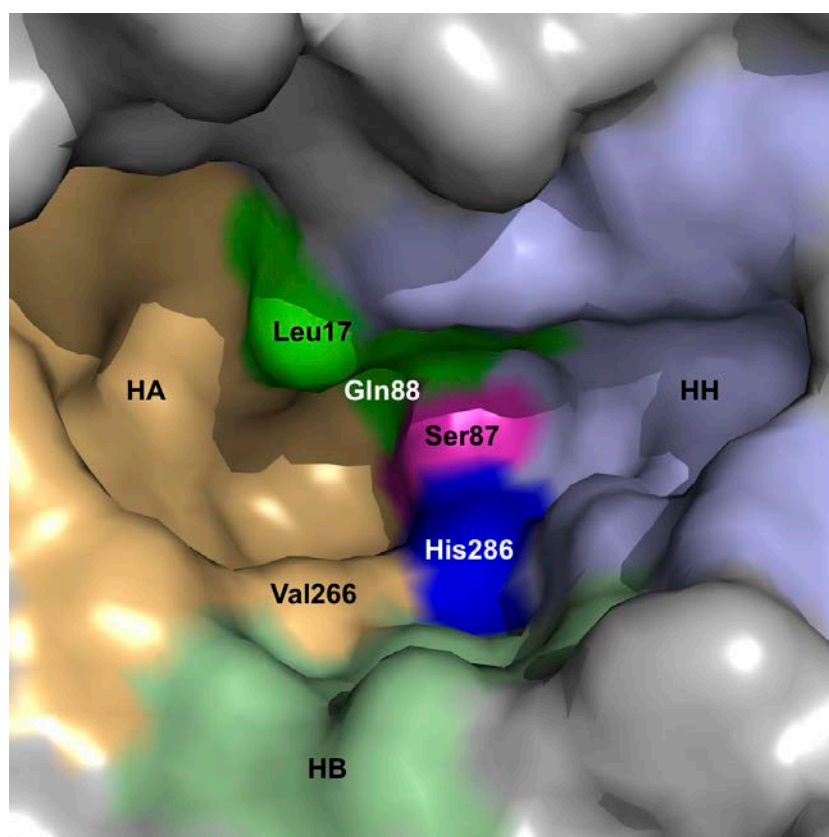


Figure I.16. Représentation en surface de la cavité catalytique de la lipase de *Pseudomonas cepacia*. Les trois poches, HA, HH et HB sont respectivement colorées en orange, violet et vert clair. Les résidus catalytiques, Ser87 (rose) et His286 (bleu) ainsi que les résidus du trou oxyanionique, Leu17 et Gln88 (vert), séparent les poches HA et HH.

3.5. La spécificité et la sélectivité de substrat des lipases

La spécificité et la sélectivité confèrent aux lipases leurs caractéristiques qui en font des outils intéressants en biocatalyse. La diversité de substrats acceptés par ces enzymes ainsi que le solvant, les liquides ioniques, la température, le pH, la concentration en sel, etc sont autant de paramètres qui influencent la spécificité et la sélectivité des lipases. Dans cette section, on s'est focalisé sur deux cas particulier : la spécificité des lipases CALB et PCL envers la quercétine; et la sélectivité de substrat de la CALB incluant la régiosélectivité de celle-ci envers les composés poly-hydroxylés.

3.5.1. La spécificité de substrat

La spécificité de substrat décrit l'aptitude d'une enzyme à catalyser la conversion d'un substrat. L'identité et la disposition des acides aminés dans le site actif définissent la spécificité envers les substrats. La complémentarité structurelle entre le substrat et le site actif, les interactions électrostatiques ainsi que le caractère hydrophobe du site actif sont des propriétés qui déterminent la spécificité. L'exemple de la spécificité des lipases CALB et PCL envers le flavonoïde aglycone quercétine est illustré dans la Figure I.17.

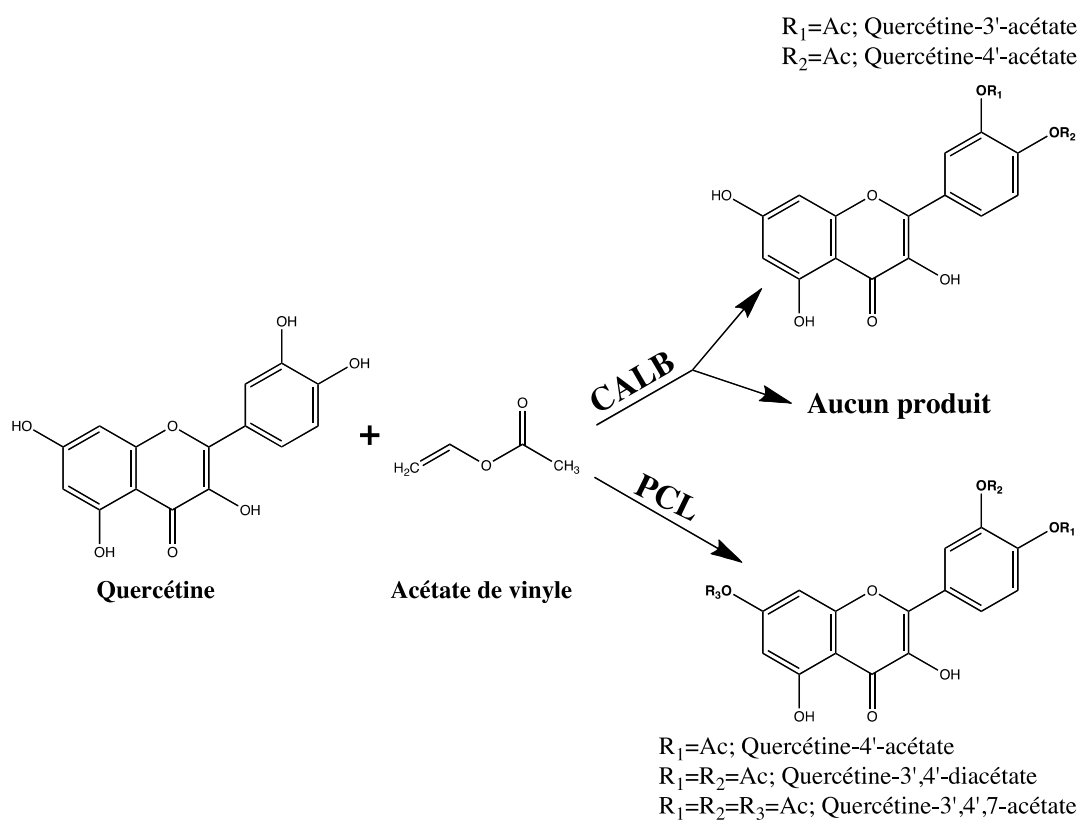


Figure I.17. Spécificité des lipases CALB et PCL envers le flavonoïde aglycone quercétine (d'après les réf. [78, 79, 120, 121]).

3.5.2. La sélectivité de substrat

Une caractéristique importante des enzymes est leur sélectivité, c'est-à-dire leur capacité à discriminer deux substrats différents. Il existe différents types de sélectivité de substrats tels que :

- **Chimio-sélectivité** : les enzymes peuvent discriminer des substrats ayant des groupements chimiques différents. Par exemple, la CALB montre une sélectivité 105 fois supérieure pour les alcools que pour les thiols dans les réactions de transacylation (Figure I.18) [122].

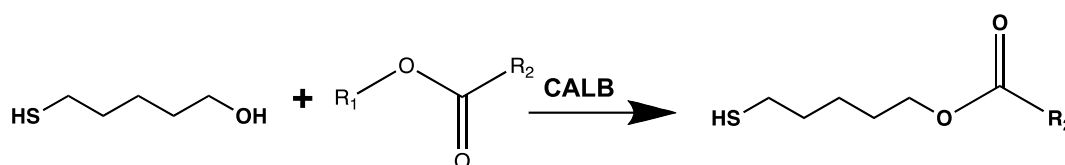


Figure I.18. Réaction d'acylation d'une fonction alcool et d'une fonction thiol par la lipase B de *Candida antarctica*.

- **Stéréosélectivité** : les enzymes présentent une sélectivité entre les stéréoisomères d'une molécule de substrat chiral. Cela signifie que l'un des isomères réagit plus rapidement que les autres. Quand les stéréoisomères sont des énantiomères, le phénomène est appelé énantiosélectivité. La résolution cinétique d'énantiomères d'alcools secondaires catalysée par la lipase B de *Candida antarctica* est une caractéristique de cette lipase largement étudiée expérimentalement et théoriquement [123-125]. L'intérêt de ces études est d'établir des modèles prédictifs pour la reconnaissance moléculaire des énantiomères qui dépend du substrat, de l'enzyme mais également du solvant, de la température, du donneur d'acyle et de la préparation de l'enzyme. La Figure I.19 donne un exemple de la résolution énantiomérique de la réaction d'acylation du 1-phényléthanol avec du butanoate de vinyle [123].

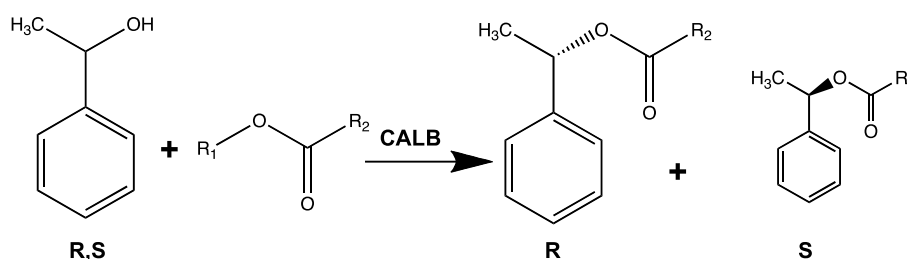


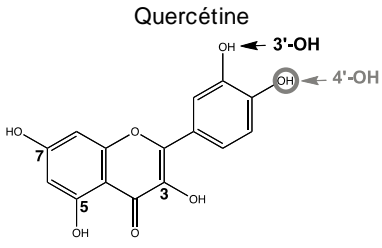
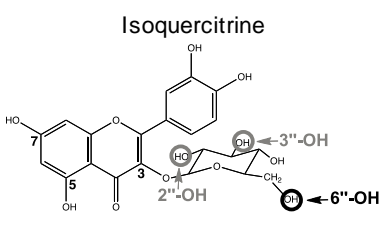
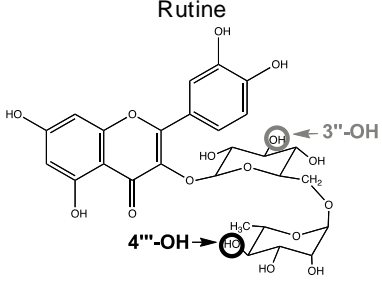
Figure I.19. Réaction de transacylation sur un mélange énantiomérique de 1-phényléthanol catalysée par CALB. L'acylation de l'énantiomère R est prédominante.

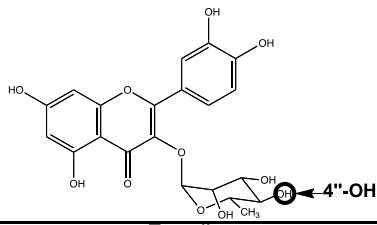
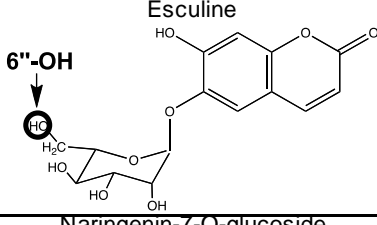
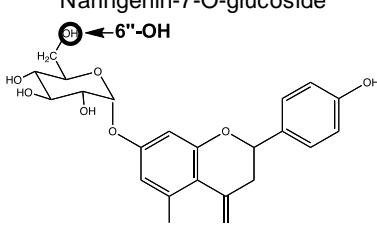
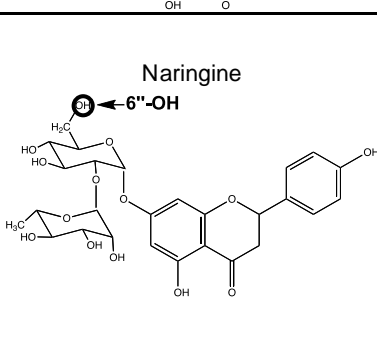
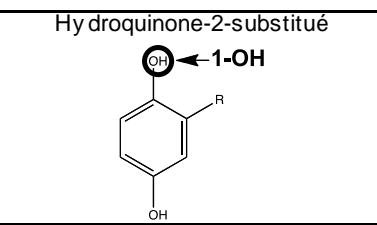
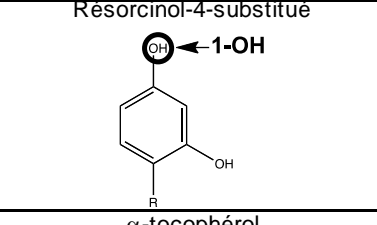
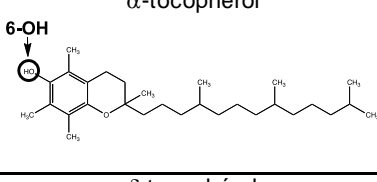
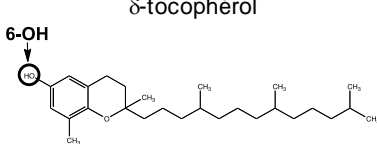
- **Régiosélectivité** : les enzymes peuvent être sélectives pour l'un des groupements similaires sur la molécule de substrat. Cette capacité de l'enzyme permet la protection et/ou la déprotection régiosélective de molécules. Notre intérêt s'est focalisé plus particulièrement sur la régiosélectivité de la lipase B de *Candida antarctica* envers des substrats poly-hydroxylés. Plusieurs études concernant l'acylation de composés poly-hydroxylés sont reprises dans le *Tableau I.10*.

L'analyse de ces données montre que la régiosélectivité de la CALB catalysant la réaction d'acylation de flavonoïdes est spécifique à la position C-6 (l'alcool primaire) de la fraction

glucose. En revanche, la sélectivité envers la rutine ou la quercitrine, qui n'a pas d'alcool primaire libre, est spécifique à la position C-4 du rhamnose. En présence d'un sucre dans la structure du composé, aucune acylation à l'égard des groupes hydroxyles phénoliques n'a été observée dans le squelette des flavonoïdes. Par ailleurs, une équipe de travail a constaté une acylation régiosélective de la quercétine par la CALB. Dans toutes les autres conditions testées, l'acylation de ce flavonoïde aglycone n'a pas été catalysée par la lipase B de *Candida antarctica*. D'autre part, un nombre réduit de travaux relate l'acylation directe d'hydroxyle phénolique catalysé par la CALB. Cela s'explique par la différence de nature des groupements hydroxyles des sucres et des groupements hydroxyles phénoliques. Ces derniers sont moins nucléophiles que les hydroxyles alcooliques.

Tableau I.10. Quelques travaux traitant de la régiosélectivité de la CALB dans l'acylation de composés poly-hydroxylés.

Substrat (S)	Conditions réactionnelles		Régiosélectivité	Réf.	
	Donneur d'acyle (DA)	Solvant, Ratio molaire DAS (R), température, temps de réaction			
Flavonoïdes	 <p>Quercétine</p>	Acide hexanoïque (C6 :0) Acide nanoïque (C9 :0) Acide laurique (C12 :0) Acide myristique (C14 :0) Acide palmitique (C16 :0) Acide stéarique (C18 :0) Acide oléique (C18 :1)	<i>tert</i> -amyl alcool ; R = 1 ; 60°C ; 150h	Aucun produit	[120]
		Acide palmitique (C16 :0) Acide cinnamique Acide phénylpropionique Acide 2-hydroxyphénylpropionique Acide 4-hydroxyphénylpropionique Acide 3,4-dihydroxyphénylpropionique	2-méthyl-2-propanol ; R = 2 ; 60°C ; 168h	Aucun produit	[121]
		Acétate de vinyle	<i>tert</i> -amyl alcool, acétonitrile ; R = 40 ; 50°C ; 96h	Aucun produit	[78]
		Acétate de vinyle	<i>tert</i> -amyl alcool, acétonitrile, acétone ; R = 330 ; 50 ou 60°C ; 72h	3'-O-acétate 4'-O-acétate	[79]
	 <p>Isoquercitrine</p>	Acétate de vinyle	Acétone ; R = 10 ; 45°C ; 60h	6''-O-acétate 3'',6''-O-diacétate	[126]
		Cinnamate de vinyle	Acétone ; R = 5 ; 37°C ; 168h	6''-O-cinnamate	[127]
		Acétate de vinyle	Acétone ; R = 40, 50°C, 96h	6''-O-acétate 3'',6''-O-diacétate 2'', 3'', 6''-O-triacétate	[78]
	 <p>Rutine</p>	Acétate de vinyle	Acétone ; R = 10 ; 45°C ; 45h	4'''-O-acétate 3'',4'''-O-diacétate	[126]
		Butyrate de vinyle	Liquides ioniques [bmim]BF ₄ et [bmim]PF ₆ ; R = 40 ; 60°C ; 96h	4'''-O-butyrate	[128]
		Acide oléique (C18:1) Acide linoléique (C18:2) Acide linoléique (C18:3) Linoate de vinyle	Acétone ; R = 5 ; 50°C ; 96h	4'''-O-oleate 4'''-O-linoléate 4'''-O-linolénate 4'''-O-linoate	[129]

	<p>Quercitrine</p>  <p>4''-OH</p>	Acétate de viny le	Acétone ; R = 1 ; 45°C ; 150h	4''-O-acétate	[126]
	<p>Esculine</p>  <p>6''-OH</p>	Buty rate de viny le	Liquides ioniques [bmim]BF ₄ , [bmim]PF ₆ , acétone ; R = 10 ; 50 ou 60°C ; 72h	6''-O-buty rate	[130]
	<p>Naringenin-7-O-glucoside</p>  <p>6''-OH</p>	Acétate de viny le Buty rate de viny le Laurate de viny le Décanoate de viny le Stéarate de viny le	Acétonitrile, Acétone, THF, <i>tert</i> - butanol ; R = 1 à 40 ; 50°C ; 6h	6''-O-laurate 6''-O-buty rate 6''-O-laurate 6''-O-décanoate 6''-O-stéarate	[131]
	<p>Naringine</p>  <p>6''-OH</p>	Laurate de viny le	Acétonitrile, Acétone, THF, <i>tert</i> - butanol ; R=10 ; 50°C ; 8h	6'''-O-laurate	[131]
		Buty rate de viny le	Liquides ioniques [bmim]BF ₄ , [bmim]PF ₆ , acétone ; R = 10 ; 50 ou 60°C ; 72h	6'''-O-buty rate	[130]
		Acide décanoïque	acétone, <i>tert</i> - butanol ; R = 6 ; 45°C ; 45h	6'''-O-decanoate	[132]
Composés phénoliques	<p>Hydroquinone-2-substitué</p>  <p>1-OH</p>	Propanoate de viny le	diisopropy le éther ; R= 3 ; 45°C ; 24h	1-O-propanoy l	[133]
	<p>Résorcinol-4-substitué</p>  <p>1-OH</p>	Propanoate de viny le	diisopropy le éther ; R= 3 ; 45°C, 72h	1-O-propanoy l	[133]
	<p>α-tocophérol</p>  <p>6-OH</p>	Acétate de viny le	<i>tert</i> -amy l alcool ; R = 4 ; 60°C ; 168h	Tocophéry l acétate	[134]
	<p>δ-tocophérol</p>  <p>6-OH</p>	Acétate de viny le	<i>tert</i> -amy l alcool ; R = 4 ; 60°C ; 168h	Tocophéry l acétate	[134]

*n.r. : non renseigné

4. Le design de protéines

La conception de protéines ou 'protein design' consiste à développer de nouvelles protéines dont les propriétés structurales et/ou fonctionnelles auraient été modifiées. Cette discipline récente peut être abordée par des approches expérimentales et/ou computationnelles. Deux stratégies prédominent dans ces approches. La première, dite rationnelle, repose sur la connaissance de la structure 3D de la protéine d'intérêt et consiste à muter des régions particulières de la protéine en vue d'obtenir les modifications désirées. Dans un premier temps, cette approche s'appuie sur des techniques computationnelles permettant la sélection d'acides aminés et éventuellement, la prédiction de l'effet des mutations rationalisées. Dans un second temps, la concrétisation des protéines mutées s'effectue via des techniques maîtrisées de mutagenèse dirigée qui présente l'avantage d'être relativement facile à mettre en place. La seconde méthode dite 'méthode combinatoire', consiste à muter aléatoirement la protéine cible et à mettre en place une procédure de sélection de façon à identifier les mutants présentant les propriétés souhaitées. Cette méthode mime les mécanismes de l'évolution naturelle et est connue sous le nom d'évolution dirigée. A l'inverse de la première approche, l'évolution dirigée ne requiert ni la connaissance de la structure 3D de la protéine ni de comprendre le mécanisme de la protéine. D'ailleurs, il arrive avec cette méthode d'obtenir la propriété recherchée en introduisant des mutations complètement inattendues. Cette technique est couramment utilisée depuis quelques années pour développer de nouveaux ligands et catalyseurs, et joue un rôle de plus en plus important dans les domaines pharmaceutique et biotechnologique. Néanmoins, cette approche s'avère coûteuse, longue à réaliser et ne permet pas une exploration complète des séquences disponibles pour la région à muter.

En revanche, l'approche rationnelle assistée par des outils de modélisation moléculaire, permet une exploration plus exhaustive et contrôlée des mutations possibles plutôt que de générer expérimentalement et aléatoirement des bibliothèques de mutants. Cette approche fut au départ essentiellement appliquée au renforcement de la stabilité des protéines existantes et au développement de protéines thermostables présentant un intérêt industriel indéniable [135, 136]. Par la suite ses applications furent étendues notamment au développement de protéines présentant de nouvelles fonctions ou de nouvelles propriétés catalytiques, à la modification de la sélectivité et de la spécificité des enzymes [137-139].

4.1. Design rationnel de la lipase B de *Candida antarctica*

Dans les procédés industriels, les lipases sont notamment exploitées pour des propriétés qui diffèrent de celles présentes dans leurs environnements physiologiques. Cela concerne principalement la spécificité de réaction, l'énantiosélectivité, la régiosélectivité, la thermostabilité, et la stabilité dans les milieux organiques. Dans de nombreux cas, la réactivité et/ou la sélectivité envers des substrats non naturels pour l'enzyme est relativement faible pour une réaction souhaitée. L'objectif du design rationnel est donc de développer des biocatalyseurs appropriés pour des

applications particulières. Les sections suivantes décrivent plusieurs travaux de mutation de la CALB menés ces dernières années.

4.1.1. Impact des mutations de la CALB sur sa spécificité de réaction

Plusieurs équipes ont entrepris des mutations sur la CALB afin d'altérer ou d'améliorer sa spécificité dans diverses réactions. Plus particulièrement, ces réactions incluent la formation de liaison carbone-carbone et des procédés oxydatifs. Le *Tableau I.11* reprend les mutations et les réactions ciblées par les différentes études.

La formation de liaison carbone-carbone a été étudiée par Branneby et al. [140] au cours d'une réaction d'addition aldol de l'hexanal et du propanal catalysée par la CALB. Ils ont démontré expérimentalement que le remplacement de la sérine par une alanine ou une glycine montre une activité spécifique quatre fois supérieure à la réaction catalysée par l'enzyme sauvage. Dans la réaction d'addition de Michael entre le composé 1,3-dicarbonyle et un aldéhyde ou une cétone α/β -insaturé, la réaction avec le variant Ser105Ala est 36 fois plus rapide que celle catalysée par l'enzyme sauvage [141]. Linder et al. [138] ont étudié la faisabilité de la réaction Diels-Alder catalysée par des mutants de la CALB *via* des calculs de dynamique moléculaire et de mécanique quantique. Il en conclut que la mutation Ser105Ala est nécessaire pour permettre un espace suffisant dans le voisinage du trou oxyanionique ; la mutation Ile189Ala permet d'obtenir un espace suffisant pour le substrat diénophile ; et le double variant Ser105Ala/Ile189Ala exécute la réaction Diels-Alder avec une performance catalytique améliorée par rapport à l'enzyme sauvage.

Les procédés oxydatifs impliquant des mutants de la CALB ont notamment été rapportés dans des réactions d'époxydation directes d'alcènes [142] et de la réaction d'oxydation de Baeyer-Villiger [137]. Dans ces deux types de réaction, la sérine a été substituée par une alanine, et une investigation théorique de la réaction a permis aux auteurs de proposer un mécanisme pour la réaction.

Tableau I.11. Mutants de la lipase B de *Candida antarctica* impliqués dans l'amélioration de diverses réactions. Les études comprenant une partie expérimentale et une partie modélisation moléculaire sont colorées en gris.

Variant de la CALB	Biotransformation	Réaction ciblée	Expérimental	Modélisation moléculaire	Réf.
S105A S105G	Formation de liaison carbone-carbone	Réaction aldol			[140]
S105A		Réaction d'addition de Michael			[141]
S105A I189A S105/I189A I285A S105/I285A		Réaction Diels-Alder			[138]
S105A	Procédé oxydatif	Réaction d'oxydation de Baeyer-Villiger			[137]
S105A		Epoxydation directe des alcènes			[142]

Dans les études rapportées ci-dessus, la modélisation moléculaire classique et quantique sont utilisés pour évaluer le positionnement des substrats dans la cavité du site actif, pour comprendre les mécanismes et pour déterminer la faisabilité de ces réactions spécifiques. Les calculs de mécanique

quantique et les simulations de DM ont contribué à déterminer que la mutation de la sérine nucléophile en un résidu alanine non polaire permet à l'histidine (His224) d'agir comme un catalyseur acide/base dans la formation de liaison carbone-carbone et les processus oxydatifs.

4.1.2. Impact des mutations de la CALB sur son énantiosélectivité

Un grand nombre de travaux concerne l'amélioration ou l'inversion de l'énantiosélectivité de la CALB par mutagenèse dirigée pour des substrats particuliers. L'énantiosélectivité des lipases suit la règle de Kazlauskas qui prédit une sélectivité R pour la plupart des substrats [143]. Chez la CALB, cette énantiosélectivité est expliquée en se basant sur les différents modes de liaison des énantiomères dans le site actif. L'énantiomère réagissant rapidement oriente son groupement de taille moyenne dans la poche stéréospécifique et son groupement large vers l'entrée du site actif. La taille de la poche stéréospécifique de la CALB est limitée par les chaînes latérales des résidus Thr42, Ser47 et Trp104. Le groupement alkyle le plus large qui peut s'insérer dans cette poche est un groupe éthyle. Cette limitation stérique rend la CALB hautement énantiosélective envers les alcools secondaires lorsque l'un des groupes est de taille inférieure ou équivalente à un groupement éthyle. Les différentes mutations entreprises pour modifier la synthèse énantiosélective et l'hydrolyse énantiosélective de la CALB sont reprises respectivement dans le *Tableau I.12* et le *Tableau I.13*.

Tableau I.12. Mutations de la CALB affectant la synthèse énantiosélective de différents substrats. L'augmentation ou la diminution d'énantiosélectivité est relatif au rapport énantiomérique lors de la synthèse énantiosélective en comparaison avec l'enzyme sauvage.

Variant de la CALB	Substrat ciblé	Donneur d'acyle	Enantio-sélectivité (ee)	Réf.
W104H	2-hexanol 1-phényl-éthanol 2-butanol 3-hexanol	2-propényl acétate 2-propényl acétate propanoate de vinyle 2-propényl acétate	↘ ↘ ↗ ↘	[144]
T42V S47A T42V/S47A W104H	1-bromo-2-octanol	butyrate de vinyle	= ↗↗ ↗ ↘	[145]
T42V S47A T42V/S47A W104H	1-chloro-2-octanol		= ↗↗ ↗ ↘	
T42V S47A T42V/S47A W104H	1-bromo-2-butanol		↘ ↘ ↘ ↘↘	
T42V S47A T42V/S47A W104H	2-butanol		= = ↘ ↘↘	
T42V S47A T42V/S47A W104H	Meso-2,3-butanediol		= = = ↘↘	
W104A	1-phényléthanol		Butanoate de vinyle	
T42V S47A T42V/S47A	Pentan-2-ol	Propanoate de méthyle	↗ ↗ =	[146]

La mutation du résidu Trp104 a permis dans plusieurs études d'inverser l'énantiosélectivité envers un substrat donné [123, 124, 147]. Par ailleurs, le remplacement des acides aminés Thr42 et Ser47 peuvent conduire à une augmentation ou une réduction de la synthèse énantiosélective. Cela dépend également de plusieurs autres facteurs tel que la température, le solvant et le substrat ciblé.

Tableau I.13. Impact de mutations de la CALB sur l'hydrolyse énantiosélective de différents substrats. L'augmentation ou la diminution d'énantiosélectivité est relatif au rapport énantiomérique lors de l'hydrolyse énantiosélective en comparaison avec l'enzyme sauvage.

Variant de la CALB	Substrat ciblé	Enantiosélectivité (ee)	Réf.
T40V T40A	2-hydroxy-propanoate	↗	[148]
T40V T40A	3-hydroxy-butanoate	↘	
CALB-N. crassa CALB- G. zeae	acide 2-méthylbutyrique	↗ ↘	[112]
L278A W104F S47A L278A/W104F L278A/S47A L278A/W104F/S47A	1-phényl éthylpropionate	↘↘ ↘↘ = ↘↘ ↘ ↘↘	[125]

4.1.3. Impact des mutations de la CALB sur son activité hydrolytique

Certaines études complètent la caractérisation des mutants en comparant l'activité hydrolytique de ce dernier avec celle de l'enzyme sauvage. Les résultats issus de ces travaux sont reportés dans le *Tableau I.14*.

Tableau I.14. Activité hydrolytique de mutants de la CALB.

Variant de la CALB	Résultats expérimentaux	Réf.
T103G S47A S47N S47H T42N T42A T42H T42V W104H T42V/S47A W104F	↘ l'activité spécifique envers la tributyrine	[149, 150] [145]
W104H M72L L278Y S47A	activité spécifique équivalente envers la tributyrine	[150] [125]
L278A L278V L278F L278A/W104F L278A/S47A L278A/W104F/S47A	↗ l'activité spécifique envers la tributyrine	[125]
Leu278P L278P/L219Q	↗ l'activité envers le p-nitrophényl palmitate	[151]
Couvercle substitué par celui d'homologues	↗ l'activité hydrolytique envers des esters de p-nitrophényl	[112]

Les modifications des résidus de la poche stéréospécifique (Thr42, Ser47 et Trp104) réduisent l'activité hydrolytique contrairement à la substitution du résidu Leu278 qui est situé sur l'hélice α 10 à l'entrée du site actif.

4.1.4. Impact des mutations de la CALB sur sa thermostabilité

En solvant organique, l'activité résiduelle de la CALB sauvage, selon la méthode employée, est inférieure ou au alentour de 50% pour des températures avoisinant 50°C [136, 149, 150, 152, 153]. La stabilité thermique des biocatalyseurs est importante dans de nombreux processus étant donné que les procédés industriels se produisent généralement à des températures élevées. Les avantages de ces conditions sont une vitesse de réaction plus rapide, une augmentation de la solubilité des substrats, une diminution de la viscosité et des risques de contamination.

Des mutations ponctuelles dans la CALB ont permis d'accroître ou de réduire la thermostabilité de cette enzyme (Tableau I.15). Les résidus mutés pour changer cette propriété ne sont généralement pas localisés dans le site actif de l'enzyme. Le repérage de ces points de mutations se fait selon les deux stratégies introduites au début de la section 4 : l'évolution dirigée et l'approche rationnelle. La stratégie d'évolution dirigée identifie le variant avec la propriété désirée [135, 152]. Néanmoins, cette approche aléatoire n'a qu'un succès marginal dans l'amélioration de la thermostabilité de la CALB. Quant à l'approche rationnelle, elle emploie des outils de modélisation moléculaire pour identifier des acides aminés intéressants qui sont par la suite modifiés par mutagenèse dirigée. Par exemple, Le et al. [153] ont utilisé plusieurs outils de calculs pour désigner des variants potentiels et pour réduire le nombre de mutation en analysant la flexibilité des acides aminés de ces variants. Kim et al. ont sélectionné leurs points de mutation en se basant sur la flexibilité des résidus de la structure cristallographique. Ils ont ensuite substitué ces sites avec des acides aminés plus rigides et optimisé ces structures. Les protocoles de MM se diversifient pour permettre de localiser des points de mutations intéressants dans une enzyme.

Tableau I.15. Mutations de la CALB qui agissent sur la thermostabilité de la lipase. Le substrat utilisé pour déterminer l'activité résiduelle, la température et l'activité résiduelle sont donnés dans le tableau. A titre indicatif, l'activité résiduelle de la CALB sauvage est indiquée entre parenthèse.

Variant de la CALB	Résultats expérimentaux				Réf.
	Thermostabilité	Substrat	Température (°C)	Activité résiduelle (%)	
T103G	↗	Tributy rine	60	71 (55)	[149]
N317Y		Esters C ₁₀ synthétique	90	30 (5)	[152]
A162C-K308C		para-nitrophényl	60	55 (46.5)	[153]
R249L		palmitate	52	80 (45)	[136]
A281E		Para-nitrophényl	70	50 (5)	[135]
V221D		butry rate	70	50 (5)	
W104H	↘	Tributy rine	60	12 (55)	[150]
M72L			60	15 (55)	
K13L			52	40 (45)	
K13L-R249L		para-nitrophényl	52	30 (45)	[136]
A281E		palmitate	52	10 (45)	

Certaines études ont exploré préalablement la structure de la CALB et les interactions avec le substrat à l'aide d'outil de modélisation moléculaire tels que le docking, la minimisation d'énergie et la dynamique moléculaire. En générale, l'aboutissement est la conception par mutagenèse dirigée des variants identifiés par MM. En revanche, d'autres études ont uniquement basé leur mutagenèse dirigée sur la connaissance de la structure tridimensionnelle du biocatalyseur. L'ingénierie enzymatique avec l'aide d'outils de modélisation moléculaire contribue à l'identification de résidus en se basant sur la connaissance des mécanismes catalytiques. Toutefois, la prédiction de l'activité enzymatique reste encore limitée dans le design rationnel. Des outils supplémentaires faisant appel à la mécanique quantique sont déployés afin de modéliser les différents états de transition au cours d'une réaction. Par ailleurs, cette approche exclut les résidus qui sont éloignés du site actif et qui peuvent influencer significativement les propriétés de l'enzyme [152] telle que la thermostabilité.

Le champ d'application de l'ingénierie de l'enzyme par design rationnel est d'éviter à terme d'utiliser des approches expérimentales qui prennent du temps, de l'argent et des ressources. Chaque année, l'amélioration des outils et des méthodologies de MM, et l'amélioration des connaissances sur les mécanismes catalytiques ainsi que l'accroissement des ressources informatiques contribuent à se rapprocher de cet objectif.

5. Références

1. Dinur, U. and A.T. Hagler, *New approaches to empirical force fields*, in *Reviews in Computational Chemistry*, K.B. Lipkowitz and D.B. Boyd, Editors. 1991, VCH: New York. p. 99-164.
2. Cornell, W.D., et al., *A second generation force field for the simulation of proteins, nucleic acids and organic molecules*. *Journal of the American Chemical Society*, 1995. **117**: p. 5179-5197.
3. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy minimization and dynamics calculations*. *Journal of Computational Chemistry*, 1983. **4**: p. 187-217.
4. Hermans, J., et al., *A Consistent Empirical Potential for Water-Protein Interactions*. *Biopolymers*, 1984. **23**: p. 1513-.
5. Jorgensen, W.L. and J. Tirado-Rives, *The OPLS Potential Functions for Proteins. Energy Minimization for Crystals of Cyclic Peptides and Crambin*. *Journal of American Chemical Society*, 1988. **110**: p. 1657-1666.
6. Brooks, B.R., et al., *CHARMM: The biomolecular simulation program* *Journal of Computational Chemistry*, 2009. **30**: p. 1545-1614.
7. Allen, M.P., *Introduction to molecular dynamics simulation*, in *Computational soft matter: from synthetic polymers to proteins*, K.B. Norbert Attig, Helmut Grubmüller, Kurt Kremer, Editor 2004, John von Neumann Institute for Computing: Jülich. p. 1-28.
8. Leach, A.R., *Molecular modelling: principles and applications*. 2nd edition ed2001, Harlow: Pearson Education.
9. Dias, R. and F. de Azevedo, *Molecular docking algorithms*. *Current Drug Targets*, 2008. **9**: p. 1040-1047.
10. Mohan, V., et al., *Docking : successes and challenges*. *Current Pharmaceutical Design*, 2005. **11**: p. 323-333.
11. Warren, G.L., et al., *A critical assessment of docking programs and scoring functions*. *Journal of Medicinal Chemistry*, 2006. **49**: p. 5912-5931.
12. Helperin, I., et al., *Principles of dockings : an overview of search algorithms and a guide to scoring functions*. *PROTEINS: Structure, Function and Bioinformatics*, 2002. **47**: p. 409-443.
13. Coupez, B. and R.A. Lewis, *Docking and Scoring - theoretically easy, practically impossible?* *Current Medicinal Chemistry*, 2006. **13**: p. 2995-3003.
14. Leach, A.R., B.K. Shoichet, and C.E. Peishoff, *Prediction of protein-ligand interactions. Docking and scoring : successes and gaps*. *Journal of Medicinal Chemistry*, 2006. **49**: p. 5851-5855.
15. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *Protein-ligand docking : current status and future challenges*. *PROTEINS: Structure, Function and Bioinformatics*, 2006. **65**: p. 15-26.
16. Teague, J.T., *Implications of protein flexibility for drug discovery*. *Nature Reviews Drug Discovery*, 2003. **2**: p. 527-541.
17. Meiler, J. and D. Baker, *ROSETTALIGAND: Protein-small molecule docking with full side-chain flexibility*. *PROTEINS: Structure, Function, and Bioinformatics*, 2006. **65**: p. 538-548.
18. Corbeil, C.R., P. Englebienne, and N. Moitessier, *Docking ligands into flexible and solvated macromolecules -1. Development and validation of FITTED 1.0*. *Journal of Chemical Informatic Modelling*, 2007. **47**: p. 435-449.
19. Huang, N., et al., *Physics-Based Scoring of Protein-Ligand Complexes: Enrichment of Known Inhibitors in Large-Scale Virtual Screening*. *Journal of Chemical Information and Modeling*, 2005. **46**(1): p. 243-253.
20. Bottegoni, G., *Protein-ligand docking*. *Frontiers in Bioscience*, 2011. **16**(1): p. 2289-2306.

21. Schulz-Gasch, T. and M. Stahl, *Scoring functions for protein-ligand interactions: a critical perspective*. Drug Discovery Today: Technologies, 2004. **1**(3): p. 231-239.
22. Huang, S.-Y., S.Z. Grinter, and X. Zou, *Scoring function and their evaluation methods for docking: recent advances and future directions*. Physical Chemistry Chemical Physics, 2010. **12**(40): p. 12899-12908.
23. Thomas, P.D. and K.A. Dill, *An iterative method for extracting energy-like quantities from protein structures*. Proceedings of the National Academy of Sciences, 1996. **93**: p. 11628-11633.
24. Tanaka, S. and H.A. Scheraga, *Medium- and Long-Range Interaction Parameters between Amino Acids for Predicting Three-Dimensional Structures of Proteins*. Macromolecules, 1976. **9**(6): p. 945-950.
25. Li, X. and J. Liang, *Knowledge-based energy functions for computational studies of proteins*, in *Computational Methods for Protein Structure Prediction and Modeling*, X.D. Xu Y., Liang J., Editor 2006, Springer: New York. p. 71-124.
26. Charifson, P.S., et al., *Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking Databases of Three-Dimensional Structures into Proteins*. Journal of Medicinal Chemistry, 1999. **42**(25): p. 5100-5109.
27. Clark, R.D., et al., *Consensus scoring for ligand/protein interactions*. Journal of Molecular Graphics and Modelling, 2002. **20**: p. 281-295.
28. Wang, R. and S. Wang, *How does consensus scoring work for virtual library screening? An idealized computer experiment*. Journal of Chemical Information and Computer Sciences, 2001. **41**(5): p. 1422-1426.
29. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models*. Accounts of Chemical Research, 2000. **33**(12): p. 889-897.
30. Rastelli, G., et al., *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA*. Journal of Computational Chemistry, 2010. **31**(4): p. 797-810.
31. Wang, J., et al., *Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA*. Journal of American Chemical Society, 2001. **123**: p. 5221-5230.
32. Muzzioli, E., A. Del Rio, and G. Rastelli, *Assessing Protein Kinase Selectivity with Molecular Dynamics and MM-PBSA Binding Free Energy Calculations*. Chemical Biology & Drug Design, 2011. **78**(2): p. 252-259.
33. Khuntawee, W., T. Rungrotmongkol, and S. Hannongbua, *Molecular Dynamic Behavior and Binding Affinity of Flavonoid Analogues to the Cyclin Dependent Kinase 6/cyclin D Complex*. Journal of Chemical Information and Modeling, 2011. **52**(1): p. 76-83.
34. Kuhn, B., et al., *Validation and Use of the MM-PBSA Approach for Drug Discovery*. Journal of Medicinal Chemistry, 2005. **48**(12): p. 4040-4048.
35. Hou, T., et al., *Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking*. Journal of Computational Chemistry, 2011. **32**(5): p. 866-877.
36. Thompson, D.C., C. Humblet, and D. Joseph-McCarthy, *Investigation of MM-PBSA Rescoring of Docking Poses*. Journal of Chemical Information and Modeling, 2008. **48**(5): p. 1081-1091.
37. Harwood, M., et al., *A critical review of the data related to the safety of quercetin and lack of evidence of in vivo toxicity, including lack of genotoxic/carcinogenic properties* Food and Chemical Toxicology, 2007. **45**: p. 2179-2205.
38. Aherne, S.A. and N.M. O'Brien, *Dietary flavonols: chemistry, food content, and metabolism* Nutrition, 2002. **18**(1): p. 75-81.
39. Halliwell, B., *Antioxidant characterization: Methodology and mechanism* Biochemical Pharmacology, 1995. **49**(10): p. 1341-1348.

40. Bischoff, S.C., *Quercetin: potentials in the prevention and therapy of disease*. Current Opinion in Clinical Nutrition and Metabolic Care, 2008. **11**: p. 733-740.
41. Boots, A.W., G.R.M.M. Haenen, and A. Bast, *Health effects of quercetin: from antioxidant to nutraceutical*. European Journal of Pharmacology, 2008. **585**: p. 325-337.
42. Heim, K.E., A.R. Tagliaferro, and D.J. Bobilya, *Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships*. The Journal of Nutritional Biochemistry, 2002. **13**(10): p. 572-584.
43. Leopoldini, M., N. Russo, and M. Toscano, *Gas and liquid phase acidity of natural antioxidants*. Journal of Agricultural and Food Chemistry, 2006. **54**: p. 3078-3085.
44. Mendoza, E.E. and R. Burd, *Quercetin as a systemic chemopreventative agent: structural and functional mechanisms*. Mini-Review in Medicinal Chemistry, 2011. **11**: p. 1216-1221.
45. Leopoldini, M., N. Russo, and M. Toscano, *The molecular basis of working mechanism of natural polyphenolic antioxidants*. Food Chemistry, 2011. **125**(2): p. 288-306.
46. Halliwell, B., *Free radicals and antioxidants - quo vadis?* Trends in Pharmacological Sciences, 2011. **32**(3): p. 125-130.
47. Chaudière, J. and R. Ferrari-Iliou, *Intracellular Antioxidants: from Chemical to Biochemical Mechanisms*. Food and Chemical Toxicology, 1999. **37**(9-10): p. 949-962.
48. Murakami, A., H. Ashida, and J. Terao, *Multitargeted cancer prevention by quercetin*. Cancer Letters, 2008. **269**: p. 315-325.
49. Chen, C., J. Zhou, and C. Ji, *Quercetin: a potential drug to reverse multidrug resistance*. Life Sciences, 2010. **87**(11-12): p. 333-338.
50. Musialik, M., et al., *Acidity of hydroxyl groups: an overlooked influence on antiradical properties of flavonoids*. Journal of Organic Chemistry, 2009. **74**: p. 2699-2709.
51. Cornard, J.P., L. Dangleterre, and C. Lapouge, *Computational and Spectroscopic Characterization of the Molecular and Electronic Structure of the Pb(II)-Quercetin Complex*. The Journal of Physical Chemistry A, 2005. **109**(44): p. 10044-10051.
52. Chen, W., et al., *Antioxidant property of quercetin-Cr(III) complex: The role of Cr(III) ion*. Journal of Molecular Structure, 2009. **918**(1-3): p. 194-197.
53. Martins, H.F.P., et al., *Towards the predictions of the activity of antioxidants: experimental and theoretical study of the gas-phase acidities of flavonoids*. Journal of the American Society for Mass Spectrometry, 2004. **15**(6): p. 848-861.
54. Escandar, G.M. and L.F. Sala, *Complexing behavior of rutin and quercetin*. Canadian Journal of Chemistry, 1991. **69**: p. 1994-2001.
55. Georgievskii, V.P., *Acidic properties of flavonoid compounds and the choice of solvent for performing potentiometric analysis*. Chemistry of Natural Compounds (Translation of Khimiya Prirodnikh Soedinenii), 1980. **16**(2): p. 136-141.
56. Milane, H.A., et al., *Isolation of quercetin's salts and studies of their physicochemical properties and antioxidant relationships*. Bioorganic & Medicinal Chemistry, 2004. **12**(13): p. 3627-3635.
57. Rothwell, J.A., A.J. Day, and M.R.A. Morgan, *Experimental determination of octanol-water partition coefficients of quercetin and related flavonoids*. Journal of Agricultural and Food Chemistry, 2005. **53**: p. 4355-4360.
58. Chebil, L., et al., *Solubilities inferred from the combination of experiment and simulation. Case study of quercetin in a variety of Solvents*. Journal of Physical Chemistry B, 2010. **114**: p. 12308-12313.
59. Chebil, L., et al., *Solubility of flavonoids in organic solvents*. Journal of Chemical and Engineering Data, 2007. **52**: p. 1552-1556.

60. Codorniu-Hernández, E., et al., *Essential amino acids interacting with flavonoids: A theoretical approach*. International Journal of Quantum Chemistry, 2005. **103**(1): p. 82-104.
61. Codorniu-Hernández, E., A. Rolo-Naranjo, and L.A. Montero-Cabrera, *Theoretical affinity order among flavonoids and amino acid residues: An approach to understand flavonoid-protein interactions*. Journal of Molecular Structure: THEOCHEM, 2007. **819**(1-3): p. 121-129.
62. Davis, A.M. and S.J. Teague, *Hydrogen bonding, hydrophobic interactions and the failure of the rigid receptor hypothesis*. Angewandte Chemie - International Edition, 1999. **38**: p. 736-749.
63. Panigrahi, S.K. and G.R. Desiraju, *Strong and weak hydrogen bonds in the protein-ligand interface*. Proteins: Structure, Function, and Bioinformatics, 2007. **67**(1): p. 128-141.
64. Meyer, E.A., R.K. Castellano, and F. Diederich, *Interactions with aromatic rings in chemical and biological recognition*. Angewandte Chemie - International Edition, 2003. **42**: p. 1210-1250.
65. Pandey, A.K., et al., *An in-silico strategy to explore neuroprotection by quercetin in cerebral ischemia: A novel hypothesis based on inhibition of matrix metalloproteinase (MMPs) and acid sensing ion channel 1a (ASIC1a)*. Medical Hypotheses, 2012. **79**(1): p. 76-81.
66. da Silva, E.R., C. do Carmo Maquiaveli, and P.P. Magalhaes, *The leishmanicidal flavonols quercetin and quercitrin target Leishmania (Leishmania) amazonensis arginase*. Experimental Parasitology, 2012. **130**(3): p. 183-188.
67. Shirfule, A.L., A.T. Sangamwar, and C.N. Khobragade, *Exploring glycolate oxidase (GOX) as an antiuro lithic drug target: Molecular modeling and in vitro inhibitor study*. International Journal of Biological Macromolecules, 2011. **49**(1): p. 62-70.
68. Cotrim, C.A., et al., *Quercetin as an inhibitor of snake venom secretory phospholipase A2*. Chemo-Biological Interactions, 2011. **189**(1-2): p. 9-16.
69. Saragusti, A.C., et al., *Inhibitory effect of quercetin on matrix metalloproteinase 9 activity Molecular mechanism and structure-activity relationship of the flavonoid-enzyme interaction*. European Journal of Pharmacology, 2010. **644**(1-3): p. 138-145.
70. Torres-Piedra, M., et al., *A comparative study of flavonoid analogues on streptozotocin-nicotinamide induced diabetic rats: Quercetin as a potential antidiabetic agent acting via 11Beta-Hydroxysteroid dehydrogenase type 1 inhibition*. European Journal of Medicinal Chemistry, 2010. **45**(6): p. 2606-2612.
71. Carlotti, M.E., et al., *On the complexation of quercetin with methyl-B-cyclodextrin: photostability and antioxidant studies*. Journal of Inclusion Phenomena and Macrocyclic Chemistry, 2011. **70**(1-2): p. 81-90.
72. Picq, M., et al., *Pentasubstituted quercetin analogs as selective inhibitors of particulate 3',5'-cyclic-AMP phosphodiesterase from rat brain*. Journal of Medicinal Chemistry, 1982. **25**(10): p. 1192-1198.
73. Bresson Rival, D., et al., *Cosmetic, dermatological, pharmaceutical, dietetic or food compositions, e.g. for improving skin condition, comprise flavonoid esters*, in *European Patent Office* 1999: France.
74. Biasutto, L., et al., *Ester-based precursors to increase the bioavailability of quercetin*. Journal of Medicinal Chemistry, 2007. **50**(2).
75. Mattarei, A., et al., *Regioselective O-derivatization of quercetin via ester intermediates. An improved synthesis of rhamnetin and development of a new mitochondriotropic derivative*. Molecules, 2010. **15**: p. 4722-4736.
76. Herowati, R., et al., *Anti-inflammatory activity of quercetin-3-monoacetate, selective acetylation product of quercetin*. Artocarpus, 2008. **8**(2): p. 60-67.
77. Gusdinar, T., et al., *Anti-inflammatory and antioxidant activity of quercetin-3, 3', 4'-triacetate*. Journal of Pharmacology and Toxicology, 2011. **6**(2): p. 182-188.
78. Chebil, L., et al., *Enzymatic acylation of flavonoids : effect of the nature of the substrate, origin of lipase and operating conditions on conversion yield and regioselectivity*. Journal of Agricultural and Food Chemistry, 2007. **53**(23): p. 9496-9502.

79. Kyriakou, E., et al., *Unexpected enzyme-catalyzed regioselective acylation of flavonoid aglycones and rapid product screening*. Organic & Biomolecular Chemistry, 2012. **10**(9): p. 1739-1742.
80. Xie, X.-N., et al., *Acylation of quercetin with a novel thermophilic esterase as biocatalyst*. Chemical Research in Chinese Universities, 2012. **28**(2): p. 225-229.
81. Natoli, M., G. Nicolosi, and M. Piattelli, *Regioselective alcoholysis of flavonoid acetates with lipase in an organic solvent*. The Journal of Organic Chemistry, 1992. **57**(21): p. 5776-5778.
82. Li, S. and M. Onda, *Heterocycles. XXV [1]. Sodium borohydride reduction of flavanols*. Journal of Heterocyclic Chemistry, 1990. **27**(7): p. 2029-2035.
83. Montenegro, L., et al., *In vitro evaluation of quercetin-3-O-acyl esters as topical prodrugs*. International Journal of Pharmaceutics, 2007. **336**: p. 257-262.
84. Saija, A., et al., *'In vitro' antioxidant and photoprotective properties and interaction with model membranes of three new quercetin esters*. European Journal of Pharmaceutics and Biopharmaceutics 2003. **56**(2): p. 167-174.
85. Lambusta, D., et al., *Enzyme-mediated regioprotection-deprotection of hydroxyl groups in (+)-catechin*. SYNTHESIS, 1993(11): p. 1155-1158.
86. D'Antona, N., et al., *Preparation of regioprotected morins by lipase-catalyzed transesterification*. Journal of Molecular Catalysis B: Enzymatic, 2008. **52-53**: p. 78-81.
87. Lambusta, D., et al., *Application of lipase catalysis in organic solvent for selective protection-deprotection of bioactive compounds*. Journal of Molecular Catalysis B: Enzymatic, 2003. **22**: p. 271-277.
88. Gatto, M.T., et al., *Antimicrobial and anti-lipase activity of quercetin and its C2-C16 3-O-acyl-esters*. Bioorganic and Medicinal Chemistry, 2002. **10**(2): p. 269-272.
89. Nicolosi, G., et al., *Biocatalytic process for the preparation of 3-O-acyl-flavonoids*, 1999.
90. Schmid, R.D. and R. Verger, *Lipases: Interfacial Enzymes with Attractive Applications*. Angewandte Chemie International Edition, 1998. **37**(12): p. 1608-1633.
91. Holmquist, M., *Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms*. Current Protein and Peptide Science, 2000. **1**: p. 209-235.
92. Hasan, F., A.A. Shah, and A. Hameed, *Industrial applications of microbial lipases*. Enzyme and Microbial Technology, 2006. **39**: p. 235-251.
93. Ghanem, A. and H.Y. Aboul-Enein, *Application of lipases in kinetic resolution of racemates*. Chirality, 2005. **17**(1): p. 1-15.
94. Martinelle, M. and K. Hult, *Kinetics of acyl transfer reactions in organic media catalysed by Candida antarctica lipase B*. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1995. **1251**(2): p. 191-197.
95. Villeneuve, P., *Lipases in lipophilization reactions*. Biotechnology Advances, 2007. **25**: p. 515-536.
96. Hasan, F., et al., *Enzymes used in detergents: lipases*. African Journal of Biotechnology, 2010. **9**(93): p. 4836-4844.
97. Neelakantan, S., A.K. Mohanty, and J.K. Kaushik, *Production and use of microbial enzymes for dairy processing*. Current Science, 1999. **77**(1): p. 143-148.
98. Arnold, R.G., K.M. Shahani, and B.K. Dwivedi, *Application of lipolytic enzymes to flavor development in dairy products*. Journal of Dairy Science, 1975. **58**(8): p. 1127-1143.
99. Bloomer, S., P. Adlercreutz, and B. Mattiasson, *Triglyceride interesterification by lipases. 1. Cocoa butter equivalents from a fraction of palm oil*. Journal of the American Oil Chemists' Society, 1990. **67**(8): p. 519-524.

100. Woolley, P. and S.B. Petersen, *Lipases - their structure, biochemistry and application* 1994, Great Britain: Cambridge University Press. 363.
101. Gotor-Fernandez, V., R. Brieva, and V. Gotor, *Lipases : useful biocatalysts for the preparation of pharmaceuticals*. Journal of Molecular Catalysis B: Enzymatic, 2006. **40**: p. 111-120.
102. Gog, A., et al., *Biodiesel production using enzymatic transesterification - Current state and perspectives*. Renewable Energy, 2012. **39**(1): p. 10-16.
103. Tarahomjoo, S. and I. Alemzadeh, *Surfactant production by an enzymatic method*. Enzyme and Microbial Technology, 2003. **33**(1): p. 33-37.
104. Ollis, D.L., et al., *The alpha/beta hydrolase fold*. Protein Engineering, 1992. **5**(3): p. 197-211.
105. Jaeger, K.E., et al., *Bacterial lipases*. FEMS Microbiology Reviews, 1994. **15**(1): p. 29-63.
106. Anderson, E.M., K.M. Larsson, and O. Kirk, *One biocatalyst - Many applications: the use of Candida antarctica B-lipase in organic synthesis*. Biocatalysis and Biotransformation, 1998. **16**: p. 181-204.
107. Zhang, B., et al., *Enzyme immobilization for biodiesel production*. Applied Microbiology and Biotechnology, 2012. **93**(1): p. 61-70.
108. Damstrup, M.L., et al., *Solvent optimization for efficient enzymatic monoacylglycerol production based on a glycerolysis reaction*. Journal of the American Oil Chemists' Society, 2005. **82**(8): p. 559-564.
109. Uppenberg, J., et al., *The sequence, crystal structure determination and refinement of two forms of lipase B from Candida antarctica*. Structure, 1994. **2**: p. 293-308.
110. Uppenberg, J., et al., *Cristallographic and molecular-modeling studies od lipase B from Candida antarctica reveal a stereospecificity pocket for secondary alcohols*. Biochemistry, 1995. **34**: p. 16838-16851.
111. Derewenda, Z.S., *Structure and function of lipases*. Advances in Protein Chemistry, 1994. **45**: p. 1-52.
112. Skjøt, M., et al., *Understanding the plasticity of the α/β -hydrolase fold: lid swapping on the Candida antarctica lipase B results in chimeras with interesting biocatalytic properties*. ChemBioChem, 2009. **10**: p. 520-527.
113. Trodler, P. and J. Pleiss, *Modeling structure and flexibility of Candida antarctica lipase B in organic solvents*. BMC Structural Biology, 2008. **8**: p. art. 9.
114. Branco, R.J.F., et al., *Molecular mechanism of the hydration of Candida antarctica lipase B in the gas phase: water adsorption isotherms and molecular dynamics simulations*. ChemBioChem, 2009. **10**: p. 2913-2919.
115. Kim, K.K., et al., *The crystal structure of a triacyl glycerol lipase from Pseudomonas cepacia reveals a highly open conformation in the absence of a bound inhibitor*. Structure, 1997. **5**: p. 173-185.
116. Schultz, T., J. Pleiss, and R.D. Schmid, *Stereoselectivity of Pseudomonas cepacia lipase towards secondary alcohols : a quantitative model*. Protein Science, 2000. **9**: p. 1053-1062.
117. Barbe, S., et al., *Insights into lid movements of Burkholderia cepacia lipase inferred from molecular dynamics simulations*. Proteins: Structure, Function, and Genetics, 2009. **77**: p. 509-523.
118. Trodler, P., R.D. Schmid, and J. Pleiss, *Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase*. BMC Structural Biology, 2009. **9**: p. art. n° 38.
119. Lang, D.A., et al., *Structural basis of the chiral selectivity of Pseudomonas cepacia lipase*. European Journal of Biochemistry, 1998. **254**(2): p. 333-340.
120. Ardhaoui, M., et al., *Acylation of natural flavonoids using lipase of Candida antarctica as biocatalyst*. Journal of Molecular Catalysis B: Enzymatic, 2004. **29**: p. 63-67.
121. Stevenson, D.E., et al., *Direct acylation of flavonoid glycosides with phenolic acids catalyzed by Candida antarctica lipase B (Novozyme 435)*. Enzyme and Microbial Technology, 2006. **39**: p. 1236-1241.

122. Hedfors, C., K. Hult, and M. Martinelle, *Lipase chemoselectivity towards alcohol and thiol acyl acceptors in a transacylation reaction*. Journal of Molecular Catalysis B: Enzymatic, 2010. **66**(1-2): p. 120-123.
123. Magnusson, A.O., et al., *An S-selective lipase was created by rational redesign and the enantioselectivity increased with temperature*. Angewandte Chemie - International Edition, 2005. **44**: p. 4582-4585.
124. Vallin, M., P.-O. Syren, and K. Hult, *Mutant lipase-catalyzed kinetic resolution of bulky phenyl alkyl secondary alcohols: a thermodynamic analysis of enantioselectivity*. ChemBioChem, 2010. **11**(3): p. 411-416.
125. Liu, D., et al., *Rational Design of Pseudozyma antarctica lipase B yielding a general esterification catalyst*. ChemBioChem, 2010. **11**: p. 789-795.
126. Danieli, B., et al., *Regioselective acylation of polyhydroxylated natural compounds catalyzed by Candida antarctica lipase B (Novozyme 435) in organic solvents*. Journal of Molecular Catalysis B: Enzymatic, 1997. **3**: p. 193-201.
127. Nakajima, N., et al., *Lipase-catalyzed direct and regioselective acylation of flavonoid glucoside for mechanistic investigation of stable plant pigments*. Journal of Bioscience and Bioengineering, 1999. **87**: p. 105-107.
128. Katsoura, M.H., et al., *Use of ionic liquids as media for the biocatalytic preparation of flavonoid derivatives with antioxidant potency*. Journal of Biotechnology, 2006. **123**: p. 491-503.
129. Mellou, F., et al., *Enzymatic esterification of flavonoids with unsaturated fatty acids : effect of the novel esters on vascular endothelial growth factor release from K562 cells*. Process Biochemistry, 2006. **41**: p. 2029-2034.
130. Katsoura, M.H., et al., *Effect of different reaction parameters on the lipase-catalyzed selective acylation of polyhydroxylated natural compounds in ionic liquids*. Process Biochemistry, 2007. **42**(9): p. 1326-1334.
131. Céliz, G. and M. Daz, *Biocatalytic preparation of alkyl esters of citrus flavanone glucoside prunin in organic media*. Process Biochemistry, 2011. **46**(1): p. 94-100.
132. Kontogianni, A., et al., *Lipase-catalyzed esterification of rutin and naringin with fatty acid of medium carbon chain*. Journal of Molecular Catalysis B: Enzymatic, 2003. **21**: p. 59-62.
133. Miyazawa, T., et al., *Highly regioselective propanoylation of dihydroxybenzenes mediated by Candida antarctica lipase B in organic solvents*. Tetrahedron Letters, 2008. **49**(1): p. 175-178.
134. Torres, P., et al., *Acetylation of vitamin E by Candida antarctica lipase B immobilized on different carriers*. Process Biochemistry, 2008. **43**: p. 145-153.
135. Zhang, N., et al., *Improving tolerance of Candida antarctica lipase B towards irreversible thermal inactivation through directed evolution*. Protein Engineering Design and Selection, 2003. **16**(8): p. 599-605.
136. Kim, H.S., Q.A.T. Le, and Y.H. Kim, *Development of thermostable lipase B from Candida antarctica (CaLB) through in silico design employing B-factor and RosettaDesign*. Enzyme and Microbial Technology, 2010. **47**(1-2): p. 1-5.
137. Carlqvist, P., et al., *Rational design of a lipase to accommodate catalysis of Baeyer-Villiger oxidation with hydrogen peroxide*. Journal of Molecular Modeling, 2003. **9**(3): p. 164-171.
138. Linder, M., et al., *Computational design of a lipase for catalysis of the Diels-Alder reaction*. Journal of Molecular Modeling, 2011. **17**(4): p. 833-849.
139. Syren, P.O., et al., *Increased activity of enzymatic transacylation of acrylates through rational design of lipases*. Journal of Molecular Catalysis B: Enzymatic, 2010. **65**(1-4): p. 3-10.
140. Branneby, C., et al., *Carbon-carbon bonds by hydrolytic enzymes*. Journal of American Chemical Society, 2003. **125**: p. 874-875.
141. Svedendahl, M., K. Hult, and P. Berglund, *Fast Carbon-Carbon Bond Formation by a Promiscuous Lipase*. Journal of the American Chemical Society, 2005. **127**(51): p. 17988-17989.

142. Svedendahl, M., et al., *Direct epoxidation in Candida antarctica lipase B studied by experiment and theory*. ChemBioChem, 2008. **9**: p. 2443-2451.
143. Kazlauskas, R.J., et al., *A rule to predict which enantiomer of a secondary alcohol reacts faster in reactions catalyzed by cholesterol esterase, lipase from Pseudomonas cepacia, and lipase from Candida rugosa*. Journal of Organic Chemistry, 1991. **56**: p. 2656-2665.
144. Patkar, S., et al., *Effect of mutations in Candida antarctica B lipase*. Chemistry and Physics of Lipids, 1998. **93**(1-2): p. 95-101.
145. Rotticci, D., et al., *Improved Enantioselectivity of a Lipase by Rational Protein Engineering*. ChemBioChem, 2001. **2**(10): p. 766-770.
146. Marton, Z., et al., *Mutations in the stereospecificity pocket and at the entrance of the active site of Candida antarctica lipase B enhancing enzyme enantioselectivity*. Journal of Molecular Catalysis B: Enzymatic, 2010. **65**(1-4): p. 11-17.
147. Engstrom, K., et al., *Mutated variant of Candida antarctica lipase B in (S)-selective dynamic kinetic resolution of secondary alcohols*. Organic & Biomolecular Chemistry, 2011. **9**(1): p. 81-82.
148. Magnusson, A., K. Hult, and M. Holmquist, *Creation of an Enantioselective Hydrolase by Engineered Substrate-Assisted Catalysis*. Journal of the American Chemical Society, 2001. **123**(18): p. 4354-4355.
149. Patkar, S.A., et al., *Effect of mutation in non-consensus sequence Thr-X-Ser-X-Gly of Candida antarctica lipase B on lipase specificity, specific activity and thermostability*. Journal of Molecular Catalysis B : Enzymatic, 1997. **3**: p. 51-54.
150. Patkar, S., et al., *Effect of mutations in Candida antarctica B lipase*. Chemistry and Physics of Lipids, 1998. **93**(1,2): p. 95-101.
151. Kim, S.-Y., et al., *In vitro evolution of lipase B from Candida antarctica using surface display in hansenula polymorpha*. Journal of Microbiology and Biotechnology, 2007. **17**(8): p. 1308-1315.
152. Chodorge, M., et al., *Rational strategies for directed evolution of biocatalysts - application to Candida antarctica lipase B (CALB)*. Advanced Synthesis and Catalysis, 2005. **347**: p. 1022-1026.
153. Le, Q.A.T., et al., *Development of thermostable Candida antarctica lipase B through novel in silico design of disulfide bridge*. Biotechnology and Bioengineering, 2012. **109**(4): p. 867-876.

Chapitre II. Matériels et Méthodes

A. PARTIE EXPERIMENTALE

1. Matériels

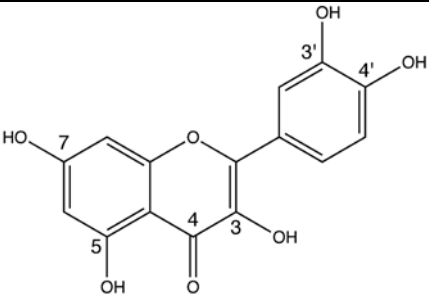
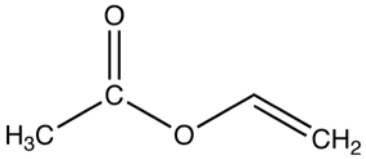
1.1. Descriptif du biocatalyseur

L'enzyme utilisée pour catalyser les réactions d'acylation est le Novozyme 435®. C'est une préparation industrielle produite par Novozyme Corporation et commercialisée par Sigma, résultant de l'immobilisation de la lipase B de *Candida antarctica* (CALB) sur une résine acrylique macroporeuse. Cette préparation se présente sous forme de granules blanches de diamètre compris entre 0.3 et 0.9 mm. Elle est fournie avec une activité de 7000 PLU.g⁻¹ (*Propyl Laurate Synthesis*), où une unité de propyle laurate (1PLU) est définie comme la quantité d'enzyme qui permet la transformation d'1 µmol d'acide laurique en laurate de propyle, par minute, en conditions standards et équimolaires (1-propanol/acide laurique = 1/1).

1.2. Réactifs et solvants

Le Tableau II.1 ci-dessous donne un descriptif exhaustif de l'ensemble des réactifs et solvants utilisés pour cette étude.

Tableau II.1. Descriptif des réactifs et solvants utilisés

Produit	Structure chimique	Pureté (%)	Fournisseur
Substrats			
Quercétine (M = 302 g.mol ⁻¹)		≥ 98	Sigma
Acétate de vinyle (M = 86,09 g.mol ⁻¹)		99	Aldrich
Solvants organiques			
Acétonitrile (M = 41,05 g.mol ⁻¹)	$\text{H}_3\text{C}-\text{C}\equiv\text{N}$	99.8	Aldrich
Acétonitrile (M = 41,05 g.mol ⁻¹)	$\text{H}_3\text{C}-\text{C}\equiv\text{N}$	Qualité CLHP	Fischer

2. Mise en œuvre des réactions d'acylation enzymatique

Les réactions de synthèse sont menées dans des bouteilles Schott hermétiques, chauffées à 60°C par l'intermédiaire d'un bain d'huile thermostaté (plaque chauffante thermorégulée IKA). Pour toutes les synthèses mises en œuvre, le volume réactionnel est de 10 mL. Le protocole général est le suivant :

- Séchage des bouteilles Schott dans une étuve à 120°C pendant 24h,
- Ajout du solvant (10 ml) séché au préalable avec du tamis moléculaire,
- Ajout de la quercétine et solubilisation pendant 15h à 60°C sous agitation par un barreau aimanté,
- Ajout de l'acétate de vinyle (5 ml) en excès molaire par rapport à la quercétine (ratio molaire de 330),
- Prélèvement et filtration (0.2 µm) de 1 ml du milieu réactionnel avant le début de la synthèse pour analyse,
- Déclenchement de la réaction par ajout de la lipase CALB (60 g.L⁻¹),
- Réaction pendant 3 jours d'incubation,
- Prélèvement et filtration (0.2 µm) du milieu réactionnel,
- Analyse des substrats et produits en utilisant la chromatographie liquide à haute performance.

En parallèle, deux témoins : l'un destiné à mettre en évidence d'éventuels problèmes de stabilité de la quercétine dans les conditions réactionnelles (solution de quercétine dans l'acétonitrile, 60°C pendant 3 jours sous agitation) ; l'autre correspondant au milieu de synthèse sans l'enzyme.

3. Méthodes analytiques

3.1. Analyse des milieux de synthèse par chromatographie liquide haute performance

3.1.1. Matériels et protocole de séparation

Le système CLHP utilisé est constitué de

- une colonne Alltima C18 (150*2,1 mm, 5 µm, Grace/Alltech, Darmstadt, Germany) munie d'une pré-colonne Alltima C18 (7,5*2,1 mm, 5 µm, Grace/Alltech, Darmstadt, Germany),
- un injecteur automatique intégrant un four à effet Peltier (ThermoFisher Scientific, SURVEYOR PLUS),
- une pompe (ThermoFisher Scientific, SURVEYOR MS),
- un détecteur PDA ou à barrette de diodes (ThermoFisher Scientific, SURVEYOR PDA PLUS),
- un logiciel d'acquisition et de traitement des chromatogrammes Xcalibur (version 2.1).

La colonne C18 permet de travailler en phase inverse : les constituants du mélange sont séparés via des interactions hydrophobes avec les chaînes aliphatiques greffées sur la phase stationnaire. Cette colonne est adaptée pour la séparation de la quercétine et de ses esters. Une méthode CLHP par

gradient est employée pour séparer les différents constituants présents dans les échantillons prélevés (Tableau II.2). Les analyses sont effectuées à une température de 25°C, avec un débit d'élution de 0,2 mL.min⁻¹.

Tableau II.2. Gradient d'élution utilisé pour la séparation de la quercétine et de ses esters.

Temps (min)	CH ₃ CN	H ₂ O (0,1 % d'acide acétique)
0	10	90
25	100	0
30	100	0

Le flavonoïde aglycone et ses esters sont détectés dans l'UV à une longueur d'onde de 400 nm.

3.1.2. Calcul du pourcentage de conversion des substrats

Le taux de conversion de la quercétine est calculé à partir des aires relevées sur les chromatogrammes (détection UV, 400 nm).

$$\%conversion = \frac{[A_i(ua)] - [A_f(ua)]}{[A_i(ua)]} * 100 \quad (II.1)$$

où A_i : aire de la quercétine présente dans le milieu initialement

A_f : aire de la quercétine présente dans le milieu à l'issue de la réaction

3.2. Analyse des milieux de synthèse par spectrométrie de masse

Dans cette étude, le spectromètre de masse utilisé est le LTQ (Thermo electron corporation®, E.U.) avec une source d'ionisation de type ESI (en anglais, *ElectroSpray Ionisation*). Cette dernière est utilisée en mode d'ionisation positif. Un couplage CL/SM, dans lequel les composés de l'échantillon analysé subissent une première séparation (cf méthode CLHP décrite ci-dessus) avant d'être introduits dans le spectromètre de masse, est appliqué aux différents échantillons. Ainsi, chaque pic chromatographique peut être associé à un spectre de masse permettant d'identifier la molécule par sa masse moléculaire (Figure II.1).

Pour déterminer précisément la structure des produits séparés par CL, certains spectres de masse sont réalisés en mode fragmentation. On parle de SMⁿ, n étant le nombre de niveaux de fragmentation imposé à un ion d'intérêt. L'ion d'intérêt est sélectionné par sa masse moléculaire grâce à un premier détecteur, puis il est fragmenté dans une chambre de collision. Un deuxième analyseur mesure les valeurs m/z des fragments ; on parle alors de SM-SM ou SM² (spectrométrie de masse en tandem). Pour ces analyses, la température du capillaire est fixée à 300°C et la tension à 48V. La confirmation structurale des ions fils obtenus est réalisée à l'aide du logiciel Xcalibur® soft (THERMO).

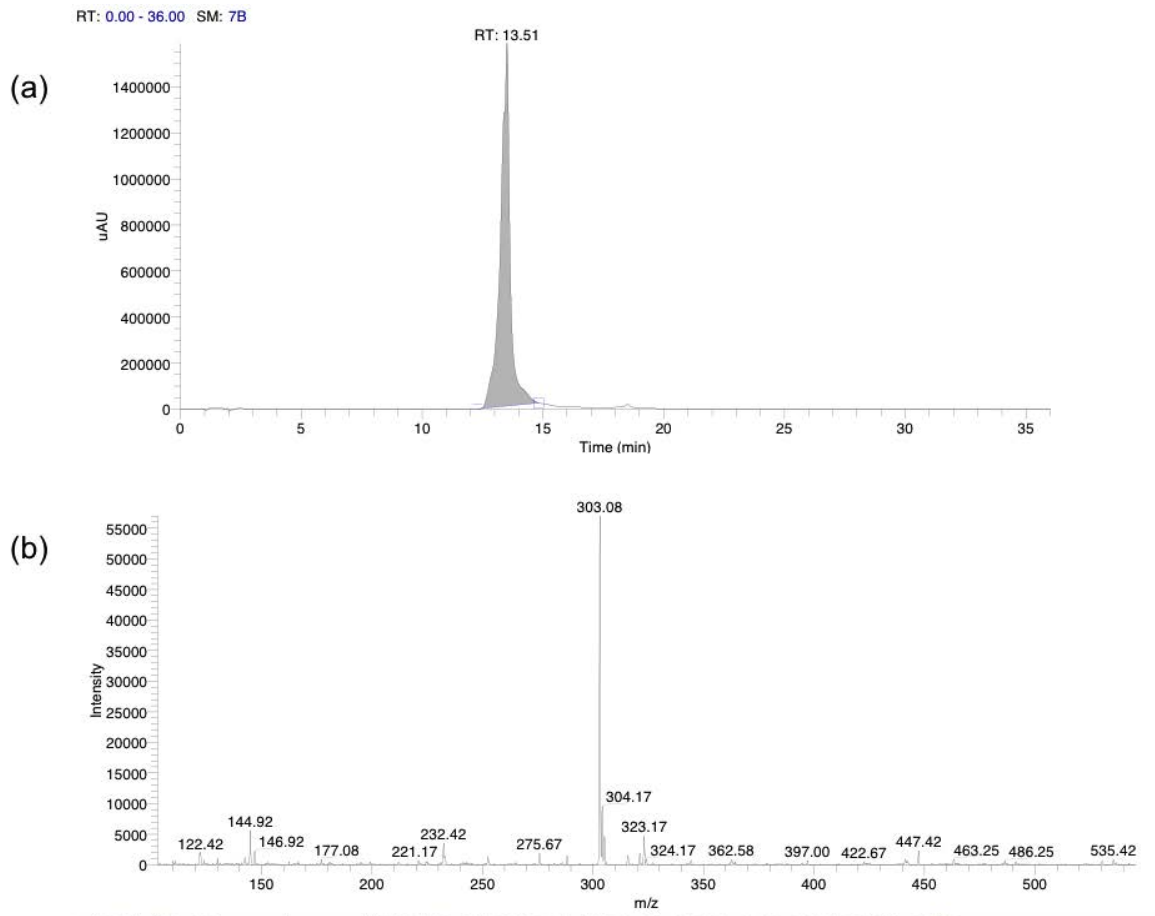


Figure II.1. Analyse CL-SM d'une solution témoin de quercétine dans l'acétonitrile. (a) Chromatogramme UV à 400 nm (b) Spectre de masse du pic présentant un temps de rétention de 13,51 min (mode d'ionisation positif)

B. MODÉLISATION MOLÉCULAIRE

1. Ressources informatiques

1.1. Ordinateurs

Un ordinateur, une station de travail pour visualisation stéréoscopique Alineos ainsi que deux serveurs de calcul sont utilisés pour l'exécution et l'analyse des simulations :

- Un processeur Intel Core Duo CPU 6550 (2.33 Ghz), équipé avec 2 Go de mémoire RAM
- Un bi-processeur AMD Opteron Shanghai 2384 4CPU (2.7 Ghz, 16 Go de mémoire)
- Un bi-processeur 4CPU (1.8 Ghz, 16 nœuds de calcul, 8 Go de mémoire RAM)
- Un bi-processeur Intel Xeon 4CPU L5420 (2.5 Ghz, 30 nœuds de calcul, 16 Go de mémoire RAM)

1.2. Logiciels de simulation moléculaire

- *Discovery Studio 2.5 (Accelrys, Inc)* : Ce logiciel comporte de nombreux modules de simulation moléculaire voués à l'étude structurale des protéines et des acides nucléiques, ainsi que de leurs interactions avec d'autres molécules. Les éléments utilisés dans cette étude sont : le champ de forces CHARMM [1] pour les simulations de mécanique moléculaire et de dynamique moléculaire ; le module *Solvation* pour ajouter des molécules de solvant et des contre-ions ; le module *LigandFit* [2] pour les simulations de docking ; le module *Consensus Score* [3] pour le classement des complexes à l'issue du docking ; le module *Analyze Ligand Poses* pour la classification des poses selon leur ressemblance ; le module *Calculate Binding Energies* [4] pour estimer l'énergie libre de liaison.
- *NAMD v.2.7 & v.2.8 (Not Another Molecular Dynamics program)* [5] : Ce logiciel est un code de dynamique moléculaire efficace pour les calculs en parallèle, conçu pour les simulations de systèmes biomoléculaires de grandes tailles.
- *VMD v1.8.7 & v1.9.1 (Visual Molecular Dynamics)* [6] : Ce logiciel est un programme de visualisation moléculaire pour l'affichage animé et l'analyse des systèmes biomoléculaires de grandes tailles à l'aide d'une interface graphique et de la conception de scripts.

2. Préparation des structures

2.1. Structures des ligands

Les structures tridimensionnelles des molécules de quercétine, de ses dérivés et de l'isoquercitrine (Figure II.2) ont été construites en utilisant les outils du module *Builder* de Discovery Studio 2.5. Les types d'atomes des substrats ont été fixés en appliquant le champ de forces CHARMM général. Les charges partielles des atomes ont été assignées avec la méthode CFF. Les

structures résultantes ont été soumises à une correction géométrique ('*clean geometry*') afin d'ajuster la longueur des liaisons. L'énergie des systèmes a été minimisée pour s'assurer de la viabilité énergétique des différents substrats. Pour cela, le protocole appliqué est une minimisation de 2000 itérations avec l'algorithme de descente abrupte (en anglais, *Steepest Descent*) en considérant l'effet de solvation grâce au modèle '*Generalized Born with a Molecular Volume*' (GBMV).

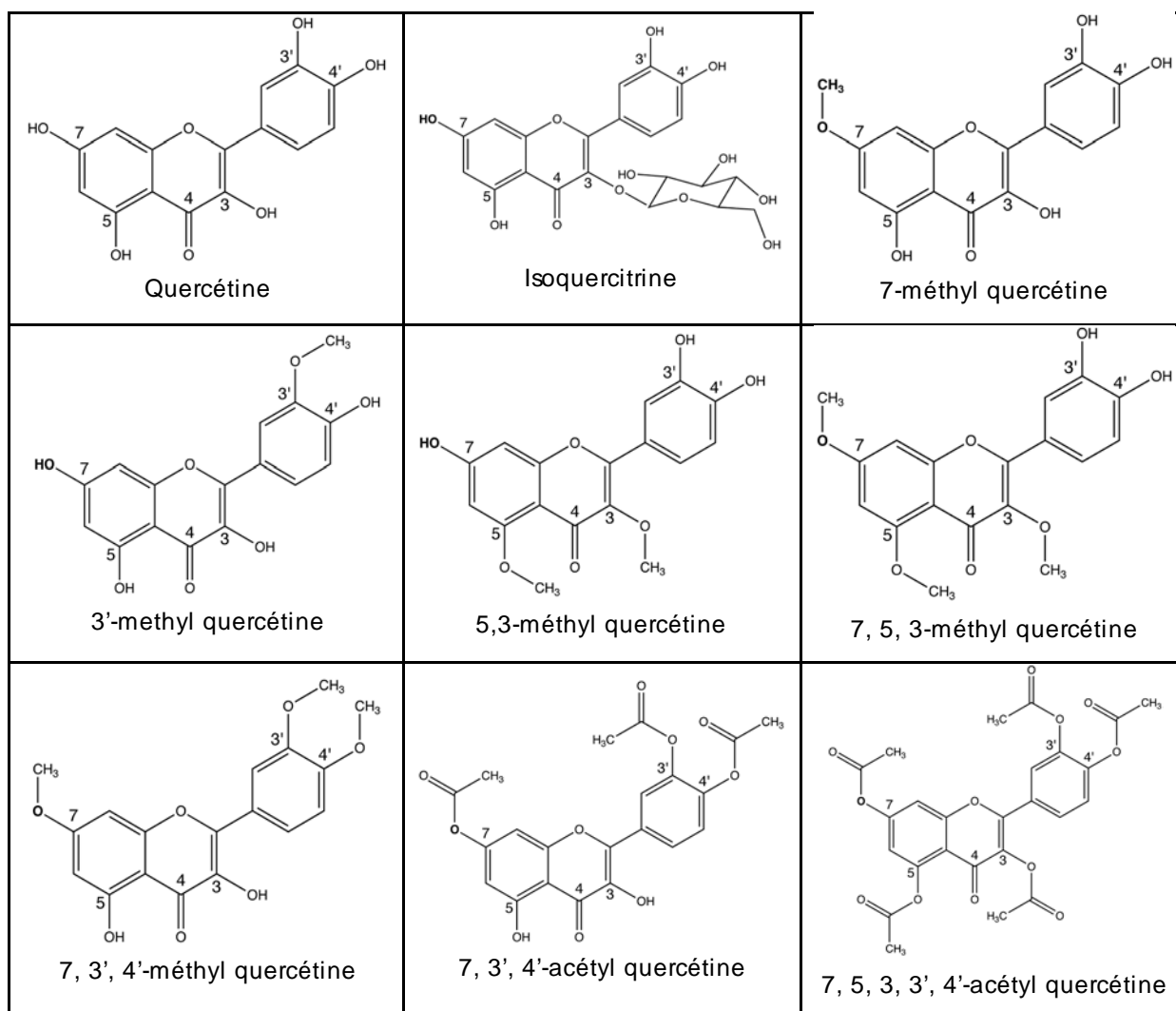


Figure II.2. Structure chimique, numérotation systématique et dénomination des ligands étudiés

2.2. Structures des lipases

Les travaux de modélisation ont porté sur deux biocatalyseurs: la lipase B de *Candida antarctica* et la lipase de *Pseudomonas cepacia*.

2.2.1. Structures de la lipase B de *Candida antarctica* (CALB)

Cinq formes cristallines de la lipase B de *Candida antarctica* ont été résolues et sont disponibles sur la PDB. Trois d'entre elles, une orthorhombique (PDB ID: 1TCA) et deux monocliniques (PDB ID: 1TCB et 1TCC), sont issues de la CALB « wild-type », c'est-à-dire sans

aucune molécule de ligand lié au site actif [7]. Les deux autres structures ont été obtenues par co-cristallisation de la CALB soit avec un tensioactif Tween80 (PDB ID : 1LBT), soit avec un inhibiteur de type phosphonate (PDB ID : 1LBS) [8]. Dans ces systèmes, l'inhibiteur et le tensioactif sont liés à la sérine du site actif de façon covalente et non covalente, respectivement.

Dans les études de modélisation moléculaire de cette lipase, les structures cristallines 1LBS et 1TCA sont les plus citées. En effet, la première présente l'avantage d'être résolue avec un inhibiteur phosphonate fixé de manière covalente sur la sérine catalytique. La présence de ce ligand fournit des informations expérimentales sur le mode d'ancrage et l'orientation du ligand vis-à-vis du trou oxyanionique. La seconde structure, quant à elle, possède la meilleure résolution (1.55 Å vs 2.60 Å pour 1LBS). Cependant, aucun ligand co-cristallisé n'est présent dans la cavité du site actif. Le *RMSD* résultant de la superposition du squelette peptidique de ces deux structures est de 0.33 Å, indiquant qu'elles ne présentent pas de différence conformationnelle significative. Dans le cadre de ce travail, le choix s'est porté sur la structure 1LBS qui est un compromis entre une résolution correcte et la présence d'un ligand co-cristallisé.

Le cristal 1LBS est composé de six chaînes protéiques par cellule. Les dimensions de chaque cellule sont $a = 229.5 \text{ \AA}$, $b = 95.6 \text{ \AA}$, $c = 86.8 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$. Chaque chaîne est constituée de 317 acides aminés, un inhibiteur éthyl-héxyl-phosphonate (HEE) lié de manière covalente à la sérine catalytique (Ser105) et un dimère de N-acétyl-glucosamine (NAG) lié au résidu Asn74. De plus, 92 molécules d'eau sont adsorbées à la chaîne A.

La préparation des différentes structures utilisées pour les simulations est détaillée dans le Tableau II.3. La CALB 'native' et la CALB acétylée sont des structures de référence. La première est dépourvue de ligand ; elle est obtenue par suppression de l'inhibiteur HEE dans la structure 1LBS. La seconde présente un acétate lié de façon covalente à la Ser105 ; elle est obtenue en remplaçant l'inhibiteur HEE par une entité acétate dans le système 1LBS. La dénomination CALB mutée correspond à des structures dans lesquelles un ou deux acides aminés ont été remplacés par d'autres acides aminés.

Tableau II.3. Détail des différentes étapes de préparation des structures initiales de la CALB.

	CALB 'native'	CALB acétylée	CALB mutée
Conservation de la chaîne A et suppression des autres chaînes	X	X	X
Suppression de l'inhibiteur HEE et du dimère NAG	X	X	X
Conservation des molécules d'eau adsorbées	X	X	X
Remplacement de la sérine catalytique par une sérine acétylée		X	
Ajustement manuel de la position de l'acétate			X
Remplacement du/des résidu(s) choisi(s) par un/des autre(s) résidu(s)			X
Addition des atomes d'hydrogène	X	X	X
Ajustement à pH neutre de l'état de protonation des résidus	X	X	X
Solvatation de la structure avec des molécules d'eau TIP3P	X	X	X
Neutralisation des systèmes avec des contre-ions	X	X	X

2.2.2. Structures de la lipase de *Pseudomonas cepacia* (PCL)

Neufs structures cristallographiques de la lipase de *Pseudomonas cepacia* sont disponibles sur la PDB. Rappelons que cette enzyme possède un clapet qui recouvre le site actif ; de ce fait, deux structures ont été déterminées, correspondants aux formes 'ouverte' et 'fermée' de la lipase. Dans le cadre de ce travail, l'étude de l'ancrage de ligands dans le site actif de l'enzyme oriente le choix vers une forme ouverte de la lipase. La structure 3LIP possédant la meilleure résolution (2.00 Å) parmi les différentes formes ouvertes disponibles a été retenue.

La cellule du cristal 3LIP est constituée d'une seule chaîne protéique et présente les dimensions suivantes : $a = 91.3 \text{ \AA}$, $b = 47.3 \text{ \AA}$, $c = 85.4 \text{ \AA}$, $\alpha = \gamma = 90.0^\circ$, $\beta = 121.4^\circ$. Cette chaîne est composée de 320 acides aminés, 193 molécules d'eau adsorbées et un cation calcium complexé par les carboxylates des chaînes latérales des résidus Asp242 et Asp288. Deux structures ont été préparées (tableau II.4) : la PCL 'native' dépourvue de ligand et la PCL acétylée résultant du greffage covalent d'une entité acétate sur la sérine catalytique (Ser87).

Tableau II.4. Détail des différentes étapes de préparation des structures de la PCL.

	PCL 'native'	PCL acétylée
Conservation des molécules d'eau adsorbées	X	X
Conservation de l'ion calcium	X	X
Remplacement de la sérine catalytique par une sérine acétylée		X
Addition des atomes d'hydrogène	X	X
Ajustement à pH neutre de l'état de protonation des résidus	X	X
Solvatation de la structure avec des molécules d'eau TIP3P	X	X
Neutralisation du système avec des contre-ions	X	X

3. Simulations de dynamique moléculaire

3.1. Paramètres des simulations

3.1.1. La température

La température (T) à laquelle sont menées les simulations de dynamique moléculaire est un paramètre important car c'est en fonction de celle-ci que les vitesses initiales des atomes sont établies. Le plus souvent, elles sont déterminées aléatoirement selon une distribution de Maxwell-Boltzmann (Eq. II.2). En outre, la température est directement reliée à l'énergie cinétique du système :

$$\frac{k_B T}{2} (3N - N_c) = \sum_{i=1}^N \frac{|p_i|^2}{2m_i} \quad (\text{II.2})$$

où k_B est la constante de Boltzmann, T est la température, N_c est le nombre de contraintes, $(3N - N_c)$ est le nombre total de degrés de liberté, p_i est le moment total de la particule i et m_i la masse de la particule.

3.1.2. Le pas d'intégration

Le pas d'intégration (Δt) (en anglais, *timestep*) est choisi selon la fréquence de vibration la plus rapide des molécules étudiées. Pour les systèmes biologiques, c'est la fréquence vibrationnelle des liens hydrogène qui sont les plus rapides (10^{-14} s.). Pour assurer une intégration idéale des équations du mouvement, le pas d'intégration doit être un dixième de cette fréquence, soit 1 fs (10^{-15} s.).

3.1.3. L'algorithme SHAKE

L'algorithme SHAKE [9] est utilisé pour contraindre les liens hydrogène à une longueur fixe. L'avantage est d'accélérer les calculs car la position des atomes d'hydrogène n'est pas calculée selon les équations du mouvement mais plutôt en fonction de la position de l'atome lourd auquel il est lié. L'application de cet algorithme permet notamment l'utilisation d'un pas d'intégration de 2 fs.

3.1.4. Les ensembles thermodynamiques

Les ensembles thermodynamiques permettent de simuler différents types de phénomènes physiques :

- L'ensemble microcanonique (ou *NVE*) est caractérisé par un nombre N de particules, un volume et une énergie totale constants. La pression et la température fluctuent. Ceci correspond à un processus adiabatique sans échange de chaleur.
- L'ensemble canonique (ou *NVT*) est caractérisé par un nombre N de particules constant, un volume constant et une température constante. Dans ce cas l'énergie du système est échangée avec un thermostat afin de maintenir la température constante.
- L'ensemble isothermale-isobare (ou *NPT*) est caractérisé par un nombre N de particules, une pression et une température constants. La pression est maintenue constante par fluctuation du volume. De nombreuses mesures expérimentales sont réalisées dans des conditions de température et de pression constantes ; de ce fait, l'ensemble *NPT* apparaît comme le plus pertinent pour corréler les modèles aux données expérimentales. Par ailleurs, cet ensemble doit être utilisé pour des systèmes soumis à des conditions périodiques frontalières.

3.1.5. Distance de troncation et méthode 'Particle Mesh Ewald'

Le calcul de l'énergie des termes non liés au cours d'une simulation de dynamique moléculaire ou d'une minimisation d'énergie est la partie la plus coûteuse en temps. Le nombre de termes non liés qui doit être évalué est proportionnel au carré du nombre de particules et est donc de l'ordre N^2 . Par ailleurs, les interactions intermoléculaires sont négligeables au-delà d'une certaine distance. Comme illustré à la Figure II.3, le potentiel de Lennard-Jones, modélisant les interactions de van der Waals, tend rapidement vers zéro et les interactions électrostatiques sont inversement proportionnelles à la distance entre deux atomes. De ce fait, les interactions entre atomes espacés d'une certaine distance sont omises pour accélérer le temps de calcul. Cette distance est appelée *distance de troncation* (en anglais, *cutoff*).

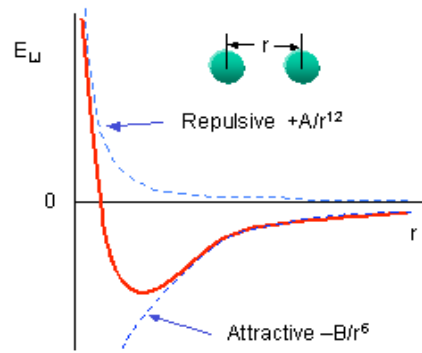


Figure II.3. Représentation graphique du potentiel de Lennard-Jones 6-12.

L'introduction d'une coupure radicale à la distance de troncation introduit des artefacts dans les interactions moléculaires. En effet, même si les interactions sont plutôt faibles au-delà de la distance de troncation, l'effet global dans un système comportant plusieurs milliers d'atomes est important. Une façon d'atténuer cet artefact est d'employer une fonction mathématique. Celle-ci ramène progressivement à zéro les interactions exercées par un atome sur les autres atomes du système. Parmi les différentes méthodes existantes, la fonction dénommée '*switching*' est représentée à la Figure II.4.

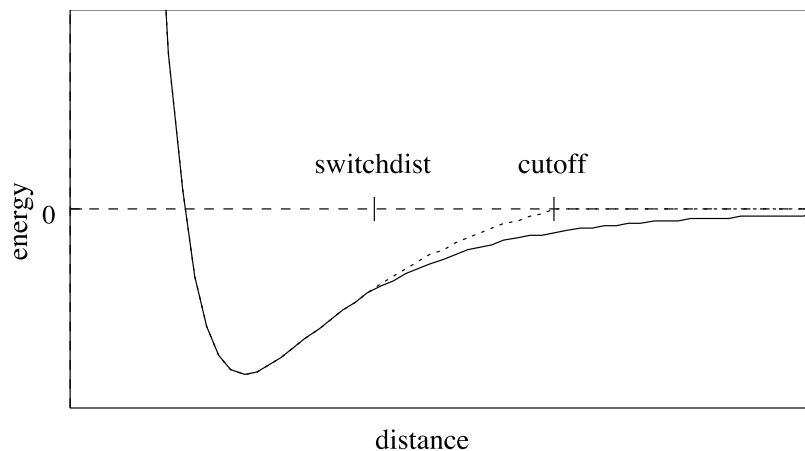


Figure II.4. Potentiel de van der Waals en présence (trait discontinu) et en absence (trait continu) de la fonction 'switching'. L'application de la fonction 'switching' réduit progressivement le potentiel à zéro entre la distance 'switch' et la distance de troncation. (d'après [10]).

La méthode 'Particle Mesh Ewald' est une méthode permettant d'éviter les approximations liées aux distances de troncation pour le calcul des interactions électrostatiques. Dans cette méthode, le calcul des interactions électrostatiques est séparé en deux parties : (i) les interactions de courte portée (E_{sr}) qui sont calculées avec une sommation traditionnelle et (ii) les interactions de longue portée (E_{lr}) qui incluent les interactions supérieures à la distance de troncation et qui sont calculées en utilisant un algorithme de Transformation de Fourier Rapide (en anglais *Fast Fourier Transform*, FFT).

3.1.6. Conditions périodiques frontières

Lors de simulations menées sur des systèmes solvatés, deux problèmes sont rencontrés aux frontières du système : (i) les interactions non liées sont discontinues et (ii) les molécules en

mouvement dans cette région ne sont pas contraintes et pourraient s'échapper. La solution à ces problèmes est l'application de conditions périodiques aux frontières du système (en anglais *Periodic Boundary Conditions*). Selon ces conditions, le système de base composé de plusieurs molécules est positionné au centre, et pour chacune des faces exposées, il y a une réplique virtuelle du système (Figure II.5). Chaque particule interagit avec les autres particules du système ainsi qu'avec ses images virtuelles. Ces répliques sont utilisées pour calculer les interactions non liées des molécules localisées aux frontières. En outre, l'application de conditions périodiques frontières permet la disparition de l'effet frontière. En effet, le déplacement d'une molécule vers une réplique virtuelle entraîne l'intégration d'une molécule de la réplique de la frontière opposée au système de base.

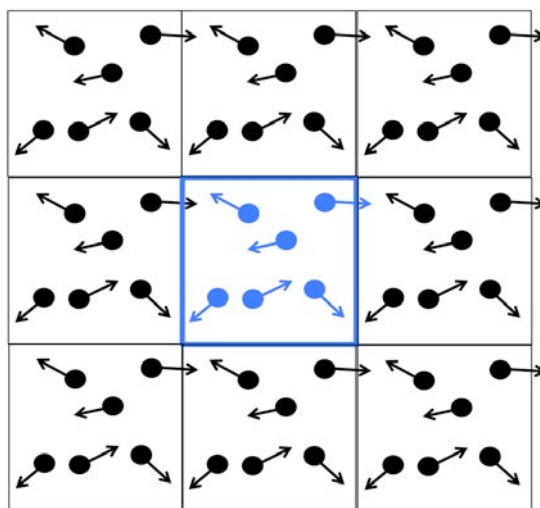


Figure II.5. Représentation en deux dimensions des conditions périodiques frontières. Chaque côté du système réel (bleu) est délimité par une réplique virtuelle.

Théoriquement, le système de base peut être reproduit indéfiniment dans les trois directions de l'espace. Néanmoins, sa reproduction en triplicata dans chacune des trois directions représente le cas le plus fréquent.

3.2. Conditions des simulations de dynamique moléculaire

Toutes les simulations de dynamique moléculaire ont été menées sur des serveurs de calcul en utilisant le programme NAMD v.2.7 & v.2.8 ainsi que les paramètres et les fichiers de topologies du champ de forces CHARMM implémentés dans *Discovery Studio 2.5*. Les différentes conditions appliquées lors des simulations sont les suivantes :

- La température du système a été fixée à 300 K
- Le pas d'intégration (Δt) a été fixé à 1 fs
- Les interactions électrostatiques à courte portée ont été calculées toutes les 1 fs
- Les interactions électrostatiques à courte et à longue portée ont été évaluées toutes les 2 fs
- L'ensemble canonique ou NPT (nombre de particules, une pression et une température constante) a été considéré comme ensemble thermodynamique.
 - Le contrôle de la température a été effectué via une dynamique de Langevin où la température a été fixée à 300 K et la fréquence de collision à 1 ps⁻¹.

- Le contrôle de la pression a été réalisé en utilisant le piston Nosé-Hoover de Langevin où la pression a été fixée à 1 atm, la période d'oscillation du barostat à 75 fs et l'échelle de temps de friction à 25 fs.
- L'utilisation de l'ensemble canonique requiert l'application de conditions périodiques frontières sur le système.
- Les interactions électrostatiques ont été traitées avec la méthode 'Particle Mesh Ewald'.
- La distance de troncation (cutoff) a été fixée à 12 Å.
- La fonction 'switching' a été appliquée aux interactions de van der Waals entre 10 et 12 Å.
- Les coordonnées ont été sauvées toutes les 1 ps.
- La liste des atomes voisins pour les interactions non liées a été actualisée toutes les 10 étapes d'intégration.
- Les liens entre les atomes lourds et les atomes d'hydrogène ont été maintenus rigides (algorithme SHAKE)

3.3. Protocole des simulations de dynamique moléculaire

Des simulations de dynamique moléculaire ont été menées d'une part pour l'étude de la stabilité structurale des enzymes cibles et d'autre part pour l'étude structurale des complexes flavonoïde-lipase. Le protocole détaillé ci-dessous a été appliqué pour toutes les simulations réalisées dans le cadre de ce travail (Figure II.6).

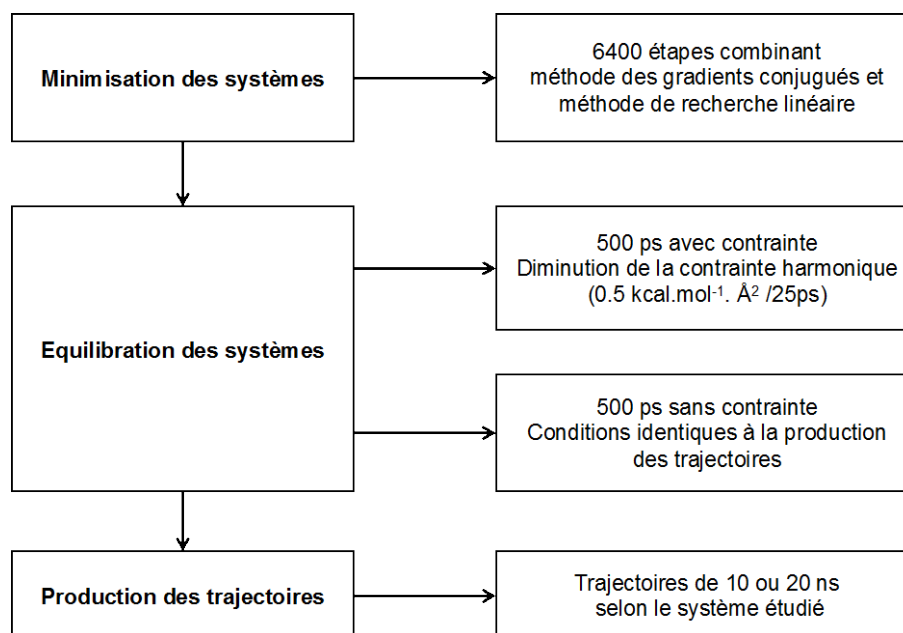


Figure II.6. Représentation schématique du protocole des simulations de dynamique moléculaire.

3.3.1. Minimisation des systèmes

Pour chaque système moléculaire étudié, une minimisation de 6400 étapes a été effectuée avec la méthode des gradients conjugués et la méthode de recherche linéaire qui sont des

algorithmes implémentés dans NAMD. Brièvement, la méthode des gradients conjugués est employée pour choisir les directions sur la surface d'énergie. Et la méthode de recherche linéaire est utilisée pour localiser le minimum dans chaque direction.

3.3.2. Équilibration des systèmes

Cette étape correspond au début de l'intégration des équations de mouvement et donc, au début de la dynamique. L'objectif de la phase d'équilibration est de permettre au système d'évoluer de la configuration de départ vers une configuration d'équilibre atteinte à la température choisie pour les simulations. L'équilibration se poursuit jusqu'à ce que les valeurs d'un ensemble de propriétés soient stables. En général, les propriétés suivies sont des grandeurs thermodynamiques telles que l'énergie, la température et la pression ainsi que des propriétés structurales.

La phase d'équilibration des systèmes a été divisée en 2 parties :

- durant les 500 premières ps, les systèmes ont été soumis à une contrainte harmonique. Celle-ci a été fixée initialement à $10 \text{ kcal.mol}^{-1}.\text{Å}^2$ et progressivement diminuée, à raison de $0.5 \text{ kcal.mol}^{-1}.\text{Å}^2$, toutes les 25 ps.
- Ensuite, l'équilibration a été menée sans contrainte dans les mêmes conditions que la phase de production durant 500 ps complémentaire.

3.3.3. Production des trajectoires de dynamique moléculaire

C'est l'étape du protocole la plus coûteuse en temps de calcul. A titre indicatif, une nanoseconde de trajectoire est calculée en six heures pour un système composé d'environ 50000 atomes, sur 12 nœuds d'un serveur de calcul. La phase de production des différentes simulations a été effectuée avec les paramètres présentés dans la section 3.2. Selon les systèmes étudiés, des durées de trajectoires de 10 ou 20 ns ont été appliquées.

3.3.4. Analyse des résultats

Les trajectoires de dynamique moléculaire fournissent des informations sur les propriétés conformationnelles des systèmes moléculaires et la façon dont les changements de conformation s'opèrent avec le temps. L'analyse des simulations est facilitée par la réalisation de graphes montrant l'évolution de plusieurs paramètres au cours du temps (angles, distances, RMSD, ...).

Les trajectoires ont été exploitées dans un premier temps pour l'analyse structurale des différents systèmes moléculaires étudiés : les lipases natives CALB et PCL en milieu solvato, les complexes flavonoïde-lipases, les lipases mutées et les complexes flavonoïde-lipases mutées. Des paramètres tels que le RMSD (en anglais, *Root Mean Square Deviation*), le RMSD 2D, le RMSF (en anglais, *Root Mean Square Fluctuation*), le volume de la poche catalytique ont été calculés pour les différentes trajectoires afin de vérifier la stabilité des systèmes, de mettre en évidence et de localiser des changements structuraux importants. Le RMSD ou l'écart quadratique moyen est une mesure de la différence entre les positions des atomes (généralement le C α) des conformations de la protéine au cours de la trajectoire :

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2} \quad (II.X)$$

où N est le nombre d'atome et δ est la distance entre les N paires d'atomes. Le RMSD 2D est une comparaison de toutes les conformations entre elles au cours du temps. Le RMSF ou la fluctuation quadratique moyenne est une mesure de la fluctuation entre la position d'un atome i (généralement le $C\alpha$) et une position de référence :

$$RMSF = \frac{1}{T} \sum_{t=1}^T (x_i(t) - \bar{x}_i)^2 \quad (II.X)$$

où T est le temps sur lequel la position moyenne \bar{x}_i est calculée.

Dans un second temps, les distances entre l'oxygène de l'acétate (Ace:O) lié à la sérine catalytique et les atomes d'hydrogène des résidus du trou oxyanionique ont été mesurées. Ces distances indiquent si l'acétate présente ou non une orientation correcte dans le site actif et donc si le complexe substrats-lipase est susceptible de mener à la formation d'un ester. En outre, les distances entre les résidus constituant la triade catalytique ont été suivies au cours des trajectoires afin de déterminer l'établissement ou non de liaisons hydrogène. Ces liaisons hydrogène stabilisent l'histidine catalytique dans une position adéquate pour le transfert du proton lors de l'attaque nucléophile du substrat sur le carbone de l'acétate fixé à la sérine catalytique.

La démarche d'analyse des trajectoires de dynamique moléculaire est la suivante :

- (i) Découpage des trajectoires de 10 et de 20 ns toutes les 5 et 10 ps, respectivement. Les analyses ont ainsi été effectuées sur 2000 frames.
- (ii) Mesure des distances d'intérêt à travers les 2000 frames
- (iii) A partir de la visualisation des animations, identification des interactions entre les résidus de la cavité et les substrats.

Dans un troisième temps, les liaisons hydrogène formées entre les résidus des sites actifs et les groupements hydroxyles de la quercétine ont été identifiées. Les critères retenus pour une liaison hydrogène sont (Figure II.7) :

- (i) une distance entre le donneur et l'accepteur d'hydrogène inférieure ou égale à 3 Å ; et,
- (ii) un angle entre le donneur d'hydrogène, l'hydrogène et l'accepteur d'hydrogène supérieure ou égale à 120 °.

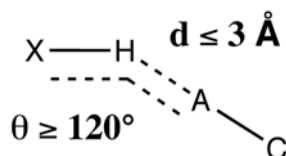


Figure II.7. Critères pour établir la présence d'une liaison hydrogène entre un donneur (X) d'hydrogène et un accepteur (A) d'hydrogène.

4. Simulations de docking

4.1. Considérations générales

Le docking moléculaire (ancrage ou amarrage moléculaire en français) est une méthodologie de simulation informatique qui prédit la structure de complexes intermoléculaires formés par 2 molécules ou plus. Les complexes intermoléculaires sont généralement constitués d'un récepteur macromoléculaire qui est la cible du docking et dont la structure tridimensionnelle est connue, et d'un ligand ou d'un substrat qui est une petite molécule. La plupart des algorithmes de docking sont capables de générer un grand nombre de structures et de ce fait, ils requièrent un moyen de classification afin d'identifier les structures d'intérêt. Une simulation de docking comporte donc une phase de génération des complexes intermoléculaires, suivie d'une phase d'évaluation des structures plausibles de ces complexes :

- La génération de complexes intermoléculaires s'exécute généralement par l'intermédiaire d'algorithmes qui positionnent le ligand dans le récepteur et échantillonnent l'espace conformationnel du ligand et parfois du récepteur.
- L'évaluation des complexes intermoléculaires générés est réalisée par des fonctions de score qui estiment les poses en termes d'énergie d'interaction avec le récepteur. Elles doivent permettre d'identifier la ou les orientations qui correspondent le mieux à la véritable structure du complexe intermoléculaire et d'évaluer l'affinité de liaison entre le ligand et le récepteur. Les fonctions de score évaluent les différentes contributions énergétiques dans les complexes ligand-récepteur. Ces derniers sont alors classés de manière à identifier les modes d'interactions ligand-récepteur les plus probables.

La mise en œuvre d'une simulation de docking comprend la définition du site actif, la génération des complexes, la sélection des complexes, le classement des complexes sélectionnés. Toutes les simulations de docking ont été menées avec le module LigandFit de *Discovery Studio 2.5*. (utilisation de la complémentarité de forme entre ligand et récepteur).

4.2. Définition du site actif

4.2.1. Principe

Le module LigandFit offre un outil pour la détection de cavités dans les protéines. Cet outil se base sur l'algorithme de remplissage par diffusion (en anglais, *flood-filling algorithm*) et fonctionne en deux étapes successives :

- (i) Une grille tridimensionnelle avec un espacement de 0.5 Å est construite en se basant sur les dimensions (x, y, z) de la protéine. Chaque point de la grille est classé comme 'occupé' si la distance de contact la plus proche de l'atome lourd de la protéine est inférieure ou égale à 2.5 Å ; dans le cas contraire, le point de la grille est classé comme 'libre'. (Figure II.7.a)

- (ii) En partant de l'extérieur de la grille, les points libres sont effacés jusqu'à ce que les nuages de points restants aient une largeur maximale prédéfinie par l'utilisateur. Chaque nuage de points conservé représente une cavité (Figure II.7.b).

Les protéines ayant très souvent des topologies de surface assez irrégulières, plusieurs cavités sont détectées par cette procédure automatique. Il appartient au modélisateur, en se basant sur ses connaissances de la protéine étudiée, de sélectionner parmi les cavités détectées celle qui constitue le site de liaison dans lequel sera effectué le docking du ligand.

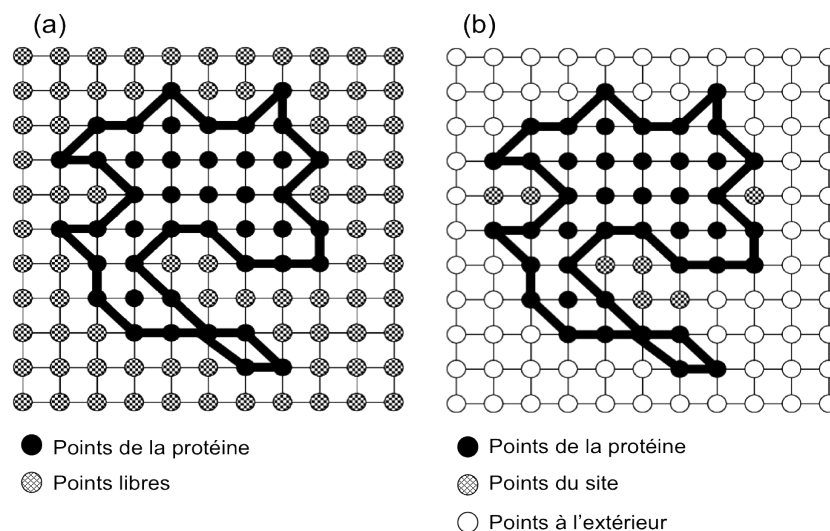


Figure II.8. Représentation schématique de la grille étendue à travers la protéine. (a) Les points de la grille 'occupés' par la protéine sont représentés par des cercles noirs tandis que les points 'libres' sont représentés par des cercles ombragés. (b) Les cercles blancs représentent des points de la grille qui ont été effacés et qui laissent apparaître des cercles ombragés identifiés comme des points de cavités [2].

4.2.2. Définition du site actif des lipases CALB et PCL

Le site actif des lipases pour le docking des ligands a été défini en se basant sur la description des cavités, à partir des données cristallographiques trouvées dans la littérature. Une largeur maximale de 5 Å a été choisie pour la détection des cavités.

- (i) La cavité catalytique de la lipase CALB [7] est étroite et profonde, avec des dimensions approximatives de 10 Å x 4 Å de largeur et 12 Å de profondeur. La triade catalytique, constituée par les résidus Ser105, His224 et Asp187, est localisée au fond de la cavité. Les résidus du trou oxyanionique sont Thr40 et Gln106. En outre, la cavité de cette enzyme est essentiellement hydrophobe, tapissée par les résidus aliphatiques Leu140, Ala141, Leu144, Val149, Val154, Ile189, Leu278, Ala281, Ala282, Ile285 et Val286. Seule une petite région autour des résidus catalytiques possède un caractère hydrophile, dû à la présence des résidus Thr40, Asp134 et Gln157.
- (ii) La cavité du site actif de la lipase PCL [11] est large et allongée avec des dimensions approximatives de 17 Å x 7 Å de largeur et 8 Å de profondeur. La triade catalytique, localisée dans la partie centrale du fond de la cavité, est constituée par les résidus Ser87, His286 et Asp264. Le trou oxyanionique est constitué des résidus Leu17 et Gln88. La

partie interne de la cavité présente deux régions : une région hydrophobe constituée par les résidus Leu17, Pro113, Phe119, Val123, Leu164, Leu167, Val266 et Val267, et une région hydrophobe/hydrophile, délimitée par les résidus Gln16, Thr18, Tyr23, Tyr29, His86, Leu287, Ile290, Gln292 et Leu293. Ces deux régions sont séparées par un étranglement formé par les acides aminés Val266 et Leu17.

Après la détection des cavités des lipases, une sphère de 12 Å de rayon a été centrée sur le site actif détecté. La spécification d'une sphère permet d'ajuster les atomes d'intérêt *via* des minimisations d'énergies localisées.

4.3. Génération des complexes flavonoïde-lipase

4.3.1. Principe

La génération et la sélection de conformères du ligand se font de la façon suivante :

- Le nombre maximal de conformères du ligand à générer (N_{trials}) et le nombre maximal de ces conformères à sauvegarder (N_{save}) sont définis par le modélisateur.
- En partant d'une conformation initiale arbitraire n du ligand, chaque angle de torsion de la molécule est soumis à une incrémentation aléatoire [12], ce qui conduit à un nouveau conformère $n+1$. Tous les conformères ainsi générés sont placés dans le site actif selon quatre orientations différentes (Figure II.8), conduisant à quatre poses du ligand (une pose correspond à une position et à une orientation données d'un conformère). Les complexes résultants sont conservés dans la liste de sauvegarde, jusqu'à ce que le nombre maximal de conformères de cette liste (N_{save}), prédéfini par le modélisateur, soit atteint. Quand la liste est pleine, la prochaine pose générée est comparée avec chacune de celles de la liste précédemment constituée, selon deux critères : l'adéquation géométrique avec la cavité du site actif et l'énergie potentielle d'interaction dans le complexe formé.

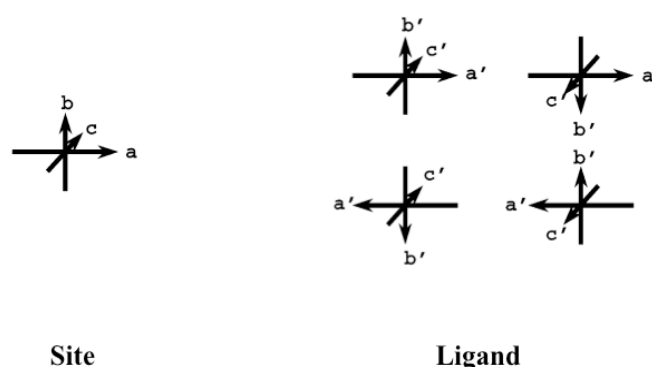


Figure II.9. Représentation schématique des quatre orientations du ligand dans le site actif du récepteur [2].

- L'adéquation de forme entre les poses du ligand et le site actif est évaluée à l'aide d'un principe mathématique de comparaison. Le site actif ainsi que les conformations du ligand sont représentés par un ensemble de points dans l'espace cartésien tridimensionnel. A partir des sommations des

coordonnées (x, y, z) de ces points, une matrice M est définie pour la pose du ligand et une autre pour le site du récepteur :

$$M = \begin{bmatrix} \sum x^2 & \sum xy & \sum xz \\ \sum xy & \sum y^2 & \sum yz \\ \sum xz & \sum yz & \sum z^2 \end{bmatrix} \quad (\text{II.3})$$

En utilisant les valeurs propres de la matrice de divergence de forme du ligand (L_1, L_2, L_3 ; $L_1 > L_2 > L_3$) et celle du site du récepteur (S_1, S_2, S_3 ; $S_1 > S_2 > S_3$), l'écart de forme ρ (Eq. II.4) est calculé. Plus la valeur de ρ s'approche de zéro, plus les deux formes se ressemblent.

$$\rho = \sqrt{\left(\frac{S_1 - L_1}{S_2 - L_2}\right)^2 + \left(\frac{S_2 - L_2}{S_3 - L_3}\right)^2 + \left(\frac{S_1 - L_1}{S_3 - L_3}\right)^2} \quad (\text{II.4})$$

Si la valeur de ρ est supérieure à celle de la structure la plus mauvaise dans la liste (soit celle présentant la valeur de ρ la plus élevée), la pose en test est rejetée. Dans le cas contraire, la pose est confrontée à un second filtre de type énergétique avant d'être acceptée dans la liste de sauvegarde.

- Pour les poses qui ont passé le filtre géométrique présenté précédemment, l'énergie d'interaction E (Eq. II.5) est calculée comme la somme des contributions de van der Waals (représentées par un terme de Lennard Jones 9-6) et électrostatiques (représentées par un terme de Coulomb).

$$E = \sum_{i,j} \varepsilon_{ij} \left[2 \left(\frac{r_{ij}^*}{r_{ij}} \right)^9 - 3 \left(\frac{r_{ij}^*}{r_{ij}} \right)^6 \right] + \frac{332.0716}{\varepsilon} \sum_{i,j} \frac{q_i q_j}{r_{ij}} \quad (\text{II.5})$$

où $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$; $r_{ij}^* = \sqrt{r_i^* r_j^*}$; r_i^* et ε_i sont respectivement le rayon de van der Waals et les paramètres de l'énergie de l'atome i du ligand, tandis que r_j^* et ε_j sont les paramètres similaires pour l'atome j de la protéine ; r_{ij} est la distance entre l'atome i du ligand et l'atome j de la protéine. Dans le terme électrostatique, q_i et q_j sont les charges respectives de l'atome i du ligand et de l'atome j de la protéine et ε est la constante diélectrique.

Si la valeur de l'énergie d'interaction de la pose en test est supérieure à celle de la structure la plus mauvaise dans la liste précédemment établie (soit celle présentant la valeur E la plus élevée), la pose est rejetée ; dans le cas contraire, elle remplace la structure classée dernière de la liste.

Cette procédure est répétée pour chaque conformère généré aléatoirement, jusqu'à ce que le nombre maximal de conformères (N_{trials}), défini par le modélisateur, soit atteint. La liste de sauvegarde est continuellement mise à jour, en éliminant les poses défavorables et les poses redondantes. Seules les meilleures structures, d'un point de vue géométrique et énergétique, sont donc accumulées au cours du processus.

4.3.2. Docking des flavonoïdes quercétine, isoquercitrine et des dérivés de la quercétine

Les paramètres appliqués dans la procédure de docking des différents ligands dans le site actif des lipases sont présentés dans le Tableau II.5. Les paramètres qui ne sont pas spécifiés dans le tableau adoptent les valeurs par défaut proposées par le logiciel.

Tableau II.5. Spécification des paramètres utilisés lors des simulations de docking.

Paramètres	Valeurs	Remarques
Champ de forces pour le calcul de l'énergie	CFF	Ce champ de forces est employé pour le calcul des énergies d'interaction
Nombre maximal de conformères générés (N_{trials})	15000	Un nombre élevé de conformères a été généré, afin de s'assurer que l'espace conformationnel des ligands est couvert
Nombre maximal de poses retenues (N_{save})	20 ou 50	La rétention des 50 ou des 20 meilleurs poses permet éventuellement d'identifier des familles d'orientations possibles
Ecart de forme maximal entre le site et le ligand (ρ)	4 Å	Cette limite représente un bon compromis d'après Venkatachalam [2] : des valeurs supérieures engendrent une rétention de poses ayant une adéquation de forme récepteur/ligand insatisfaisante ; des valeurs inférieures mènent à l'exclusion excessive de poses pouvant former des complexes énergétiquement favorables
Ecart minimal pour la discrimination des poses du ligand	1.5 Å	Cette valeur évite que des poses trop similaires soient retenues dans la liste de sauvegarde
Fonctions de score sélectionnés	LigScore1, LigScore2, PLP1, PLP2, PMF, Jain	Les six fonctions de score disponibles dans Discovery Studio 2.5 ont été sélectionnées
Algorithme de minimisation	Steepest Descent, 1000 étapes	Cette étape supplémentaire permet de minimiser les poses des ligands ainsi que les atomes du récepteur englobés dans la sphère précédemment définie

4.4. Classement des complexes issus du docking

Le classement des complexes issus de la procédure de docking a été effectué d'une part à l'aide d'un consensus des fonctions de score implémenté dans Discovery Studio 2.5, et d'autre part, en regroupant les poses selon leur similarité conformationnelle.

- *Consensus des fonctions de score* : les poses retenues dans la procédure de docking ont été soumises à un consensus de fonctions de score. La méthode implémentée dans Discovery Studio est le CScore qui est une technique basée sur le *coarse quantile*. Chaque fonction de score émet un vote si le résultat se situe dans le quantile supérieur de la gamme de valeurs obtenues pour cette fonction de score. Le consensus de score est le nombre total de votes obtenus pour chaque pose. Ce filtre énergétique permet d'identifier les conformations présentant les affinités les plus élevées pour le récepteur.

Les fonctions utilisées pour ce consensus sont classées selon les approches qui ont été décrites dans la section 1.2.2. du chapitre 1 et elles sont de deux types (Tableau II.6). Les fonctions de score empiriques qui consistent en un ensemble de termes caractérisant divers aspects des interactions protéine-ligand et la fonction de score basée sur la connaissance qui est issue d'une collecte statistique de distances interatomiques dans une base de structures protéine-ligand.

Tableau II.6. Les types de fonctions de score implémentés dans *Discovery Studio 2.5*

Type de fonctions de score	Fonction de score
Empirique	PLP1 [13], PLP2 [13], LigScore1 [14], LigScore2 [14], Jain [15]
Basé sur la connaissance	PMF[16, 17]

- *Analyse géométrique des poses* : les poses obtenues à l'issue du docking ont été regroupées d'après leur similarité conformationnelle. La classification est réalisée en calculant le RMSD entre chaque paire de conformations et est représentée sous forme de dendrogramme (Figure II.10). Ce filtre géométrique permet de repérer rapidement les poses présentant des similarités mais également des familles de poses.
- La *Combinaison des filtres énergétique et géométrique* a conduit à la sélection des différents complexes. La similarité structurale entre les poses ayant obtenu un consensus de score de 6 ou 5 a été analysée comme indiqué dans la Figure II.10. Pour chaque famille conformationnelle, c'est-à-dire pour les poses qui présentent une orientation similaire dans le site actif de l'enzyme, un ou deux représentant(s) a (ont) été sélectionné(s) parmi les conformères ayant un consensus de 6 ou 5.

A l'issue du processus de sélection, les complexes sont solvatés explicitement avec des molécules d'eau. Les systèmes sont ensuite soumis au protocole de simulations de dynamique moléculaire décrit dans la section 3.3. Des trajectoires de 10 ns ont été produites pour les complexes lipase/flavonoïde.

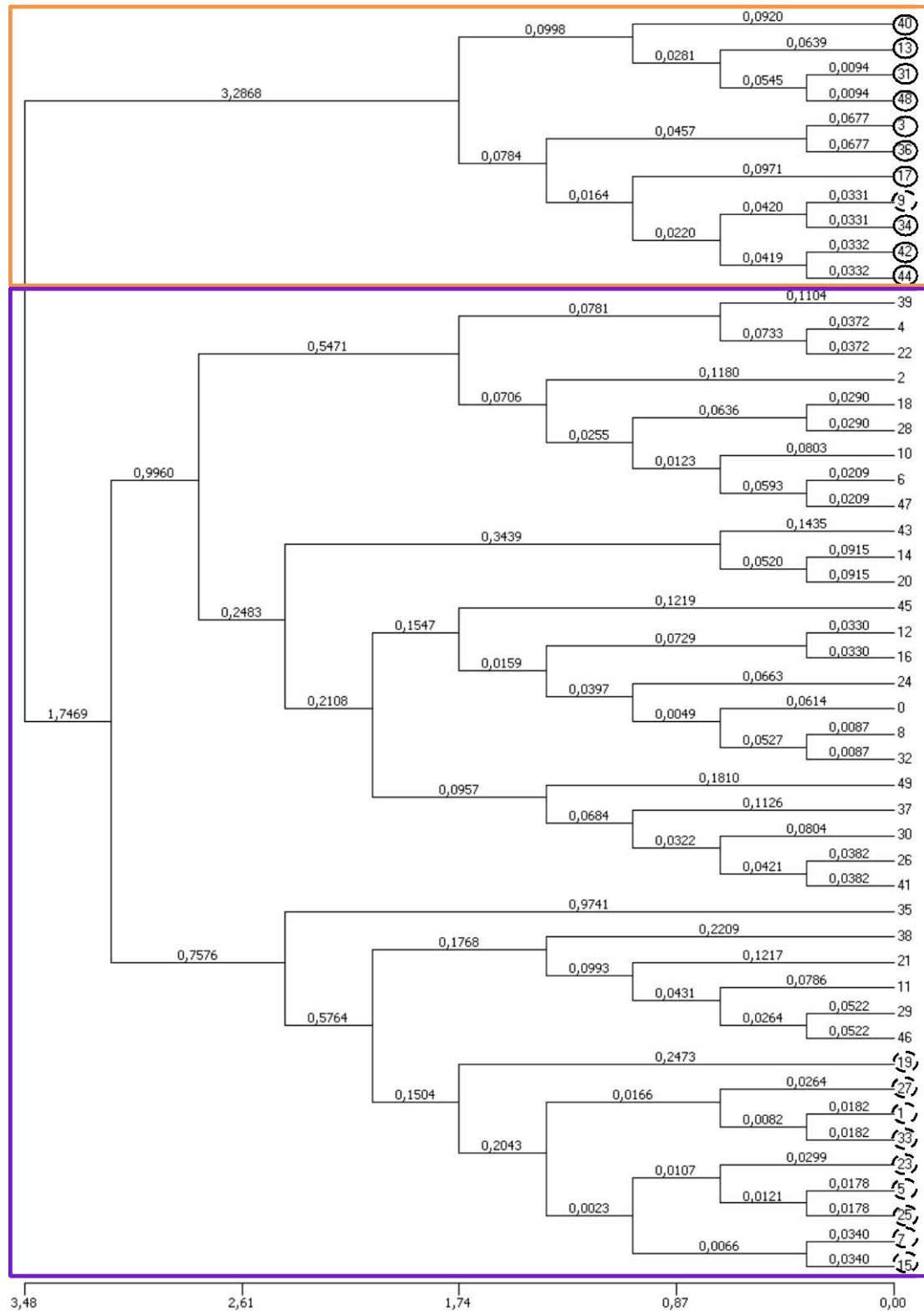


Figure II.10. Dendrogramme des poses issues du docking de la quercétine dans la CALB acétylée. Le regroupement des poses selon leur similarité structurelle permet de visualiser aisément les différentes familles (encadré orange et violet) ainsi que les sous-familles. Les poses ayant obtenu un consensus de fonctions de score de 6 et de 5 sont entourées respectivement d'un cercle en trait plein et d'un cercle en trait discontinu.

4.5. Calcul de l'énergie libre de liaison

4.5.1. Impact de la modification du substrat sur l'affinité de liaison

Les poses obtenues à l'issue du protocole de docking ont été évaluées selon leur affinité de liaison pour le site actif de la lipase par la méthode MM-PBSA (en anglais, *Molecular mechanics-Poisson Boltzmann surface area*) implémentée dans le module 'Calculate Binding Energy' de *Discovery Studio 2.5*. Le calcul d'énergie est effectué sur des structures statiques des complexes lipase/flavonoïde qui ont été préalablement minimisées (1000 étapes de descente abrupte (en anglais, *Steepest Descent*)). L'expression générale de l'énergie libre de liaison est donnée dans l'équation II.6.

$$\Delta G_{\text{liaison}} = \Delta G_{\text{complexe}} - \Delta G_{\text{protéine}} - \Delta G_{\text{ligand}} \quad (\text{II.6})$$

où l'énergie libre du complexe ($\Delta G_{\text{complexe}}$), de la protéine ($\Delta G_{\text{protéine}}$) et du ligand (ΔG_{ligand}) sont calculées d'après les équations suivantes :

$$\begin{aligned} \Delta G &= \Delta E_{\text{MM}} + \Delta G_{\text{solvation}} - T\Delta S \\ \Delta E_{\text{MM}} &= \Delta E_{\text{int}} + \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} \\ \Delta G_{\text{solvation}} &= \Delta G_{\text{polaire}} + \Delta G_{\text{nonpolaire}} \end{aligned} \quad (\text{II.7})$$

ΔE_{MM} est l'énergie de la mécanique moléculaire en phase gazeuse calculée pour chaque *frame* désolvatée. $\Delta G_{\text{solvation}}$, l'énergie libre de solvation, est divisée en deux termes, la composante électrostatique $\Delta G_{\text{polaire}}$ est calculée en utilisant l'approche de Poisson-Boltzmann et, la composante non polaire est déterminée en utilisant le modèle de l'aire de la surface accessible au solvant. L'entropie ($T\Delta S$) est estimée en utilisant l'analyse de mode normal. L'entropie peut être omise dans le cas où l'énergie libre relative est évaluée pour une série de composés structurellement similaires. Les différents paramètres sont repris dans le Tableau II.7.

Ce protocole a été utilisé pour comparer l'énergie libre de liaison:

- (i) de la quercétine pour la CALB native et la CALB acylée ;
- (ii) entre une réaction de synthèse et une réaction d'alcoololyse ;
- (iii) entre la quercétine native et différents dérivés de la quercétine pour le site actif de la CALB.

4.5.2. Impact de mutations dans CALB sur l'affinité de liaison

L'énergie libre de liaison a été déterminée pour des complexes quercétine/CALB sauvage ainsi que pour des complexes quercétine/CALB mutée dans le but d'évaluer l'impact des mutations introduites dans l'enzyme CALB sur l'affinité de liaison avec le substrat quercétine. Elle a été calculée, en utilisant l'exécutable de CHARMM implémenté dans *Discovery Studio 2.5.*, par la méthode MM-PBSA. Le calcul a été réalisé sur un ensemble conformationnel (500 frames) issu de la trajectoire de dynamique moléculaire réalisée sur les complexes (1 frame toutes les 20 ps). Les molécules d'eau ainsi que les contre-ions ont été éliminés des frames. Les différents paramètres sont donnés dans le Tableau II.7.

Tableau II.7. Paramètres pour le calcul de l'énergie libre de liaison.

Paramètre	Modification du substrat	Impact des mutations
Nombre de frame pour le calcul	1	500
Paramètre pour la contribution polaire		
Constante diélectrique du solvant	80	80
Constante diélectrique du soluté	1	1
Espacement de la grille (Å)	0.4	0.4
Rayon pour les molécules de soluté (Å)	Rayon de van der Waals	Rayon de van der Waals
Paramètre pour la contribution non polaire		
γ (kcal.mol ⁻¹ .Å ⁻²)	0.00542	0.00542
β (kcal.mol ⁻¹)	0.92	0.92
Rayon pour les molécules de solvant (Å)	1.4	1.4
Température (K)	298	300

4.6. Validation des performances de LigandFit

Il existe divers logiciels de docking et de fonctions de score. Nous avons cherché à vérifier les performances du logiciel choisi avec les systèmes étudiés. La méthode la plus simple pour évaluer la précision d'une procédure de docking est de déterminer si la pose la mieux classée à l'issue du docking est structuralement proche de la conformation du ligand natif indiquée par les données cristallographiques. Selon la littérature [18, 19], la prédiction est dite réussie si la valeur de l'écart quadratique moyen (RMSD) de la meilleure conformation est inférieure à 2 Å par rapport aux données expérimentales. La précision de LigandFit a été évaluée par le docking de l'inhibiteur éthyle hexylphosphonate (HEE) dans le site actif de la CALB. La structure cristallographique utilisée dans le cadre de cette validation est celle correspondant au code PDB 1LBS. Les résultats indiquent que la conformation obtenue pour l'inhibiteur se superpose parfaitement avec celle présente dans la structure cristallographique comme montré en Figure II.11. La valeur de RMSD des atomes lourds de la meilleure conformation du HEE est de 0.77 Å en prenant l'inhibiteur présent dans la structure 1LBS comme référence. Ces résultats mettent en évidence la bonne précision et les performances de LigandFit pour les simulations de docking.

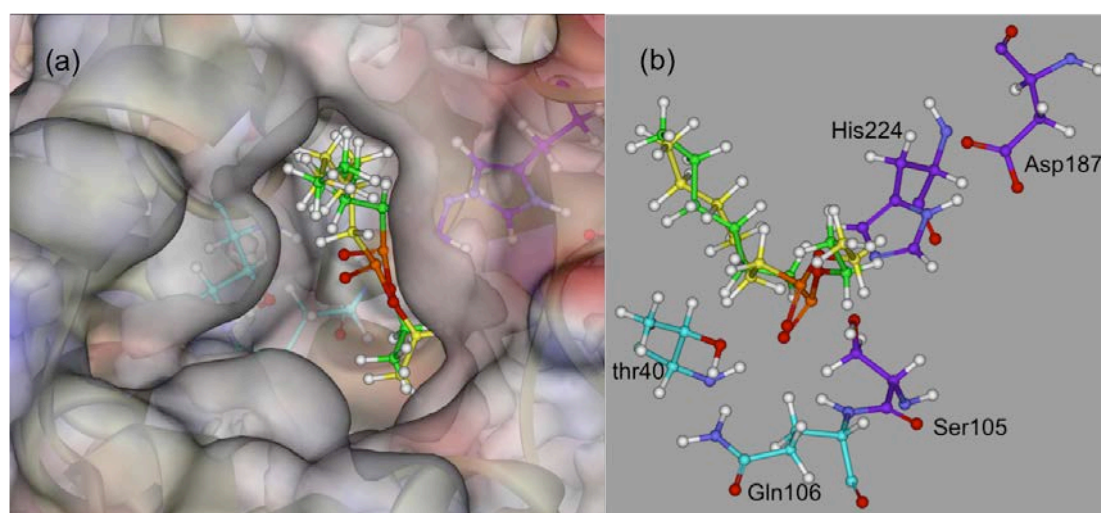


Figure II.11. Superposition de la structure cristallographique 1LBS et du meilleur complexe obtenu par docking de l'inhibiteur HEE dans la cavité de CALB (code PDB : 1LBS), en utilisant LigandFit. Le ligand

HEE docké et le ligand HEE présent dans la structure cristallographique sont colorés respectivement en jaune et vert. Le RMSD défini sur les atomes lourds des structures est de 0,77 Å.

5. Résumés schématiques des protocoles de modélisation moléculaire

Dans le cadre de ce travail, trois axes de recherche ont été développés : l'étude de la spécificité des lipases CALB et PCL, un travail d'ingénierie de la CALB pour l'acylation régiosélective de la quercétine et l'étude des interactions CALB/quercétine par modification structurale du substrat. Les protocoles généraux appliqués lors de ces études ont été schématisés pour clarifier l'ordre des événements.

5.1. Étude de la spécificité de la lipase B de *Candida antarctica* et de la lipase de *Pseudomonas cepacia* dans l'acétylation de la quercétine

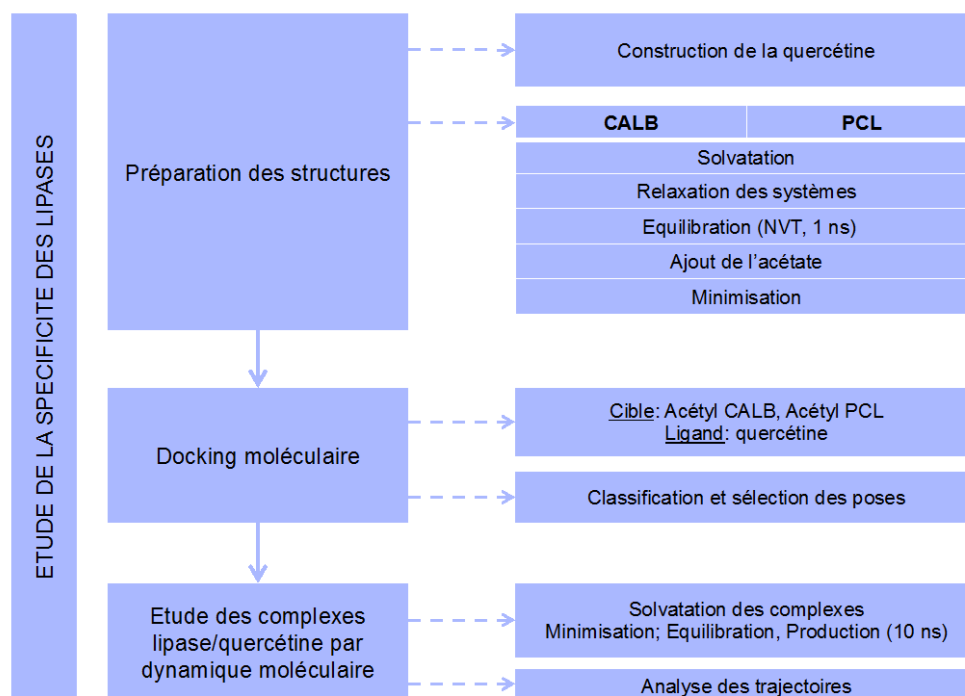


Figure II.12. Représentation schématique du protocole appliqué pour l'étude de la spécificité des lipases CALB et PCL dans la réaction d'acétylation de la quercétine.

5.2. Ingénierie de la lipase B de *Candida antarctica*

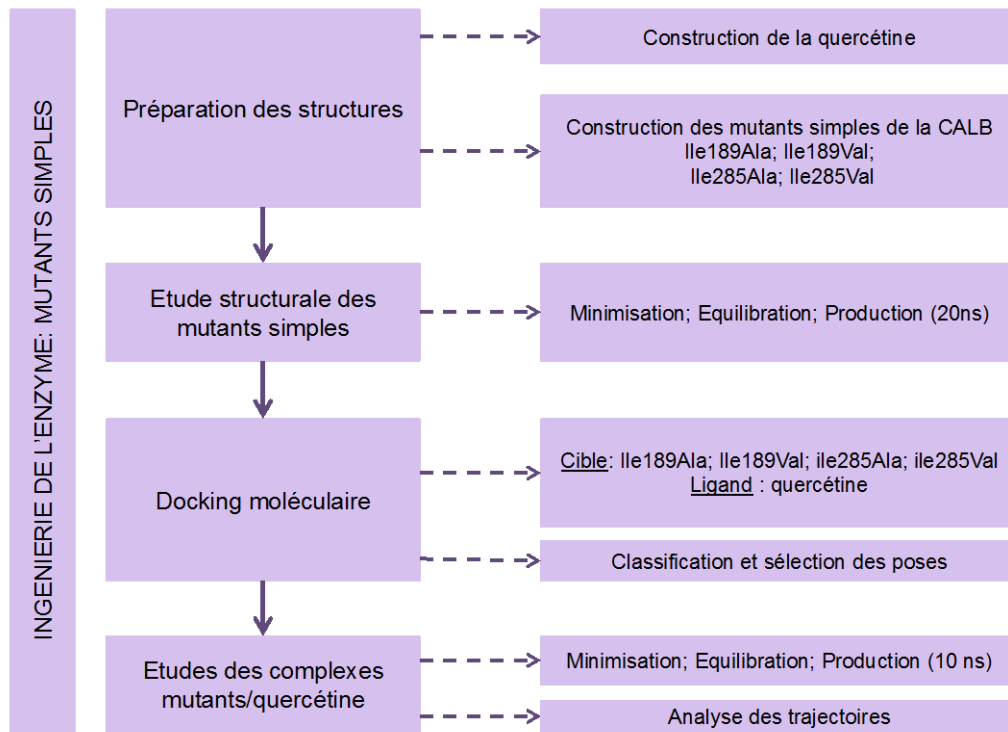


Figure II.13. Représentation schématique du protocole appliqué pour l'étude structurale des variants simples de CALB.

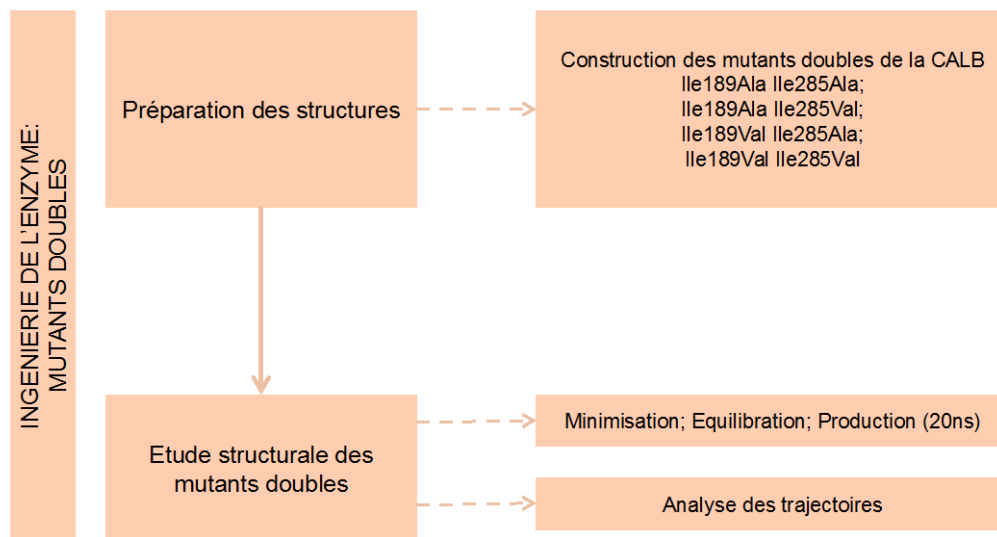


Figure II.14. Représentation schématique du protocole suivi pour l'étude structurale des variants doubles de CALB.

5.3. Étude des interactions CALB/quercétine par modification structurale du substrat

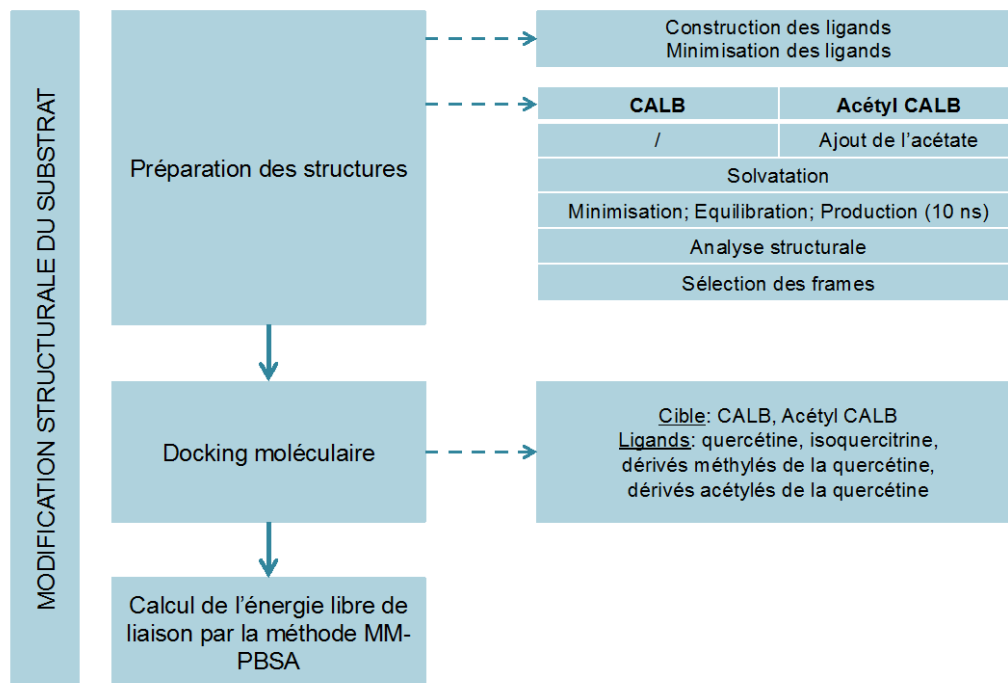


Figure II.15. Représentation schématique du protocole appliqué pour l'étude des interactions CALB/quercétine, par modification structurale du substrat.

6. Références

1. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy minimization and dynamics calculations*. Journal of Computational Chemistry, 1983. **4**: p. 187-217.
2. Venkatachalam, C.M., et al., *LigandFit : a novel method for the shape-directed rapid docking of ligands to protein active sites*. Journal of Molecular Graphics and Modelling, 2003. **21**: p. 289-307.
3. Feher, M., *Consensus scoring for protein–ligand interactions*. Drug Discovery Today, 2006. **11**: p. 421-428.
4. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models*. Accounts of Chemical Research, 2000. **33**(12): p. 889-897.
5. <http://www.ks.uiuc.edu/Research/namd>.
6. Humphrey, W., A. Dalke, and K. Schulten, *VMD : Visual Molecular Dynamics*. Journal of Molecular Graphics, 1996. **14**: p. 33-38.
7. Uppenberg, J., et al., *The sequence, crystal structure determination and refinement of two forms of lipase B from Candida antarctica*. Structure, 1994. **2**: p. 293-308.
8. Uppenberg, J., et al., *Crystallographic and molecular-modeling studies of lipase B from Candida antarctica reveal a stereospecificity pocket for secondary alcohols*. Biochemistry, 1995. **34**: p. 16838-16851.
9. Ryckaert, J.P., G. Ciccotti, and J.C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes*. Journal of Computational Physics, 1977. **23**(3): p. 327-341.
10. Bhandarkar, M., et al., *NAMD User's Guide version 2.5*. 2003.
11. Kim, K.K., et al., *The crystal structure of a triacyl glycerol lipase from Pseudomonas cepacia reveals a highly open conformation in the absence of a bound inhibitor*. Structure, 1997. **5**: p. 173-185.
12. Oldfield, T.J., *A number of real-space torsion-angle refinement techniques for proteins, nucleic acids, ligands and solvent*. Acta Crystallographica Section D, 2001. **57**(1): p. 82-94.
13. Gehlhaar, D.K., et al., *Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming*. Current Biology, 1995. **2**: p. 317-324.
14. Krammer, A., et al., *LigScore : a novel scoring function for predicting binding affinities*. Journal of Molecular Graphics and Modelling, 2005. **23**: p. 395-407.
15. Jain, A.N., *Scoring non covalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities*. Journal of Computer-Aided Molecular Design, 1996. **10**(5): p. 427-440.
16. Muegge, I., *PMF revisited*. Journal of Medicinal Chemistry, 2006. **49**: p. 5895-5902.
17. Muegge, I. and Y.C. Martin, *A general and fast scoring function for protein-ligand interactions : a simplified potential approach*. Journal of Medicinal Chemistry, 1999. **42**: p. 791-804.
18. Wang, R., Y. Lu, and S. Wang, *Comparative evaluation of 11 scoring functions for molecular docking*. Journal of Medicinal Chemistry, 2003. **46**: p. 2287-2303.
19. Kramer, B., M. Rarey, and T. Lengauer, *Evaluation of the FlexX incremental construction algorithm for protein-ligand docking*. PROTEINS: Structure, Function and Bioinformatics, 1999. **37**: p. 228-241.

Chapitre III. Étude Expérimentale de
l'Acétylation de la Quercétine en Présence
de la Lipase B de *Candida antarctica*

1. Introduction

Plusieurs études expérimentales et théoriques publiées ces dernières années, aussi bien par notre équipe que par d'autres laboratoires, ont démontré l'incapacité de la lipase B de *Candida antarctica* (CALB) à catalyser l'acylation du flavonoïde aglycone quercétine, en milieu organique. Le bioprocédé d'acylation enzymatique a été testé avec des esters de vinyle comme donneur d'acyle, un rapport molaire donneur d'acyle-flavonoïde compris entre 5 et 40, différents solvants organiques (acétone, acétonitrile et le 2-méthyl-2-butanol) ainsi qu'une température de 50 ou 60°C. Par ailleurs, une équipe grecque a identifié les dérivés 3'-O-acétyl et 4'-O-acétyl de la quercétine en appliquant des conditions expérimentales similaires exception faite du rapport molaire donneur d'acyle-flavonoïde qu'ils ont fixé à 330. Ces résultats contradictoires sur l'activité d'acétylation de la CALB vis-à-vis de la quercétine ont été clarifiés dans ce chapitre.

L'expérience d'acétylation de la quercétine a été menée en présence et en absence de la CALB dans le milieu réactionnel. Le solvant organique acétonitrile, l'acétate de vinyle comme donneur d'acyle, un rapport molaire vinyle acétate/quercétine de 330 et une température de 60°C sont les conditions expérimentales qui ont été utilisées. Les milieux réactionnels ont été caractérisés par spectrométrie de masse (LC-ESI-MS).

Les résultats de cette étude sont reportés sous la forme d'un projet de publication

2. Article

Regioselective chemo-enzymatic acetylation of flavonoid quercetin with *Candida antarctica* lipase B

Bidouil Christelle^a, Paris Cédric^a, Humeau Catherine^a, Chebil Latifa^a, Ghoul Mohamed^a & Engasser Jean-Marc^{a*}

^a Laboratoire Ingénierie des Biomolécules, ENSAIA-INPL, Université Lorraine, 2 av. de la Forêt de Haye, 54500, Vandoeuvre-lès-Nancy, France

*Corresponding authors: Phone: (+33)3.83.59.58.41 / Fax: (+33)3.83.59.57.78
Email: jean-marc.engasser@ensaia.inpl-nancy.fr

Abstract:

In order to assess the respective contributions of chemical and enzymatic acetylation on the aglycone flavonoid quercetin when reacting under a high excess of vinyl acetate, the quercetin acetylation reaction was compared in the presence and absence of the *Candida antarctica* lipase B (CALB). In acetonitrile at 60°C, when adding an initial high molar ratio of vinyl acetate/quercetin of 330, the acetylation takes place in the absence of the enzyme catalyst. After three days about 90% of quercetin is converted mainly into quercetin monoacetate. When the immobilized CALB lipase is added, only a slightly reduced acetylation of quercetin is observed. This confirms previous reports that CALB has no acetylation activity on the aglycone quercetin but may display a hydrolytic activity on chemically synthesized quercetin acetates.

Keywords: quercetin, chemical acylation, chemo-enzymatic acylation, *Candida antarctica* lipase B

1. INTRODUCTION

Quercetin [2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-chromen-4-one, 1] is one of the most abundant flavonoids in fruits and vegetables. These polyphenolic antioxidants are drawing increasing interest in view of their potentially beneficial effects (anti-inflammatory, anti-ageing, cardioprotective, anticancer, etc.) [1-4]. Quercetin has a characteristic fifteen-carbon backbone structure consisting of two aromatic rings (A and B) joined by a three-carbon linked γ -pyrone ring (C) (Figure III.1). It carries five hydroxyl groups with different properties and biological activities. The catechol moiety on the B ring (3'-OH, 4'-OH) is largely responsible for the redox properties of the molecule [5], the 3-OH is a key group for kinase inhibition [6], the 7-OH is chiefly responsible for the weak uncoupling activity [7]. The 5-OH is the least acidic and reactive one, due to intramolecular H-bonding to the carbonyl at 4-C [8].

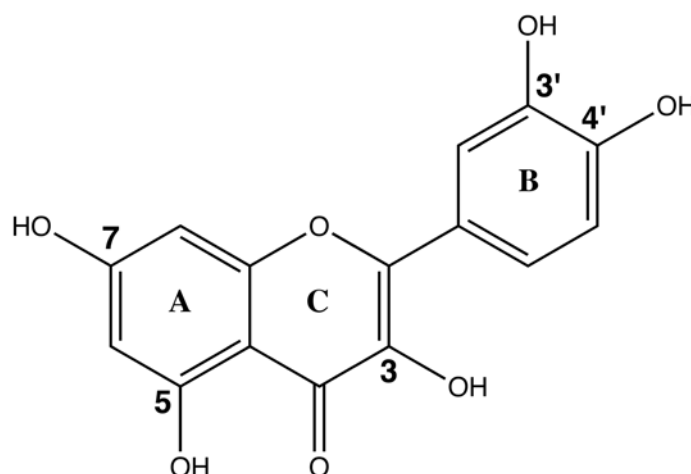


Figure III.1. Conventional nomenclature of quercetin structure

Despite promising activities, quercetin has not been used widely because it is practically insoluble in water or in oil [9] and displays low oral absorption and bioavailability [2]. Esterification of some of the hydroxyl groups is an interesting approach to overcome these limitations. Quercetin esters have been reported to withstand liver conjugation reactions when absorbed and to be reconverted into quercetin through blood esterases [10]. In addition acylation may yield quercetin analogues with higher biological activities. For instance quercetin-3-monoacetate was reported to show enhanced anti-inflammatory activity compared to quercetin [11].

In order to synthesize quercetin esters with selected numbers and positions of acylated hydroxyl groups, a chemical or enzymatic toolbox for regioselective quercetin modifications is necessary. In the past, quercetin acylation has been the subject of many studies using chemical reactions, enzymatic bioconversion, or combinations of chemical and enzymatic conversions. A wide variety of esters have been obtained depending on the nature of the catalyst, the reaction solvent, the acylating agent, the initial concentrations and the medium temperature (Figure III.2).

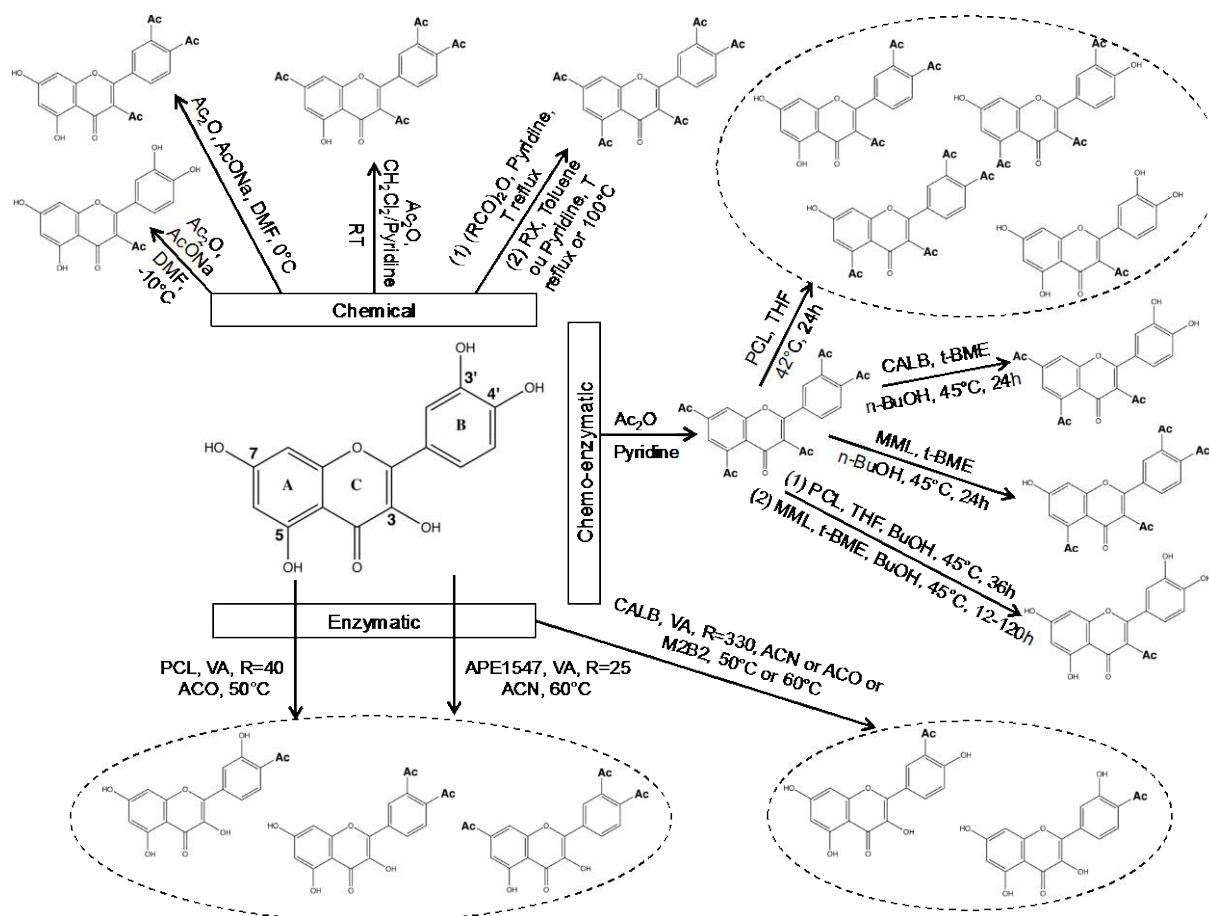


Figure III.2. Regioselectivity of quercetin acylation by chemical, enzymatic and chemo-enzymatic approach used (from data in [10-23]). ACN: acetonitrile; ACO: acetone; Ac₂O: Acetic Anhydride; AcONa: sodium acetate; APE1547: *Aeropyrum pernix* K1 esterase; BuOH: butanol; CALB: *Candida antarctica* lipase B; CH₂Cl₂: dichloromethane; DMF: dimethylformamide; M2B2: 2-methyl-2-butanol; MML: *Mucor meihei* lipase; n-BuOH: n-butanol; PCL: *Pseudomonas cepacia* lipase; R: molar ratio; (RCO)₂O: acid anhydride; Room T: Room temperature; RX: alkyl halide; T: temperature; t-BME: methyl tert-butyl ether; THF: tetrahydrofuran; VA: vinyl acetate.

For the chemical conversion of quercetin, mostly described is the non regioselective synthesis of pentaacetylquercetin using acetate anhydride, or alkyl halide in pyridine or toluene solvent at reflux temperature or 100°C. By careful control of the reaction conditions (temperature, type and equivalents of acylating agent and reaction time), it is possible to stop the acylation at the 3,3',4',7- tetraester stage [21]. The regioselectivity can be attributed to the low nucleophilic reactivity of the 5-OH group. A tetraacetyl quercetin with a free 7-OH was also prepared via imidazole-promoted hydrolysis of pentaester [21]. Synthesis of quercetin triacetate at position 3', 4', 3, was achieved by reacting quercetin dissolved in dimethylformamide with sodium acetate and acetic anhydride at 0°C [11]. On the other hand, the monoester quercetin-3-monoacetate was synthesized when quercetin was acetylated at -10°C with acetic anhydride in dimethylformamide and sodium acetate as catalyst [20]. A still broader range of quercetin esters can be obtained when reacting the chemically synthesized quercetin pentaacetate with lipase enzymes, the regioselectivity of the transesterification reaction being dependent on the lipase structure and the reaction conditions. For instance reacting pentaacetate with *Pseudomonas cepacia* lipase (PCL) in tetrahydrofuran yields a mixture of tetraester

3', 4', 3, 5, triesters 3', 4', 3 and 3', 3, 5 and monoester 3 [15]. Quercetin pentaacetate has also been transesterified through alcoholysis in butanol (BuOH). With *Mucor Meihei* lipase (MML), in t-BME and BuOH the reaction yielded tetraester 3', 4', 3, 5 [16, 17]. With *Candida antarctica* lipase B (CALB) in t-BME and n-BuOH, mainly triester 3, 5, 7 was obtained, which indicates a preferential CALB transesterification activity on the B ring 3' and 4' acetate [16, 17]. Under similar conditions the PCL lipase catalysed the transesterification on four acetates groups to produce quercetin-mono-3-acetate [18, 24, 25]. Lipases have also been tested to directly acylate quercetin in organic solvents. Acetylation activity was reported for the lipase PCL [13] and the esterase *Aeropyrum pernix* K1 (APE1547) [12] using vinyl acetate at a VA/Q ratio of 25 and 40 in acetone and acetonitrile, respectively. A mixture of quercetin monoacetate 4', quercetin diacetate 4' and 3' and quercetin triacetate at 4' 3' and 7 was then synthesized. Interestingly, under similar conditions, no acetylating activity was reported with *Candida antarctica* lipase B, whereas it was found very active in acylating the sugar hydroxyl groups on the glycosylated quercetin derivatives rutin and isoquercitrin.

In a recent paper Kyriakou et al [14] reported an acylation activity of CALB on quercetin when performing the reaction at a higher acetate/quercetin molar ratio. In acetonitrile at 60°C, after 72 hours the reaction yielded a conversion of quercetin into 3' and 4' monoacetate. As no blank experiment was done the absence of the lipase, the possibility for a chemical acetylation cannot be excluded. Consequently, no clear conclusion can be drawn on the ability of CALB to acetylate aglycone quercetin. In order to further assess the respective contributions of chemical and enzymatic acetylation under these operational conditions, and thus to clarify the esterification activity of CALB on quercetin, in the present work the acetylation reaction of quercetin was investigated under high excess of vinyl acetate in the presence and absence of the CALB lipase.

2. MATERIALS AND METHODS

2.1. Reagents

Immobilized CALB (Novozym 435, acrylic resin, EC 3.1.1.3, 7000 PLU/g : propyl laurate units synthesized per gram of catalyst) was purchased from Novo Industries. Quercetin dehydrate (Sigma) was the flavonoid aglycon tested and vinyl acetate (purity 99%, Aldrich) was used as acyl donors. The acetylation reactions were performed in acetonitrile.

2.2. Flavonoid ester synthesis procedure

The synthesis of flavonoid esters was performed by dissolving 500 mg (150 µmol) of flavonoid in 10 ml acetonitrile which has been previously dried on 4 Å molecular sieves. After 15 h of solubilisation at 60 °C, 5 ml of acyl donor (50 mmol) was added to the mixture. The enzymatic reaction was started by adding 60 g/L of Novozym 435. Reaction was carried out in stirred Schott bottle that was incubated in oil bath regulated at 60 °C. Two controls experiments (quercetin and quercetin/acyl donor) without enzymes were performed in the same condition as the synthesis reaction. The reactions lasted for 72h (3 days).

Samples taken from the reaction medium after acyl donor addition and after 72 hours reaction were diluted in the same solvent used for biotransformation and analysed by LC-MS.

2.3. HPLC - MS analysis

Screening of quercetin and acetylated forms of quercetin was performed using a HPLC-MS system (ThermoFisher Scientific, San Jose, CA, USA) equipped with an LTQ ion trap as mass analyzer (Linear Trap Quadrupole). Data were processed using Xcalibur software (version 2.1). Chromatographic separation was performed on a C18 reverse phase column (150*2.1mm, 5µm porosity – Grace/Alltech, Darmstadt, Germany) equipped with a C18 pre-column (7.5*2.1mm, 5µm porosity – Grace/Alltech, Darmstadt, Germany) at 25°C and mobile phases consisted in water modified with acetic acid (0.1%) for A and acetonitrile for B. Compounds were eluted using a linear gradient from 10% to 100% of B phase for 25 min at a flow rate of 0.2 mLmin⁻¹. A double photodiode array and mass detection was performed during all the time of the run. The mass spectrometric conditions were as follows: ESI positive ionization mode (ElectroSpray Ionization) was used; source gases were set (in arbitrary units min⁻¹) for sheath gas, auxiliary gas and sweep gas at 40, 10 and 10, respectively; capillary temperature was set at 300°C; capillary voltage was set at 48V; tube lens, split lens and front lens voltages were set at 138 V, -38V and -4.25V, respectively. The ion optics parameters were optimized by automatic tuning using a standard solution of rutin at 0.1gL⁻¹ infused in mobile phase (A/B: 50/50) at a flow rate of 5µLmin⁻¹. Full scan MS spectra were repeatedly performed between 100 to 1000 m/z ratio during the time of the run and compounds of interest were monitored as their pseudo-molecular [M+H]⁺ corresponding ions. In addition MS² fragmentation spectra were realized for quercetin and the five possible acetylated forms in order to obtain structural confirmation: MS²(303.5) for quercetin, MS²(345.5) for mono-acetylated quercetin, MS²(387.5) for di-acetylated quercetin, MS²(429.5) for tri-acetylated quercetin, MS²(471.5) for tetra-acetylated quercetin, MS²(513.5) for penta-acetylated quercetin, respectively.

3. RESULTS

3.1. Acetylation of quercetin in the absence of lipase

In a first step, the acylation of quercetin by vinyl acetate in acetonitrile as a solvent was investigated in the absence of the enzyme. As a preliminary step we checked the stability of quercetin in acetonitrile at 60°C. No thermal degradation of quercetin was detected.

When a high excess of vinyl acetate was added (VA/Q=330) after three days of reaction the HPLC analysis of the diluted sample shows the formation of a major peak identified by mass spectrometry as monoacylated quercetin and a second minor peak identified as quercetin diacetate (Figure III.3). HPLC analysis of undiluted samples also revealed traces of tri- and tetra- acetylated products. After 72h a high 90% conversion yield of quercetin was obtained.

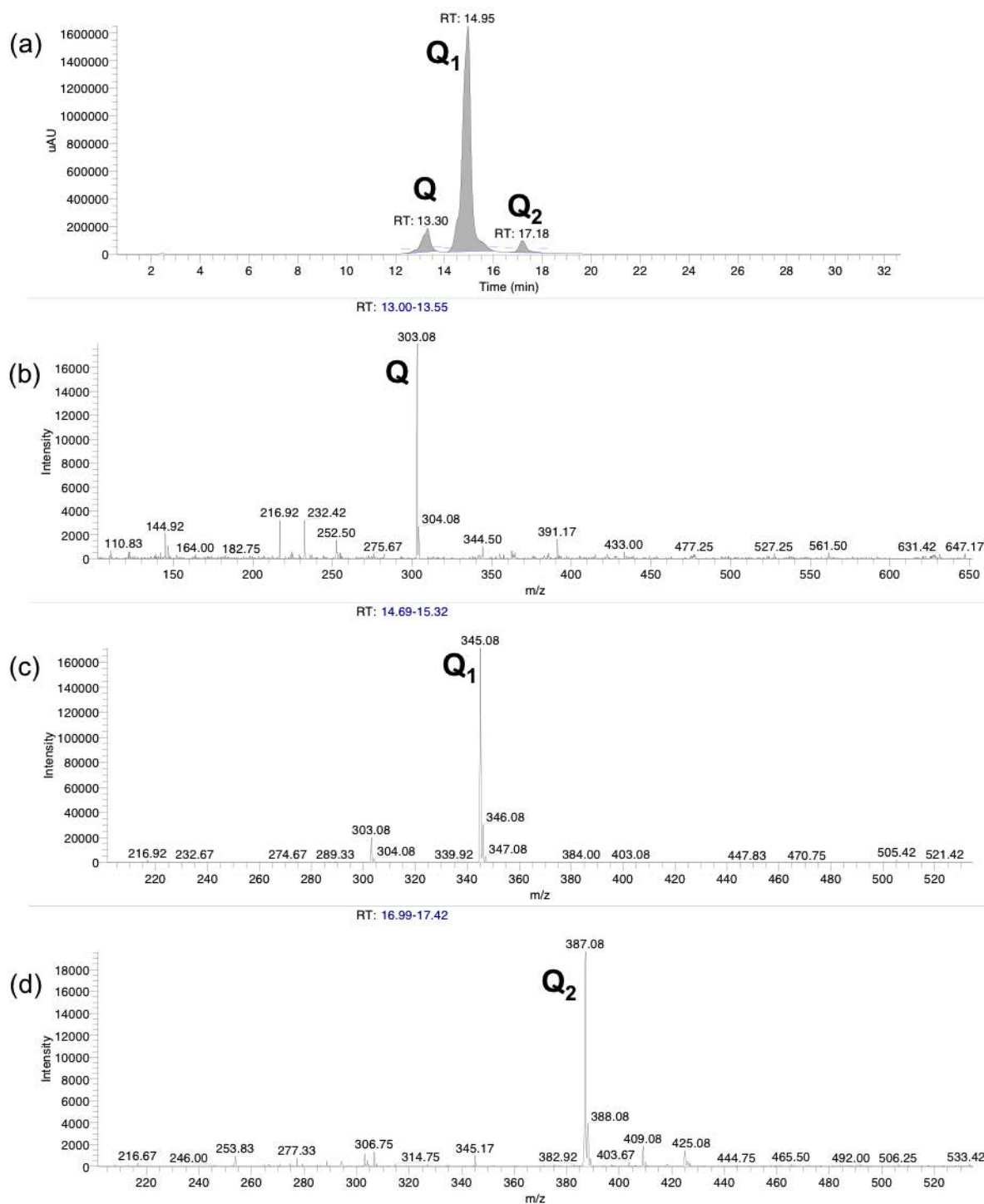


Figure III.3. MS spectrum (b, c, d) corresponding to UV chromatogram at 400 nm (a) after 72 hours reaction for products quercetin in the absence of lipase CALB.

3.2. Acetylation of quercetin in the presence of CALB lipase

In a second step, the quercetin acetylation reaction was investigated under the same conditions but by initially adding immobilized CALB lipase to the medium. After three days of reaction, HPLC analysis showed the same three peaks of unreacted quercetin, monoacetylated and diacetylated quercetin (Figure III.4).

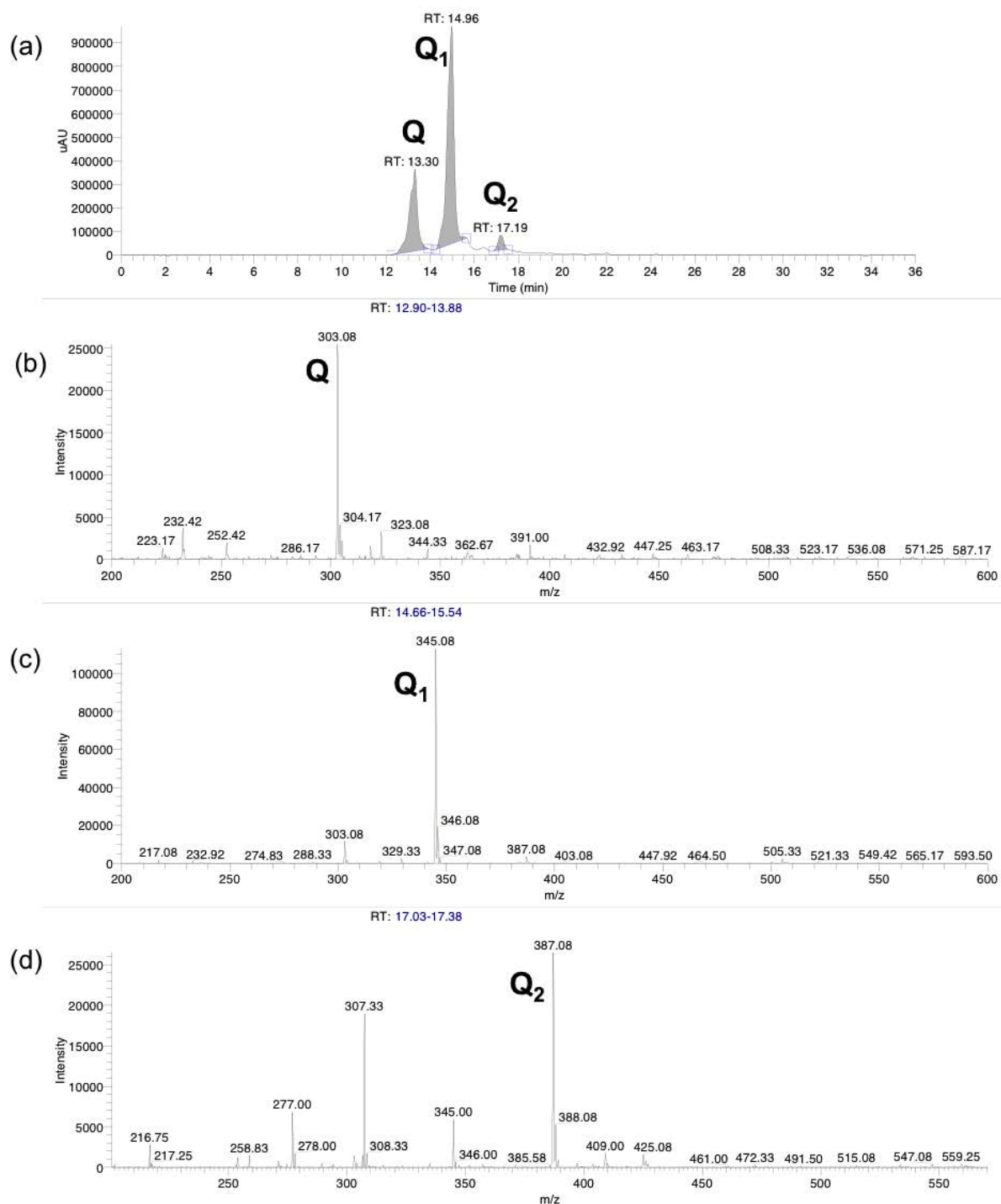


Figure III.4. MS spectrum (b, c and d) corresponding to UV chromatogram at 400nm (a) after 72 hours reaction for products quercetin in the presence of CALB.

Comparing to the previous results in the absence of lipase, only small differences can be seen : a slight decrease in the surface area of the mono- and di-acetylated quercetin, together with an increased residual quercetin. In addition traces of tri- and tetra-acetate were no longer detected in the presence of lipase (Figure III.5).

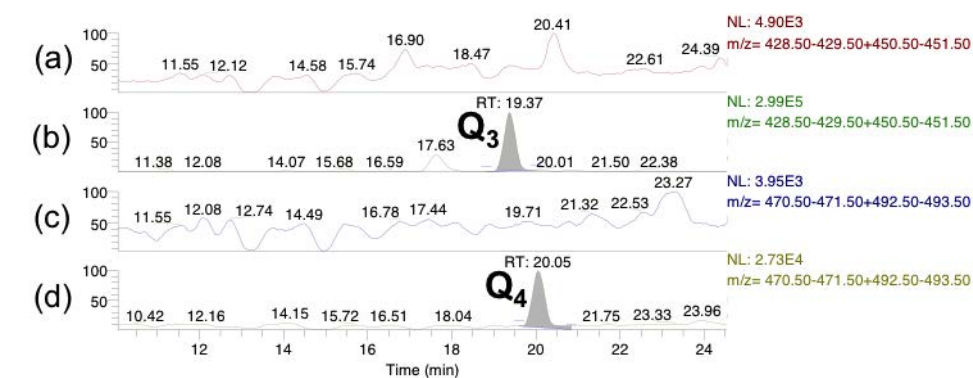


Figure III.5. Chromatogram corresponding to the masses of triacetate quercetin (a, b) $m/z [M+H]^+ = 429$ and $m/z [M+Na]^+ = 451$, and of tetraacetate quercetin (c, d) $m/z [M+H]^+ = 471$ and $m/z [M+Na]^+ = 493$ in the presence of CALB (a, c) and in the absence of CALB (b, d) for samples undiluted.

4. DISCUSSION

According to the obtained results, when adding a high excess of vinyl acetate to quercetin in acetonitrile at 60°C, the acetylation reaction takes place in the absence of enzyme catalyst. The chemical reaction leads to a predominant synthesis of quercetin monoacetate. Addition of the CALB lipase only resulted in a reduced conversion yield of quercetin.

These results are in agreement with the previously reported experimental findings of Kyriakou et al. [14] with the same medium composition and in the presence of CALB, they observed a predominant acetylation of quercetin into monoacylated quercetin. By additional NMR analysis the monoacylated products were identified as 4' and 3' quercetin monoacetate.

Comparing the results obtained with and without CALB clearly indicates that quercetin acetylation occurs by chemical reaction and that CALB has no acetylating activity on quercetin. The lipase only has slight a hydrolytic action on the quercetin acetate using traces of residual water in the acetonitrile medium. This result confirms previous reports on the absence of acylation activity of CALB on aglycone quercetin [13, 26-30].

From the reaction chemistry point of view, this study shows novel conditions for a regioselective chemical acetylation of quercetin to form predominantly quercetin monoacetate. In previous report, the synthesis of monoester quercetin-3-monoacetate was achieved when quercetin was acetylated with acetic anhydride in DMF with sodium acetate as catalyst at -10°C [20].

As a conclusion, this study brings additional contribution to the regioselective acetylation of the aglycone flavonoid quercetin by either chemical, enzymatic or chemo-enzymatic process.

5. REFERENCES

1. Harwood, M., et al., *A critical review of the data related to the safety of quercetin and lack of evidence of in vivo toxicity, including lack of genotoxic/carcinogenic properties* Food and Chemical Toxicology, 2007. **45**: p. 2179-2205.
2. Boots, A.W., G.R.M.M. Haenen, and A. Bast, *Health effects of quercetin: from antioxidant to nutraceutical*. European Journal of Pharmacology, 2008. **585**: p. 325-337.
3. Chen, C., J. Zhou, and C. Ji, *Quercetin: a potential drug to reverse multidrug resistance*. Life Sciences, 2010. **87**(11-12): p. 333-338.
4. Leopoldini, M., N. Russo, and M. Toscano, *The molecular basis of working mechanism of natural polyphenolic antioxidants*. Food Chemistry, 2011. **125**(2): p. 288-306.
5. Metodiewa, D., et al., *Quercetin may act as a cytotoxic prooxidant after its metabolic activation to semiquinone and quinoidal product*. Free Radical Biology and Medicine, 1999. **26**(1-2): p. 107-116.
6. Sarno, S., et al., *Toward the rational design of protein kinase casein kinase-2 inhibitors*. Pharmacology & Therapeutics, 2002. **93**(2-3): p. 159-168.
7. van Dijk, C., A.J.M. Driessen, and K. Recourt, *The uncoupling efficiency and affinity of flavonoids for vesicles*. Biochemical Pharmacology, 2000. **60**(11): p. 1593-1600.
8. Leopoldini, M., N. Russo, and M. Toscano, *Gas and liquid phase acidity of natural antioxidants*. Journal of Agricultural and Food Chemistry, 2006. **54**: p. 3078-3085.
9. Rothwell, J.A., A.J. Day, and M.R.A. Morgan, *Experimental determination of octanol-water partition coefficients of quercetin and related flavonoids*. Journal of Agricultural and Food Chemistry, 2005. **53**: p. 4355-4360.
10. Biasutto, L., et al., *Ester-based precursors to increase the bioavailability of quercetin*. Journal of Medicinal Chemistry, 2007. **50**(2).
11. Gusdinar, T., et al., *Anti-inflammatory and antioxidant activity of quercetin-3, 3', 4'-triacetate*. Journal of Pharmacology and Toxicology, 2011. **6**(2): p. 182-188.
12. Xie, X.-N., et al., *Acylation of quercetin with a novel thermophilic esterase as biocatalyst*. Chemical Research in Chinese Universities, 2012. **28**(2): p. 225-229.
13. Chebil, L., et al., *Enzymatic acylation of flavonoids : effect of the nature of the substrate, origin of lipase and operating conditions on conversion yield and regioselectivity*. Journal of Agricultural and Food Chemistry, 2007. **53**(23): p. 9496-9502.
14. Kyriakou, E., et al., *Unexpected enzyme-catalyzed regioselective acylation of flavonoid aglycones and rapid product screening*. Organic & Biomolecular Chemistry, 2012. **10**(9): p. 1739-1742.
15. Natoli, M., G. Nicolosi, and M. Piattelli, *Regioselective alcoholysis of flavonoid acetates with lipase in an organic solvent*. The Journal of Organic Chemistry, 1992. **57**(21): p. 5776-5778.
16. Lambusta, D., et al., *Application of lipase catalysis in organic solvent for selective protection-deprotection of bioactive compounds*. Journal of Molecular Catalysis B: Enzymatic, 2003. **22**: p. 271-277.
17. D'Antona, N., et al., *Preparation of regioprotected morins by lipase-catalyzed transesterification*. Journal of Molecular Catalysis B: Enzymatic, 2008. **52-53**: p. 78-81.
18. Saija, A., et al., *'In vitro' antioxidant and photoprotective properties and interaction with model membranes of three new quercetin esters*. European Journal of Pharmaceutics and Biopharmaceutics 2003. **56**(2): p. 167-174.
19. Gatto, M.T., et al., *Antimicrobial and anti-lipase activity of quercetin and its C2-C16 3-O-acyl-esters*. Bioorganic and Medicinal Chemistry, 2002. **10**(2): p. 269-272.

20. Herowati, R., et al., *Anti-inflammatory activity of quercetin-3-monoacetate, selective acetylation product of quercetin*. *Artocarpus*, 2008. **8**(2): p. 60-67.
21. Mattarei, A., et al., *Regioselective O-derivatization of quercetin via ester intermediates. An improved synthesis of rhamnetin and development of a new mitochondriotropic derivative*. *Molecules*, 2010. **15**: p. 4722-4736.
22. Picq, M., et al., *Pentastituted quercetin analogs as selective inhibitors of particulate 3',5'-cyclic-AMP phosphodiesterase from rat brain*. *Journal of Medicinal Chemistry*, 1982. **25**(10): p. 1192-1198.
23. Bresson Rival, D., et al., *Cosmetic, dermatological, pharmaceutical, dietetic or food compositions, e.g. for improving skin condition, comprise flavonoid esters*, in *European Patent Office* 1999: France.
24. Montenegro, L., et al., *In vitro evaluation of quercetin-3-O-acyl esters as topical prodrugs*. *International Journal of Pharmaceutics*, 2007. **336**: p. 257-262.
25. Lambusta, D., et al., *Enzyme-mediated regioprotection-deprotection of hydroxyl groups in (+)-catechin*. *SYNTHESIS*, 1993(11): p. 1155-1158.
26. Ardhaoui, M., et al., *Acylation of natural flavonoids using lipase of *Candida antarctica* as biocatalyst*. *Journal of Molecular Catalysis B: Enzymatic*, 2004. **29**: p. 63-67.
27. Stevenson, D.E., et al., *Direct acylation of flavonoid glycosides with phenolic acids catalyzed by *Candida antarctica* lipase B (Novozyme 435)*. *Enzyme and Microbial Technology*, 2006. **39**: p. 1236-1241.
28. De Oliveira, E.B., et al., *A molecular modelling study to rationalize the regioselectivity in acylation of flavonoid glycosides catalyzed by *Candida antarctica* lipase B*. *Journal of Molecular Catalysis B: Enzymatic*, 2009. **59**: p. 96-105.
29. De Oliveira, E.B., et al., *An approach based on density functional theory (DFT) calculations to assess the *Candida antarctica* lipase B selectivity in rutin, isoquercitrin and quercetin acetylation*. *Journal of Molecular Catalysis B: Enzymatic*, 2010. **66**: p. 325-331.
30. Bidouil, C., et al., *Combined docking and molecular dynamics simulations to enlighten the capacity of *Pseudomonas cepacia* and *Candida antarctica* lipases to catalyze quercetin acetylation*. *Journal of Biotechnology*, 2011. **156**: p. 203-210.

3. Contribution de l'article

Dans cette étude expérimentale, le potentiel catalytique de la lipase B de *Candida antarctica* dans la réaction d'acétylation de la quercétine (Q) a été éclairci. La réaction menée pendant 72h dans l'acétonitrile à 60°C en utilisant l'acétate de vinyle (VA) comme agent acylant et un rapport molaire VA/Q de 330 en présence et en absence du biocatalyseur permet de parvenir aux conclusions suivantes :

- En absence de la CALB dans le milieu réactionnel,
 - 90% de quercétine est converti après 72h de réaction.
 - la réaction d'acétylation est relativement régiosélective : le produit prédominant identifié est le monoacétate de quercétine et une faible quantité de diacétate de quercétine, ainsi que des traces de tri- et de tétra-acétate de quercétine sont également détectées.
- En présence de la CALB dans le milieu réactionnel,
 - 80% de quercétine est converti après 72h de réaction. Le pourcentage de conversion est légèrement inférieur à celui de la réaction conduite en absence du biocatalyseur.
 - le monoacétate de quercétine est le produit majoritaire dans le milieu et aucune trace de produits tri- et de tétra-acétylés de la quercétine n'est détectée.

Ces résultats expérimentaux sur l'acétylation de la quercétine permettent de conclure que:

- la réaction d'acétylation en absence de la CALB et en présence d'un large excès d'acétate de vinyle est essentiellement de nature chimique.
- la CALB n'a aucune activité catalytique d'acétylation mais une faible activité d'hydrolyse en utilisant les traces d'eau présentes dans le milieu.

Chapitre IV. Étude par Modélisation
Moléculaire de la Spécificité de la Lipase B
de *Candida antarctica* et de la Lipase de
Pseudomonas cepacia dans la Réaction
d'Acétylation de la Quercétine

1. Introduction

Plusieurs études expérimentales, dont celle présentée au chapitre précédent, ont mis en évidence l'absence d'activité de transestérification de la lipase B de *Candida antarctica* (CALB) dans la réaction d'acétylation du flavonoïde aglycone quercétine. Cette enzyme est pourtant connue pour sa capacité à catalyser, en milieu organique, l'estérification d'une grande variété de substrats incluant notamment deux analogues glycosylés de la quercétine, l'isoquercitrine et la rutine. Par ailleurs, la lipase de *Pseudomonas cepacia* (PCL) s'est montrée efficace dans l'acétylation enzymatique de la quercétine dans l'acétone à une température de 50°C et un ratio molaire acétate de vinyle/quercétine fixé à 40. En effet, dans ces conditions, l'ester quercétine 4'-mono-acétate est synthétisé lors des premières heures de la réaction, suivi du diester quercétine 3',4'-diacétate ainsi que du triester quercétine 3',4',7'-triacétate.

Au vu de ces résultats empiriques, il apparaît que la reconnaissance du substrat par les sites actifs des lipases CALB et PCL présente des différences significatives au niveau moléculaire. Plus particulièrement, on peut supposer que la différence de structure entre ces deux lipases a un impact important sur leur spécificité envers la quercétine. Partant de ces données, une étude a été menée avec les objectifs suivants :

- déterminer les types d'interactions de la quercétine avec le site catalytique de la lipase B de *Candida antarctica* et de la lipase de *Pseudomonas cepacia*.
- identifier pour chaque enzyme les résidus clés interagissant avec la quercétine.
- évaluer la capacité des complexes quercétine/lipase à conduire à la formation d'un produit.
- confronter les conclusions issues des modèles aux résultats expérimentaux.

Les modèles ont été construits de façon à reproduire l'acyl-enzyme. Dans le mécanisme catalytique, l'acyl-enzyme subit une attaque nucléophile par le substrat accepteur d'acyle ce qui conduit à la formation d'un intermédiaire tétraédrique puis à l'ester. Un protocole de docking a été appliqué pour étudier les positions et les orientations de la quercétine dans les sites actifs des lipases CALB et PCL acétylés (obtention de poses). En combinant une classification géométrique et énergétique des poses ainsi générées, des complexes quercétine/lipase ont été sélectionnés. Ces derniers ont ensuite été soumis à des simulations de dynamique moléculaire afin d'obtenir une évaluation temporelle de leur configuration globale et d'étudier les interactions dont ils sont le siège.

Pour évaluer si les complexes sont susceptibles de conduire à la formation d'un ester, les critères définis lors d'une étude antérieure ont été utilisés (Figure IV.1): d'une part, l'orientation de l'acétate vers les résidus du trou oxyanionique évaluée par des distances et d'autre part, l'orientation et les distances des groupements OH de la quercétine par rapport aux résidus catalytiques de la lipase et leur accessibilité.

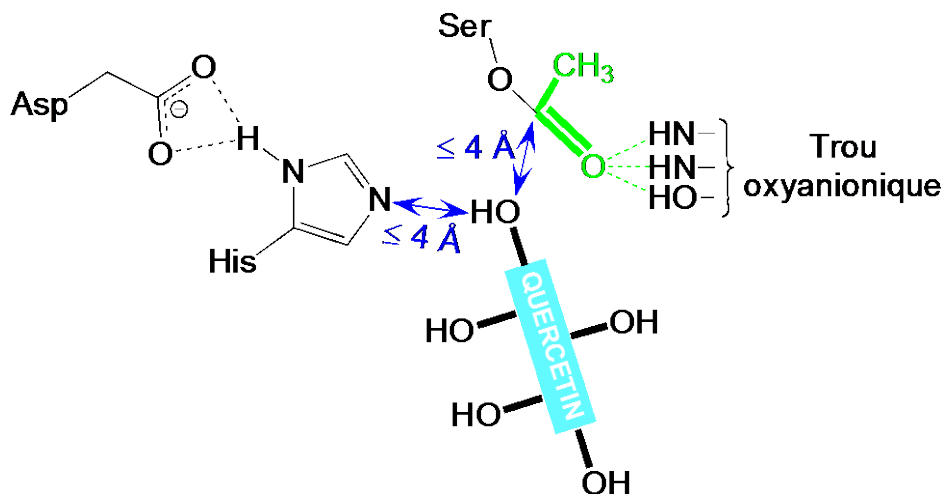


Figure IV.1. Schématisation des critères structuraux devant être respectés pour que les complexes enzyme/substrats soient considérés comme productifs. L'acétate lié à la sérine catalytique est représenté en vert. La triade catalytique et les groupements des résidus du trou oxyanionique stabilisant l'acyl-enzyme sont schématisés en noir (De Oliveira et al., 2009)

Les résultats de cette étude ont fait l'objet d'une publication intitulée « **Combined docking and molecular dynamics simulations to enlighten the capacity of *Pseudomonas cepacia* and *Candida antarctica* lipases to catalyze quercetin acetylation** » parue dans le 'Journal of Biotechnology' (volume 156, pages 203-210, 2011).

Journal of Biotechnology 156 (2011) 203–210



Contents lists available at SciVerse ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec



Combined docking and molecular dynamics simulations to enlighten the capacity of *Pseudomonas cepacia* and *Candida antarctica* lipases to catalyze quercetin acetylation

Bidouil Christelle^a, Basilio De Oliveira Eduardo^b, Chebil Latifa^a, Maia Elaine-Rose^c, Maigret Bernard^d, Ronat-Heidt Evelyne^a, Ghoul Mohamed^a, Engasser Jean-Marc^a, Humeau Catherine^{a,*}

^a Laboratoire Ingénierie des Biomolécules, ENSAIA-INPL, Nancy Université, 2 av. de la Forêt d'Haye, 54500, Vandœuvre-lès-Nancy, France

^b Departamento Tecnologia de Alimentos, Universidade Federal de Viçosa, 3 6570. 000 - Viçosa - MG - Brazil

^c Laboratório de Estudos Estruturais Moleculares (LEEM), Instituto de Química, Universidade de Brasília, CP 4478, 70904-970, Brasília-DF, Brazil

^d Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), CNRS, Université Henri Poincaré, BP 239, 54506, Vandœuvre-lès-Nancy, France

2. Article

Combined docking and molecular dynamics simulations to enlighten the capacity of *Pseudomonas cepacia* and *Candida antarctica* lipases to catalyze quercetin acetylation

Bidouil, Christelle^a, Basilio De Oliveira, Eduardo^b, Chebil, Latifa^a, Maia, Elaine-Rose^c, Maigret, Bernard^d, Ronat-Heidt, Evelyne^a, Ghoul, Mohamed^a, Engasser, Jean-Marc^a, Humeau, Catherine^{a*}

^a *Laboratoire Ingénierie des Biomolécules, ENSAIA-INPL, Nancy Université, 2 av. de la Forêt d'Haye, 54500, Vandoeuvre-lès-Nancy, France*

^b *Departamento Tecnologia de Alimentos, Universidade Federal de Viçosa, 3 6570. 000 - Viçosa - MG - Brasil*

^c *Laboratorio de Estudos Estruturais Moleculares (LEEM), Instituto de Química, Universidade de Brasília, CP 4478, 70904-970, Brasília-DF, Brazil*

^d *Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), CNRS, Université Henri Poincaré, BP 239, 54506, Vandoeuvre-lès-Nancy, France*

* Corresponding author: Phone: (+33) 3 83 59 57 84 / Fax: (+33) 3 83 59 57 78

E-mail: catherine.humeau@ensaia.inpl-nancy.fr

Abstract :

A combined docking and molecular dynamics protocol was applied to investigate quercetin binding modes within the catalytic cavity of *Candida antarctica* lipase B (CALB) and *Pseudomonas cepacia* lipase (PCL), aiming to explain the difference of specificity of these enzymes in acetylation reaction. For both lipases, docking of quercetin yielded two families of conformers with either the quercetin A or B-ring pointing towards the catalytic residues. Molecular dynamics (MD) calculations were subsequently performed on several complexes of each family. MD trajectories were analyzed focusing on the orientation of the acyl donor bound to the catalytic serine towards the oxyanion hole residues and the proximity of quercetin hydroxyl groups to the catalytic residues. Results showed that with CALB, the acetate was not correctly positioned within the oxyanion hole whatever the orientation of quercetin, suggesting that no product could be obtained. With PCL, the acetate remained within the oxyanion hole during all MD trajectories. Depending on quercetin orientation, either the 7-OH group or the 3, 5, 3', 4'-OH groups came alternatively near the catalytic residues, suggesting that all of them could be acylated. The capacity of models to explain the regioselectivity of the reaction was discussed. Key residues and interactions involved in quercetin binding modes were identified and related to the reaction feasibility.

Keywords : Lipase specificity, flavonoid, acylation, docking, molecular dynamics

1. INTRODUCTION

Quercetin (3, 5, 7, 3', 4'-pentahydroxyflavone) is a typical flavonol-type flavonoid which presents a common flavan backbone composed of two aromatic rings (A and B) joined by a three-carbon linked γ -pyrone ring (C) (Figure IV.2). Quercetin is one of the most abundant flavonoid in human diet and is known for its various biological beneficial health effects and especially its antioxidant activity (Boots et al., 2008; Chen et al., 2010; De Oliveira et al., 2009). Three structural elements were identified as determinant factors for this property: (i) a catechol moiety (B-ring), (ii) the 3- and 5-OH groups neighbouring the 4-keto group, and (iii) the 2,3-double bond (Martins et al., 2004; Musialik et al., 2009). However, the use of quercetin in several application fields is limited by its low solubility in both hydrophilic and hydrophobic media and low stability and bioavailability (Boots et al., 2008; Montenegro et al., 2007; Murakami et al., 2008). One solution consists in acylating quercetin, by grafting one or several acyl groups on the polyhydroxylated structure of the flavonoid. To preserve the functionality of the hydroxyl groups responsible for the antioxidant activity the regioselectivity of the acylation process must be controlled. In fact, Gusdinar et al. (Gusdinar et al., 2011) showed that the antioxidant activity of quercetin-3,3',4'-triacetate was lower than that of quercetin. On the contrary, Saija et al. (Saija et al., 2003) reported that the acylation of quercetin 3-OH group allowed to preserve quercetin antioxidant capacity and to enhance its photoprotective activity.

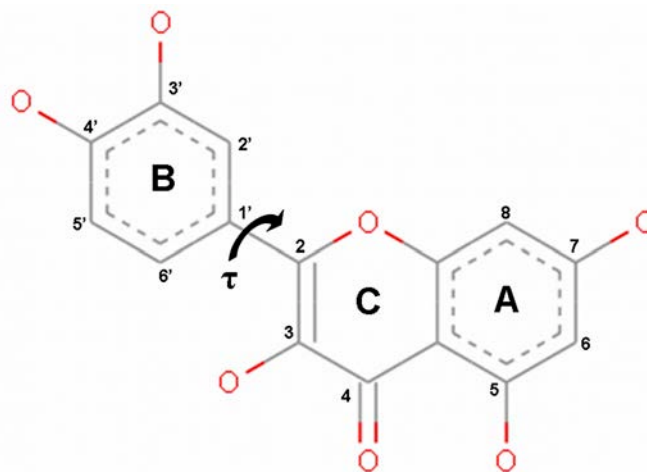


Figure IV.2. Chemical structure of quercetin, showing the systematic numbering of the carbon atoms, the nomenclature of the rings and the dihedral angle τ , defined by the atoms O1-C2-C1'-C2'.

The regioselective acylation of polyfunctional compound can be achieved using an enzymatic process catalyzed by lipases in organic medium. Among available biocatalysts, lipase B from *Candida antarctica* (CALB) is subject of numerous studies and is widely used in many industrial applications (Anderson et al., 1998; Gotor-Fernandez et al., 2006). Its main advantages are a high stability especially under immobilized forms, a broad range of substrates and interesting selectivity properties, both in hydrolysis and synthesis reactions. Many studies reported the use of this enzyme to catalyze the acylation of glycosylated flavonoids with a high degree of regioselectivity (Chebil et al., 2006;

Danieli et al., 1997; Teng et al., 2005; Xanthakis et al., 2010). On the contrary, CALB was shown to be unable to catalyze the acylation of aglycon flavonoids like quercetin (Chebil et al., 2007). To our knowledge, *Pseudomonas cepacia* lipase (PCL) was the only enzyme that was able to achieve this reaction, but non regioselectively.

Recently, a first study was reported showing the potential of molecular modelling tools to explain the regioselectivity of acetylation reaction in the case of two flavonoid glycosides catalyzed by CALB (De Oliveira et al., 2009). Computational results were shown to be in perfect accordance with experimental ones.

The present work aims to study quercetin acetylation catalyzed by CALB at molecular level. A procedure combining docking and molecular dynamics simulations was applied to investigate the binding modes of quercetin within CALB acyl enzyme. Positions, orientations and interactions of substrates in the catalytic pocket were examined and the proximity of quercetin hydroxyl groups to the catalytic residues was analyzed. Key residues involved in quercetin binding modes were identified and related to the absence of reaction. As a comparison, PCL was chosen as a model enzyme that catalyze quercetin acetylation. This work constitutes the first step for the design of pertinent mutants of CALB for the regioselective acylation of quercetin. If achieved, this goal would be a substantial step forward for the biotechnological exploitation of aglycon flavonoids.

2. MODELLING METHODOLOGY

2.1. Computational resources

Relaxation procedure of the systems and docking were carried out on a bi-processor AMD Dual Core 280, 2.4 GHz. Docking was performed using the LigandFit module (Venkatachalam et al., 2003) of the program-package Discovery Studio version 2.0 (Accelrys, Inc.). Molecular dynamics simulations were carried out on a computer cluster equipped with 2 Quad Core Intel Xeon processors L5420, 2.5 GHz and 16 GB of RAM on a Linux Platform, using NAMD (v.2.7) (Phillips et al., 2005). Visualisations and analysis of trajectories were performed with VMD (v.1.8.7) (Humphrey et al., 1996). All molecular mechanics and molecular dynamics calculations were performed with the CHARMM force field (MacKerrel Jr. et al., 1998). Partial charges of quercetin were automatically assigned by Insight II (version 2005 Accelrys, Inc) using the general CHARMM force field (Brooks et al., 1983). These non-standard force field parameters were implemented in Discovery Studio topology file (version 2.0, Accelrys, Inc.).

2.2. Preparation of the systems

2.2.1. Starting structures

The crystal structures of both PCL (Kim et al., 1997) and CALB (Uppenberg et al., 1994) were downloaded from the Protein Data Bank (PDB entries 3LIP and 1LBS, respectively). In both cases, missing hydrogen atoms were added in their theoretical positions. Protonation states of residues were assumed to be those observed at pH 7.0. Histidines were mono-protonated, aspartic and glutamic acids were negatively charged, and arginines and lysines were positively charged. In the case of PCL, the crystal has one single protein chain per unit cell of dimensions $a = 91.3 \text{ \AA}$, $b = 47.3 \text{ \AA}$, $c = 85.4 \text{ \AA}$;

$\alpha = 90.0^\circ$, $\beta = 121.4^\circ$, $\gamma = 90^\circ$. The chain system is composed of 320 amino acids, one calcium cation and 193 adsorbed water molecules. The crystal water molecules and the calcium cation were kept. The resulting structure was centered in a rectangular box with optimized dimensions of $a = 60 \text{ \AA}$, $b = 56 \text{ \AA}$, $c = 60 \text{ \AA}$; $\alpha = \beta = \gamma = 90.0^\circ$ and solvated with 4873 water molecules. In the case of CALB, the complex has six chains per unit cell of dimensions $a = 229.5 \text{ \AA}$, $b = 95.6 \text{ \AA}$, $c = 86.8 \text{ \AA}$; $\alpha = \beta = \gamma = 90.0^\circ$. Each chain system is composed of 317 amino acids, one ethylhexylphosphonate inhibitor (HEE) covalently bound to the Ser105, a dimer of N-acetyl-glucosamine covalently bound to the Asn74 and 92 adsorbed water molecules. Only one protein chain and its water molecules were kept; the five others were erased. In the remaining chain, the phosphonate inhibitor and the two N-acetyl-glucosamine dimers were removed. The resulting structure was centred in a rectangular box with optimized dimensions of $a = 64 \text{ \AA}$, $b = 59 \text{ \AA}$, $c = 63 \text{ \AA}$; $\alpha = \beta = \gamma = 90.0^\circ$ and solvated with 6145 water molecules. These new solvated systems were used as starting atomic coordinates set for the subsequent modelling steps.

2.2.2. Relaxation of the systems

The systems were relaxed following a careful relaxation protocol, in order to preserve the crystalline organization of protein atoms (Andrade et al., 2007; Da Cunha et al., 2007). In this protocol, constraints and restraints were initially added to the systems and then progressively removed throughout the procedure, totaling four steps: (a) In order to eliminate initial strains, all protein heavy atoms were fixed; only protein hydrogen atoms and water molecules were allowed to move; (b) A fix constraint was applied on the protein backbone, harmonic restraints were applied on the side chains and protein hydrogen atoms and water molecules were allowed to move; (c) An harmonic restraint was applied on the protein backbone and all the remaining parts of the system were free to move; (d) Finally, the whole system was minimized without any constraints or restraints. The step (a) was done by performing 1000 Steepest Descent iterations. The steps (b)-(d) were done using the Conjugate Gradient method until the root mean square deviation (rmsd) gradient fell below 0.01, 0.001 and 0.0001 kcal.mol⁻¹.Å⁻¹, respectively. Next, the optimized systems were equilibrated during 1ns of standard dynamic trajectories at 300 K, under periodic boundary conditions, using the NVT thermodynamic ensemble. For both PCL and CALB, the protein structure in the last frame was taken for the construction of the corresponding acetyl-lipase structures.

2.2.3. Construction of the acetyl-lipase structures

As the mechanism of lipase-catalyzed acylation involves the prior esterification of the catalytic serine by the acyl donor substrate, the side chain hydroxyls of the Ser87 of PCL and Ser105 of CALB were replaced by an acetate, in order to mimic the acetyl-lipase structure (Botta et al., 2002; De Oliveira et al., 2009; Palocci et al., 2007). Then, the energy of these new systems was further minimized by executing 1000 Conjugate Gradient steps, with a fix constraint applied to the protein backbone. The resulting acetyl-PCL and acetyl-CALB structures were taken as targets for the docking of quercetin.

2.3. Docking simulations

2.3.1. Docking and scoring procedures

Quercetin was non-covalently docked against the acetyl-CALB and acetyl-PCL using LigandFit which combines a Monte Carlo conformational search algorithm and a shape comparison filter. In addition to the geometric fit with the binding pocket, LigandFit entails an internal scoring function (DockScore) that assures to keep only the ligand poses having favourable energy of interaction with the protein, taken as the sum of van der Waals energies (represented by a 6-9 Lennard-Jones term) and electrostatic energies (represented by a Coulomb term). For both targets, 15,000 Monte Carlo trials were performed on the quercetin and only the 50 most favourable poses were retained. Conformations were clustered according to the RMSD between all pair-wise of poses. Moreover, poses were ranked considering a consensus of six different scoring functions (Ligscore1, Ligscore2, PLP1, PLP2, PMF and Jain) (Feher, 2006). The consensus score ranged from 0 to 6 and the best score was 6. Among complexes exhibiting a consensus score of 5 or 6, one complex was selected in each cluster of poses in order to sample quercetin conformations within lipase catalytic pocket. All selected complexes were submitted to subsequent modelling steps.

2.4. Molecular dynamics simulations

2.4.1. Molecular dynamics procedure

Molecular dynamics simulations were conducted in explicit water. That choice was motivated by several arguments: firstly, water is the most commonly used solvent for simulations of biomolecular systems (Roccatano, 2008); secondly, no solvent effect was experimentally observed on the selectivity of quercetin acetylation catalyzed by lipases, in protic or aprotic solvents of various logP (Chebil et al., 2007); finally, as reported in several molecular modelling studies, identical conformational behaviours were observed in organic solvents and in water, for both CALB and PCL, (Barbe et al., 2009; Roccatano, 2008; Soares et al., 2003; Trodler and Pleiss, 2008).

All selected complexes were surrounded by a periodic box of TIP3P (Jorgensen et al., 1983) water molecules with a margin of 15 Å along each dimension. The total number of water molecules was 14178 and 17912 for CALB systems and for PCL systems, respectively. The total charge of the systems was neutralized by the addition of 1 Na⁺ for CALB and 4 Na⁺ for PCL.

The energy of the solvated systems was minimized with 6400 steps of Conjugate Gradient using a NAMD minimization tool in order to optimize the relative positions of solvent molecules. Then, an equilibration phase of 1 ns was performed. During this step, enzymes and substrates were harmonically restrained. The force constant was initially fixed at 10 kcal.mol⁻¹.Å² and was progressively decreased (0.5 kcal.mol⁻¹.Å²/25 ps). The equilibration was continued over 500 ps. All productive MD simulations were performed at constant pressure (1 atm) and constant temperature (300 K), controlled *via* Langevin piston and *via* Langevin thermostat, respectively. The Particle-Mesh-Ewald (PME) method (Darden et al., 1993) was applied to calculate long range electrostatic interactions. The cutoff distance for nonbonded interactions was set to 12 Å and a switch function was applied between 10 and 12 Å. The SHAKE algorithm (Ryckaert et al., 1977) was applied to constrain

all of the covalent bonds involving hydrogen atoms. Periodic boundary conditions were applied to all dimensions. A time step of 1 femtosecond (fs) was applied. Coordinates were saved every 1 picosecond (ps). The dynamic behaviour of enzyme-substrate complexes was studied using 2000 frames from each simulation, covering the production phase (1 frame every 5 ps).

2.4.2. Analysis of the trajectories

Molecular dynamics (MD) simulations aimed to check the temporal stability of the enzyme-substrate complexes and to evaluate their propensity to lead to the formation of a tetrahedral intermediate and then to an ester. The root mean square deviation (rmsd) was computed for the protein backbone atoms referring to the starting structures of the MD simulation in order to check the stability of the trajectories. Then, MD trajectories were analyzed on the base of 2000 frames focusing on two major criteria: i) the distances between the carbonyl of the serine-bound acetate and the oxyanion hole residues. Indeed, these distances are indicative of the correct or incorrect orientation of the acetate to form the second tetrahedral intermediate, ii) the localization of quercetin hydroxyl groups with regard to the region comprised between the catalytic histidine and serine residues in order to identify the productive complexes (that means, those able to lead to the formation of the ester product) (Branneby et al., 2004; De Oliveira et al., 2009; Svedendahl et al., 2008)

3. RESULTS

3.1. Quercetin acetylation catalyzed by *Pseudomonas cepacia* lipase (PCL)

3.1.1. Quercetin-PCL docking

Docking simulations were first used to identify the possible binding modes of quercetin within the catalytic pocket of PCL after the formation of the acetyl-lipase intermediate. The PCL catalytic cavity is approximately 7 Å x 17 Å wide and 8 Å deep (Kim et al., 1997). The catalytic triad is constituted by the residues Ser87, His286 and Asp264 that are located at the center of the cavity bottom. The oxyanion hole residues are Leu17 and Gln88, whose backbone –NH groups can form two hydrogen bonds with the carbonyl oxygen of the serine-bound acetate. The internal part of the cavity exhibits a hydrophobic region including the residues Pro113, Ser117, Phe119, Val123, Leu164, Leu167 and Val267 and a slightly hydrophilic region, constituted by the residues Thr18, Val26, Leu27, Tyr29, Phe146, Ile290, Gln292 and Leu293. The cavity entrance consisting of the residues Leu248, Thr251, Val266 and Leu287 is mostly hydrophobic and includes the $\alpha 5$ helix lid (residues 139-150). The crystal structure used in simulations corresponds to the opened form of the enzyme.

The fifty most favourable complexes retained after the automated docking procedure were classified in two families of poses, PA and PB. The PA family comprises complexes in which the flavonoid A-ring was orientated towards the enzyme core whereas the PB family corresponds to complexes in which the flavonoid B-ring pointed towards the enzyme core. The complexes belonging to the PA and the PB family had consensus scores of 5 and 6, respectively. Within the PA family, the quercetin conformers presented only minor differences in their conformations and positions inside the enzyme cavity (RMSD of PA pair-wise between 0.03 and 0.40 Å). Therefore, only one PA complex (PAQ50) was selected as input structure for the subsequent molecular dynamic procedure. Within the

PB family, quercetin exhibited three major conformations and positions inside the enzyme cavity (RMSD of PB pair-wise between 0.41 and 3.90 Å). Thus, three PB complexes (PBQ12, PBQ21, PBQ22) representative of the different configurations were selected as input structures for molecular dynamic simulations.

3.1.2. Dynamics of the A-ring binding mode

The stability of the PCL structure all along the trajectory was monitored with the RMSD of the backbone atoms referring to the initial structure of the protein. After the equilibration phase, the RMSD values ranged from 1.7 to 1.9 Å during the first 5 ns of the production phase and from 2.3 to 2.7 Å during the last 5 ns. Main movements were observed on the β -hairpin residues (214 to 228). When neglecting these residues, the RMSD was slightly increasing from 1.6 to 2.0 Å showing that the presence of quercetin did not bring significant rearrangement of protein structure. These results are in accordance with a previous study that showed the high flexibility of this region through molecular dynamics simulations (Barbe et al., 2009; Lafaquière et al., 2009). As reported by Trodler et al. (Trodler et al., 2009), the movement of this region was observed to be independent of the movement of the lid. The dihedral angle O1-C2-C1'-C2' (τ) of quercetin was -110° after the docking procedure, then 180° at the end of the equilibration phase, showing that quercetin tended to come near planarity. During the production phase, this angle varied between -10° and $+10^\circ$ around the planar position.

The molecular dynamic trajectory of the A-ring binding mode was firstly analyzed with regard to the orientation of the serine-bound acetate towards the oxyanion hole in order to verify if the acyl-enzyme had a catalytically competent configuration. The carbonyl oxygen (Ace:O) of the acetate pointed towards the oxyanion hole at the end of the equilibration phase and kept this orientation all along the trajectory. The average Ace:O-Gln88:HN and Ace:O-Leu17:HN distances were 2.30 ± 0.44 Å and 2.69 ± 0.80 Å, respectively. There was only a short period (between 3.48 and 4.78 ns) during which the Ace:O moved somewhat far from Leu17 and Gln88 due to the insertion of a water molecule.

Two major orientations of quercetin within the PCL catalytic cavity were observed (Figure IV.3). At the beginning of the production phase, during 1 ns, quercetin was positioned in the catalytic cavity as follows (orientation 1) : the 7-OH group pointed towards the catalytic residues and the quercetin aromatic rings interacted with the aliphatic side chains of Phe 119 (helix α_4), Thr18, Ala247 and Leu248 (helix α_9), and Leu287 through $\text{CH}_3\text{-}\pi$ hydrophobic interactions (Morita et al., 2006; Ringer et al., 2006; Tsuzuki et al., 2006). In addition, the 4'-OH formed a H-bond interaction with the hydroxyl group of Thr251. Then quercetin moved and adopted the orientation 2 in the cavity, its 5-OH and 3-OH groups pointing towards the cavity entrance in contact with water molecules. The 7-OH group was observed to establish a H bond with the residue Thr18. The A-ring and the C-ring were shown to establish hydrophobic interactions with the side chains of aliphatic residues Leu248 and Leu287 while the B-ring interacted with the residues Leu248, Val266 and Phe119. Otherwise, two additional hydrogen bonds between 3'-OH and Thr251, and 7-OH and Thr18 were formed. Quercetin kept this position up to the end of the trajectory (during 5 ns).

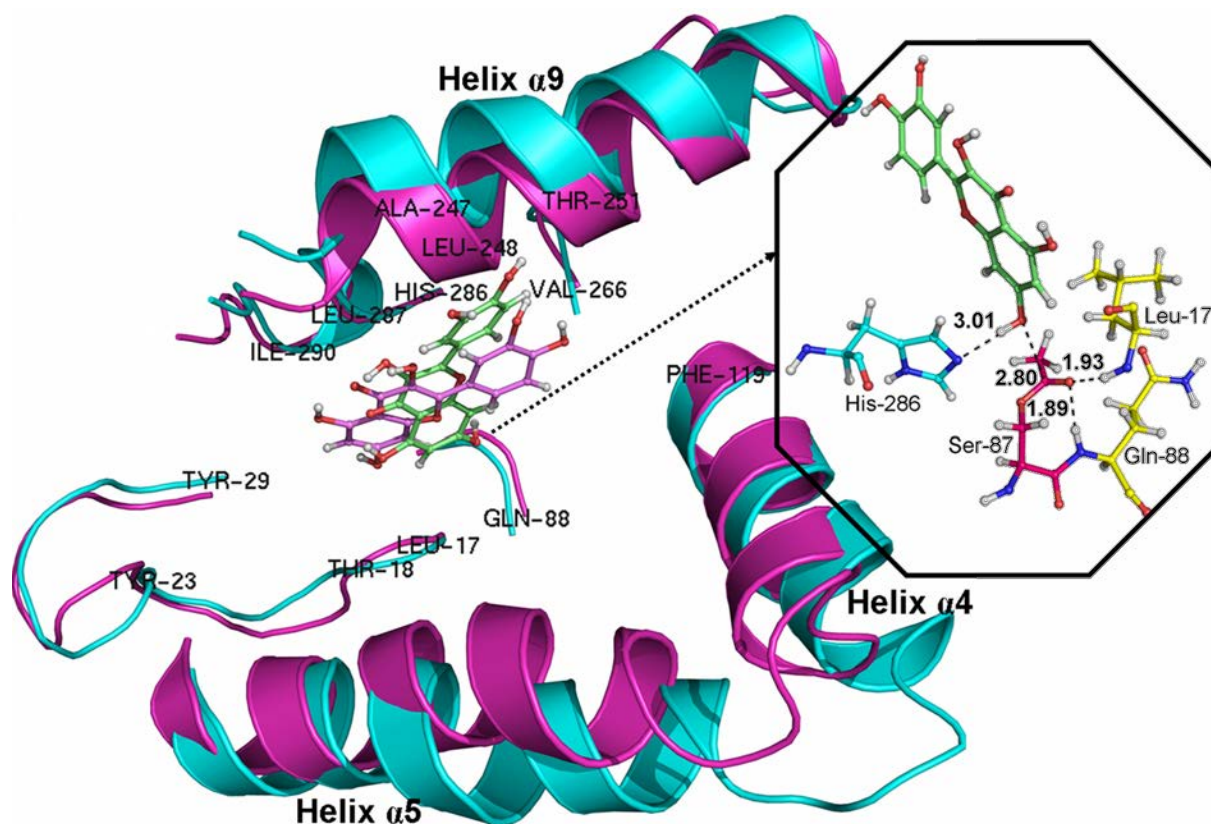


Figure IV.3. Superimposition of representative snapshots of the trajectory of the PAQ50 complex. In this frontal view, the three alpha helices constituting the cavity entrance are indicated. The two major orientations of quercetin are shown in pink (orientation 1) and green (orientation 2). Hydrogen bond interactions between the oxyanion hole residues (Leu17 and Gln88) and the carbonyl of the serine-bound acetate and distances (Å) between the quercetin 7-OH group and His286:N ϵ and Ace:C are shown in the zoom view.

In this A-ring binding mode, only the 7-OH group of quercetin was located close to the catalytic triad, during 1 ns of the trajectory, with distances to His286:N ϵ and Ace:C of 3.40 ± 0.57 Å and 3.13 ± 0.55 Å, respectively. These results respected the conditions reported in the literature (Palocci et al., 2007) as required for a complex to be considered as productive (distances to catalytic His and Ser < 4 Å), suggesting that acylation may occur on quercetin 7-OH position.

3.1.3. Dynamics of the B-ring binding modes

A similar molecular dynamic analysis was performed on the three complexes corresponding to the B-ring binding modes of quercetin within the catalytic pocket. During the MD production phase of complexes PBQ12, PBQ21 and PB22, the RMSD values using the starting structures as a reference were about 2.7 Å, 2.1 Å, 2.9 Å, respectively. As for the complex PAQ50, the main movements were observed in the β -hairpin region. The average RMSD in the absence of this region were 1.9 Å, 1.9 Å and 2.3 Å for PBQ12, PBQ21 and PBQ22 respectively, indicating that the three systems were stable throughout simulations. The dihedral angle τ of quercetin oscillated between +20° and -20° from the planar position. The average distances between the carbonyl oxygen of the serine-bound acetate (Ace:O) and the oxyanion hole residues, Leu17:HN and Gln88:HN, were 2.64 ± 0.89 Å and $2.32 \pm$

0.57 Å, respectively, that was in favour of H bond interactions. These results indicated that the acetate was in a catalytically competent orientation. Only few tilting points were observed corresponding to small disruptions of the hydrogen bond network due to the insertion of water molecules.

Moreover, the MD simulations of the three complexes showed that quercetin adopted two major orientations (orientation 3 and orientation 4) in the catalytic pocket as shown in Figure IV.4. In these orientations, the 3-OH and 5-OH groups of quercetin were orientated towards either the cavity entrance or the cavity bottom. The orientation 3 was characterized by the following interactions: hydrophobic interactions between quercetin A-ring and the residues Leu17, Phe119, Val266; hydrophobic interactions between quercetin B-ring and the residues Thr18, Leu287, Ile290; H bond interactions between the 3'-OH and 4'-OH groups of quercetin and the residues Tyr23, Tyr29. The orientation 4 was characterized by the following interactions: hydrophobic interactions between the quercetin B-ring and the residues Leu17, Thr18, Leu287, Ile290; H bond interactions between the 3'-OH and 4'-OH groups and the residues Thr18, Tyr23, Tyr29; hydrophobic interactions between the quercetin A- and C-rings and the side chain of the residues Leu17, Thr18, Leu248, Val266, Leu287.

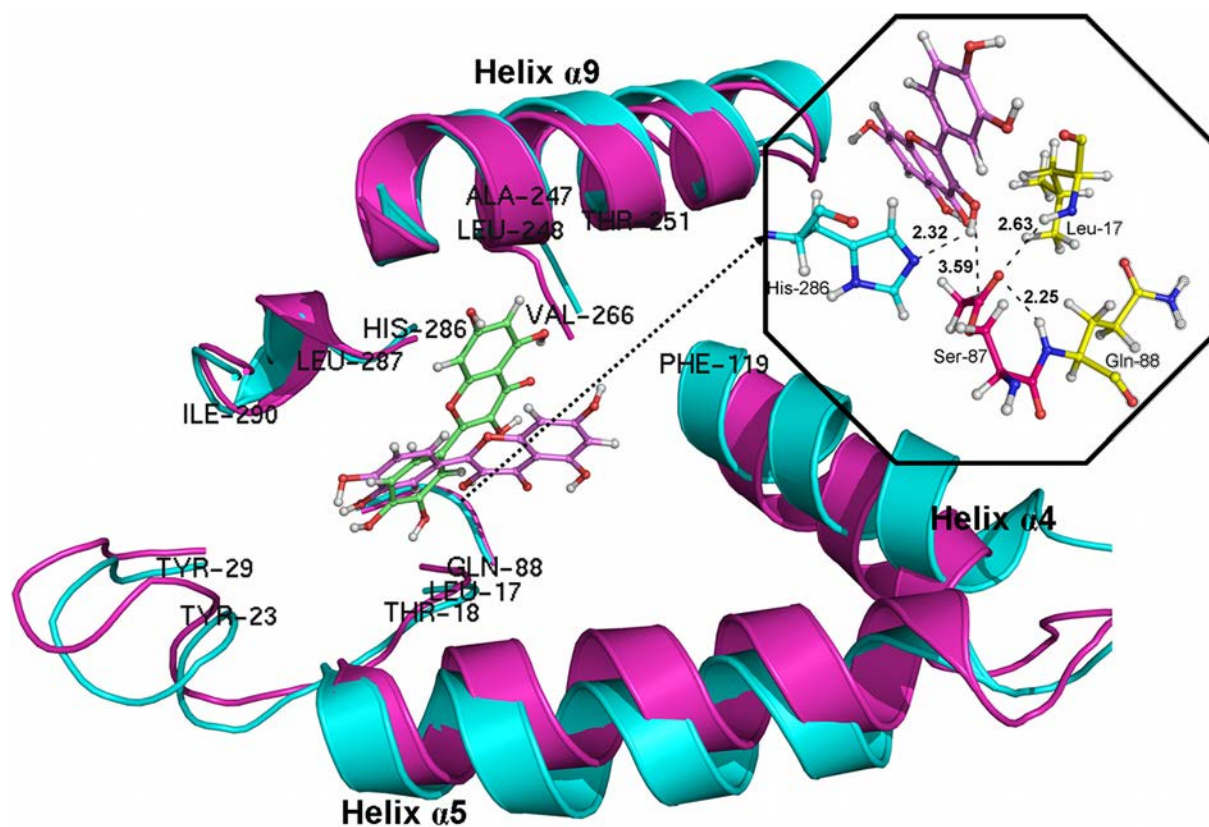


Figure IV.4. Superimposition of representative snapshots of the PBQ21 complex (frontal view). The orientations 3 and 4 of quercetin are shown in pink and in green respectively. In the zoom view, hydrogen bond interactions between Ace:O and the oxyanion hole and distances of the quercetin 3-OH group (orientation 3) to the catalytic residues are shown.

One particular event was observed for one complex (PBQ12) in which interactions between quercetin 4'-OH and 3'-OH groups and the residues Thr18, Tyr23, Tyr29 were disrupted at the end of

the trajectory (after 7.5 ns). In that case, quercetin moved between helices $\alpha 9$ and $\alpha 5$ and left the enzyme cavity for the solvent.

The trajectories were also analyzed basing on the proximity of the quercetin hydroxyl groups from the catalytic residue His286 and the serine-bound acetate. Results were shown in Table IV.1. When quercetin was in the orientation 3, the 3-OH and 5-OH groups were located close to the catalytic residues at distances inferior to 4 Å. The 5-OH group maintained an intramolecular H-bond interaction with the 4-keto group all along the trajectory (distance of 1.86 ± 0.13 Å and angle of $130.61 \pm 10.39^\circ$). In the case of the orientation 4, the nearest hydroxyl groups from the catalytic histidine and serine were the 3'-OH and 4'-OH.

Table IV.1. Average distances between the quercetin hydroxyl groups and the catalytic residue His286 and the serine-bound acetate. Only periods during which hydroxyl groups came at a distance inferior to 4 Å from Ace:C and His286:Ne were taken into account.

Complexes	Buried-OH group	Distance from Ace:C (Å)	Distance from His286:Ne (Å)	Time (ns)
PBQ12	3'-OH	3.47 ± 0.39	2.98 ± 0.40	1.8 to 2.2
	4'-OH	3.44 ± 0.61	3.48 ± 0.81	0 to 4.6
PBQ21	3-OH	3.52 ± 0.61	3.56 ± 0.48	0 to 6.7 9.2 to 10.0
	5-OH	3.32 ± 0.31	4.35 ± 0.79	6.7 to 9.2
PBQ22	3-OH	3.19 ± 0.30	3.12 ± 0.34	0 to 5.9

According to the results obtained for the two binding modes, the combined docking and molecular dynamic protocol suggested that all hydroxyl groups of quercetin could be acetylated using the PCL lipase as biocatalyst. Whatever the initial position of quercetin at the beginning of the trajectories, the flavonoid was observed to alternate between two major orientations within the catalytic cavity. Major interactions were observed between aromatic rings of the flavonoid and hydrophobic or aromatic residues of the enzyme. H bond interactions were also identified involving quercetin hydroxyl groups and residues located at the cavity bottom.

3.2. Quercetin acetylation catalyzed by *Candida antarctica* lipase B (CALB)

3.2.1. Quercetin-CALB docking

The catalytic cavity of CALB is a narrow channel with approximately 10 Å x 4 Å of width and 12 Å of deep (Uppenberg et al., 1994). Contrary to most of known lipases, CALB has no lid covering its active site. The catalytic triad is composed by the residues Ser105, His224 and Asp187 and the oxyanion hole contains the residues Thr40 and Gln106. In this case, the oxyanion of the tetrahedral intermediate can be stabilized by hydrogen bond interactions not only with the backbone -NH of Thr40 and Gln106 but also with the side chain -OH of Thr40. The cavity wall doesn't contain any aromatic residues and is essentially lined with the aliphatic hydrophobic residues Leu140, Ala141, Leu144, Val149, Val154, Ile189, Leu278, Ala281, Ala282, Ile285 and Ala286. CALB catalytic pocket also

entails a small slightly polar region comprising the catalytic Ser105 and made of the aliphatic polar side chains of Thr40, Asp134 and Gln157.

As in the case of PCL, the docking and scoring procedures gave fifty complexes, classified in two families with either the A-ring (CA) or the B-ring (CB) of quercetin orientated towards the enzyme core. Within the CA family, conformers obtained a maximal consensus score of 5 and exhibited only minor conformational differences (RMSD pair-wise between 0.04 and 0.55 Å). Two poses were selected differing in the value of the dihedral angle τ (CAQ20 and CAQ2). Conformers of the CB family obtained a consensus score of 6 and also exhibited minor differences (RMSD pair-wise between 0.02 and 0.52 Å). Therefore, only one CB complex (CBQ18) was selected for further simulations.

3.2.2. Dynamics of the A-ring binding modes

The stability of CALB in presence of quercetin, along the trajectories, was monitored with the RMSD of the backbone atoms from the initial structure of the protein. Average RMSD values were about 1.4 Å and 1.5 Å, respectively for CAQ2 and CAQ20. During the production phase, the dihedral angle τ oscillated from -20° to $+20^\circ$ around the planar position for both complexes.

The distances between the carbonyl oxygen atom of the serine-bound acetate (Ace:O) and the oxyanion hole residues were also monitored and their average values are reported in Table IV.2. In both complexes the carbonyl of the serine-bound acetate was globally not oriented towards the oxyanion hole residues during the trajectories and no hydrogen bonds with the residues Thr40 and Gln106 were established as shown by the average distances Ace:O-Gln106:HN, Ace:O-Thr40:HN and Ace:O-Thr40:HO (Figure IV.5). This position was maintained all along the trajectories. Such results indicated an unfavourable orientation of the acetate for the formation of the second tetrahedral intermediate.

Table IV.2. Average distances (Å) between the carbonyl oxygen of the serine-bound acetate and the oxyanion hole residues throughout the 10 ns molecular dynamics trajectories for both CA complexes.

Complexes	Ace:O – oxyanion hole		
	Gln106:HN	Thr40:HN	Thr40:HO
CAQ2	4.18 ± 0.18	5.54 ± 0.63	7.31 ± 0.82
CAQ20	3.63 ± 0.45	4.03 ± 0.73	3.22 ± 0.53

The positions and orientations of quercetin within the catalytic cavity of CALB were also analyzed aiming to understand the dynamic behaviour of this substrate. In both CA complexes, the flavonoid was stabilized during 10 ns of the trajectories by hydrophobic interactions between its flavan backbone and the side chains of the aliphatic residues neighbouring the cavity entrance (Leu140, Leu144, Val154, Val190, Leu 278, Ala282). More specifically, CH- π interactions between the side chains of the amino acids Ile189 and Ile285 and the flavan backbone of quercetin were observed (Table IV-3).

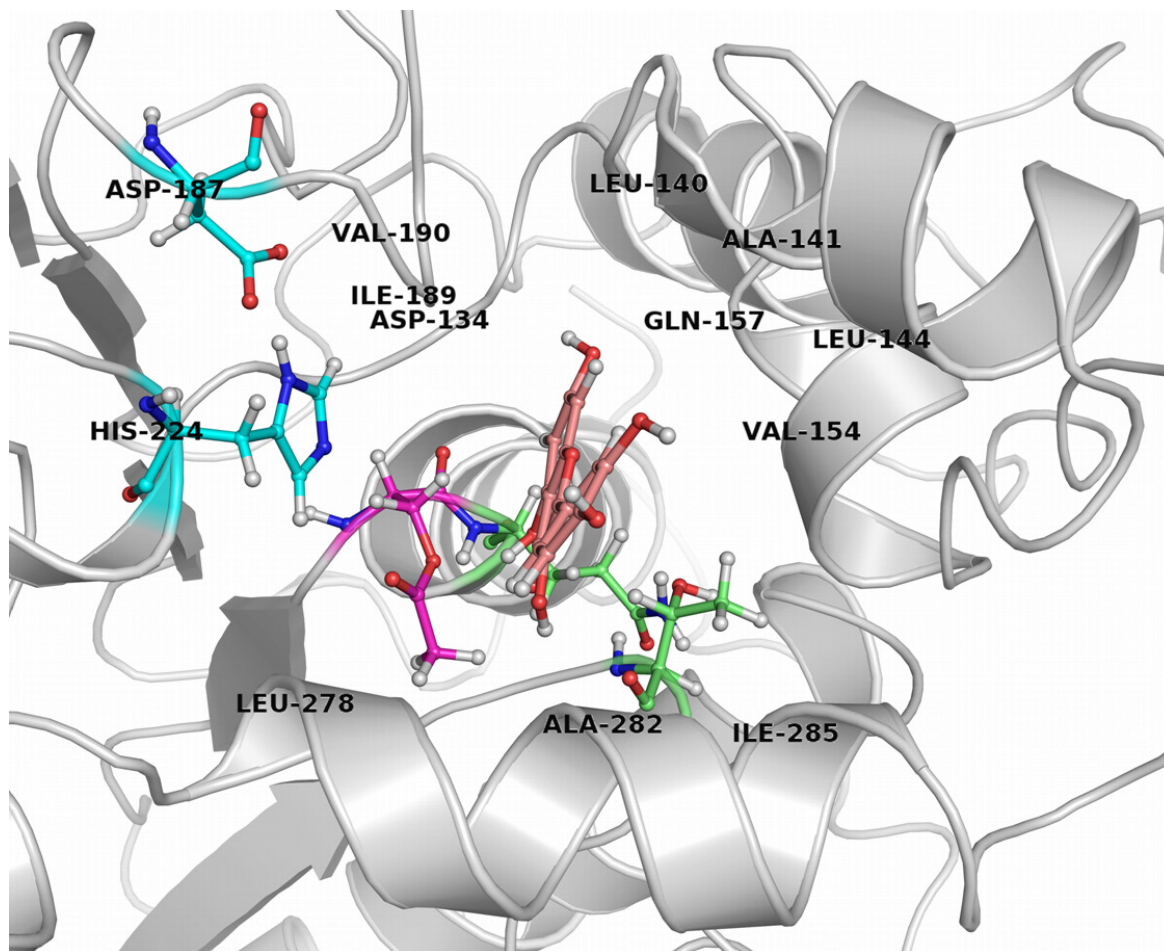


Figure IV.5. Frontal view of the CAQ2 complex. Residues of the oxanion hole (Thr40 and Gln106) are shown in green, side chains of the catalytic triad in blue for residues His224 and Asp134 and pink for the Ser105-bound acetate. The residues neighbouring the cavity entrance are labelled.

Table IV.3. Distances (Å) between the center of mass of quercetin rings and the carbon atoms of the side chains methyl group of Ile189 and Ile285.

Complexes	Quercetin rings	Residues	Distances
CAQ2	A	Ile189:C γ 2	4.06 \pm 0.32
	C	Ile189:C γ 2	3.78 \pm 0.22
		Ile285:C δ 1	4.75 \pm 0.54
	B	Ile189:C δ 1	4.09 \pm 0.37
CAQ20	A	Ile189:C γ 2	5.04 \pm 0.56
		Ile285:C δ 1	5.47 \pm 0.56
	C	Ile189:C δ 1	4.90 \pm 0.63
		Ile189:C γ 2	4.43 \pm 0.48
		Ile285:C δ 1	4.28 \pm 0.37
B	Ile189:C δ 1	4.73 \pm 0.87	
CBQ18	A	Ile 189:C δ 1	4.22 \pm 0.51
		Ile285:C δ 1	4.15 \pm 0.57
	C	Ile189:C δ 1	4.29 \pm 0.49
		Ile189:C γ 2	4.14 \pm 0.36
		Ile285:C δ 1	3.76 \pm 0.47
B	Ile189:C γ 2	4.05 \pm 0.32	

The 7-OH group was involved in a strong hydrogen bond with $O_{\delta_{Asp134}}$ (CAQ2: $2.01 \pm 0.17 \text{ \AA}$; $144.72 \pm 11.47^\circ$) or O_{Gln157} (CAQ20: $1.91 \pm 0.18 \text{ \AA}$; $157.25 \pm 11.85 \text{ \AA}$) which constitute with Thr40 a slightly polar region near the catalytic residues. The 3-OH and 5-OH groups formed an intra-molecular hydrogen bond with the 4-keto group and an inter-molecular H-bond with the residue Thr40 (CAQ2) or Ace:O (CAQ20).

The 4'-OH and 3'-OH groups were localized at the cavity entrance and interacted with water molecules. All these electrostatic and hydrophobic interactions stabilized quercetin in the binding pocket of CALB. The quercetin enone moiety was located between the Thr40 and the carbonyl of the serine-bound acetate that was diverted from the oxyanion hole.

These results suggested that the A-ring binding modes of quercetin within the CALB active site led to unproductive complexes where the two substrates (the acetate and the flavonoid) exhibited incorrect orientations with regard to the formation of the second tetrahedral intermediate.

3.2.3. Dynamics of the B-ring binding mode

During the production phase, the protein structure was stable all along the trajectory (average RMSD value of 1.5 \AA). As in the A-ring binding modes, the dihedral angle τ oscillated from -20° to $+20^\circ$ around the planar position during the production phase.

In the B-ring binding mode (complex CBQ18), the carbonyl of the serine-bound acetate interacted with the oxyanion hole residues through two hydrogen bonds during two periods of time : 0 to 2.3 ns, then 5.3 to 5.7 ns. Characteristic average distances were $1.99 \pm 0.15 \text{ \AA}$, $2.23 \pm 0.26 \text{ \AA}$ and $4.68 \pm 0.47 \text{ \AA}$ for Ace:O-Gln106:HN, Ace:O-Thr40:HN and Ace:O-Thr40:HO, respectively. Apart from these periods, the serine-bound acetate turned away from the oxyanion hole region, as shown by the average distances Ace:O-Gln106:HN ($4.12 \pm 0.18 \text{ \AA}$), Ace:O-Thr40:HN ($5.30 \pm 0.53 \text{ \AA}$), Ace:O-Thr40:HO ($6.83 \pm 0.61 \text{ \AA}$).

Quercetin was mainly stabilized by CH- π hydrophobic interactions with the side chain of residues Leu140, Leu144, Ile189, Ala281 and Ile285. As for the A-binding mode, the aromatic rings of quercetin strongly interacted with side chains of Ile189 and Ile285 (Figure IV.6) as shown by average distances reported in Table IV.3.

The 4'-OH group was implied in two hydrogen bonds with the residues Asp134 and Gln157, as shown by the average distances $H4'_{\text{quercetin}}-O_{\delta 1_{Asp134}}$ ($2.01 \pm 0.15 \text{ \AA}$, $154.78 \pm 10.90^\circ$) and $O4'_{\text{quercetin}}-H_{\epsilon_{Gln157}}$ ($2.37 \pm 0.80 \text{ \AA}$, $143.26 \pm 22.54^\circ$). The 3'-OH formed an intra-molecular hydrogen bond with the 4'-OH ($2.29 \pm 0.16 \text{ \AA}$, $1134.47 \pm 7.13 \text{ \AA}$) and sometimes interacted with the residue Gln106:HN. In this orientation, quercetin pointed its 7-OH group towards the cavity entrance all along the trajectory and was in contact with water molecules.

The proximity of the quercetin hydroxyl groups to the catalytic residues was studied during the two periods where the serine-bound acetate had a catalytically competent orientation. The closest hydroxyl groups were the 3'-OH and the 3-OH. The 3'-OH was localized at average distances of $5.51 \pm 0.96 \text{ \AA}$ and $4.42 \pm 0.29 \text{ \AA}$ from the His224:N ϵ and the Ace:C, respectively. In the case of the 3-OH group, these distances were $4.49 \pm 0.85 \text{ \AA}$ and $3.53 \pm 0.32 \text{ \AA}$. Considering simultaneously the distances relative to the histidine and the acetate, the acetylation distances criterion exceeds the

maximal values admitted as to allow the proton transfer and the nucleophilic attack. Clearly, the B-ring binding mode cannot be expected to lead to the formation of the second tetrahedral intermediate between quercetin and the acetyl-CALB.

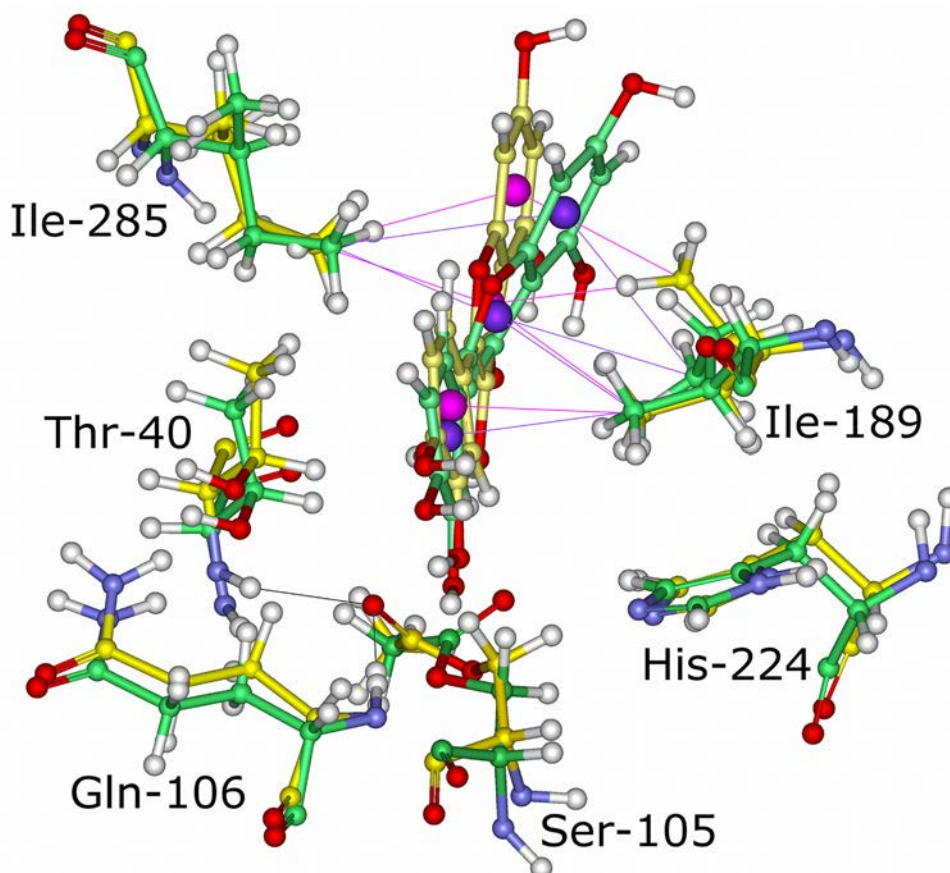


Figure IV.6. Superimposition of representative snapshots of the CBQ18 complex when the carbonyl of the serine-bound acetate pointed towards (yellow) and away (green) the oxyanion hole. CH- π interactions between the center of mass of quercetin aromatic rings, represented by a sphere, and residues Ile189 and Ile285 are indicated in solid lines. (centroid pink for yellow quercetin and centroid purple for green quercetin).

According to the modelling results obtained for the two binding modes (A-ring and B-ring), the combined docking and molecular dynamic protocol predicts that the quercetin acetylation is unlikely with CALB as biocatalyst. These modelling predictions are corroborated by previous experimental results, which demonstrated the absence of any acetylation activity of CALB towards quercetin in the presence of vinyl acetate as acyl donor (Ardhaoui et al., 2004; Chebil et al., 2007).

4. DISCUSSION

In this study, a combined docking and molecular dynamics protocol was applied to theoretically investigate the lipase catalyzed acetylation of quercetin. Within the global transesterification reaction mechanism, the modelling only considered the binding of the flavonoid to the acetylated lipase to form the second tetrahedral intermediate. It involved an initial docking procedure which identified two possible orientations of quercetin with either the A or the B-ring

pointing towards the catalytic residues, followed by molecular dynamics calculations over a 10 ns time scale, to assess the stability of the binding modes and to evaluate the proximity of the two substrates to the lipase catalytic residues. Considering the reaction feasibility, modeling results suggested that quercetin acylated derivatives can be synthesized using the PCL as biocatalyst, whereas the CALB did not catalyze the reaction. These results are in agreement with experimental data (Ardhaoui et al. 2004; Chebil et al. 2007). The outlined methodology which aims to explain the productivity of quercetin binding modes is based on several stability and proximity criteria. The formation of the second tetrahedral intermediate first requires a stabilization of the serine-bound acetate within the oxyanion hole. In the case of PCL, the over-mentioned carbonyl was located at a very close distance to the oxyanion hole residues all along the trajectories, indicating a good orientation of the first substrate in the catalytic cavity. In the case of CALB, this criterion is not respected. The presence of quercetin was suggested to be responsible for the bad orientation of the serine-bound acetate which was pushed away from the oxyanion hole. In order to assess such assumption, molecular dynamic simulations of the acyl-enzyme without quercetin were run (results not shown) and revealed a good stabilization of the serine-bound acetate within the oxyanion hole all along the trajectory.

Furthermore, the ability of the models to predict the regioselectivity was discussed. Considering the proximity of the flavonoid hydroxyl groups both to the catalytic histidine and to the serine-bound acetate, in the case of PCL, all quercetin hydroxyls groups came at distances less than 4 Å from the His286:Ne and from the Ace:C atom during the trajectories, suggesting that all of them could be acylated. However, experimental data showed that only the 4'-OH, 3'-OH and 7-OH were acylated (Chebil et al. 2007). This apparent divergence between modeling and experimental results could be explained by the fact that models didn't reflect the chemical reactivity of substrates, and especially the nucleophilic character of quercetin hydroxyl groups. Models must be carefully analyzed, including a good knowledge of all molecular entities implied in the system. Although the 5-OH group of the PBQ21 complex was close to the catalytic residues, it was maintained in an intra-molecular hydrogen bond with the 4-keto group. This strong interaction has been yet reported by several theoretical studies about quercetin (Cornard et al., 1997; Leopoldini et al., 2006; Mendonza-Wilson and Glossman-Mitnik, 2004). Moreover, the 3-OH and 5-OH groups were reported to be the least acidic groups of quercetin (Leopoldini et al., 2006; Martins et al., 2004; Musialik et al., 2009). On a chemical point of view, these conclusions strongly suggest that both 3-OH and 5-OH groups exhibit a poor nucleophilic character and seem unlikely to be acylated.

In the active site of PCL, quercetin underwent larger displacements than in CALB cavity. The main amino acids involved in interactions between quercetin and the cavity walls are Leu17, Leu248, Val266, Leu287 and Ile290. Guieysse et al. (Guieysse et al., 2003) reported similar observations and identified the network of hydrophobic side chains of PCL as an efficient dynamic guiding system for substrates. Otherwise, major shifts of quercetin in the PCL pocket may be due to a gain or loss of hydrogen bond interactions with the residues Tyr23, Tyr29 or Thr18 of the cavity bottom. On the contrary, in the CALB pocket, quercetin was strongly stabilized by electrostatic interactions with residues localized at the cavity bottom (Asp134, Gln157, and Thr40) and by hydrophobic interactions involving side chains of aliphatic residues. More specifically, methyl groups of Ile189 and Ile285

stabilized quercetin at the cavity entrance, preventing its correct orientation towards the catalytic residues. These results constitute guiding elements for the conception of a new biocatalyst able to catalyze the regioselective acylation of quercetin by the rational redesign of CALB.

5. REFERENCES

- Anderson, E.M., Larsson, K.M., Kirk, O., (1998) One biocatalyst - Many applications: the use of *Candida antarctica* B-lipase in organic synthesis. *Biocatalysis Biotransformation* 16, 181-204.
- Andrade, C.K.Z., Silva, W.A., Maia, E.R., (2007) Computational approach for the design of AP1867 analogs: aiming at new synthetic routes for potential immunosuppressant agents. *J. Biomol. Struct. Dyn.* 25, 35-48.
- Ardhaoui, M., Falcimaigne, A., Engasser, J.M., Moussou, P., Pauly, G., Ghoul, M., (2004) Acylation of natural flavonoids using lipase of *Candida antarctica* as biocatalyst. *J. Mol. Catal. B: Enzym.* 29, 63-67.
- Barbe, S., Lafaquière, V., Guieysse, D., Monsan, P., Remaud-Siméon, M., André, I., (2009) Insights into lid movements of *Burkholderia cepacia* lipase inferred from molecular dynamics simulations. *Proteins: Struct., Funct., Genet.* 77, 509-523.
- Boots, A.W., Haenen, G.R.M.M., Bast, A., (2008) Health effects of quercetin: from antioxidant to nutraceutical. *Eur. J. Pharmacol.* 585, 325-337.
- Botta, B., Zappia, G., Tafi, A., Botta, M., Manetti, F., Cernia, E., Milana, G., Palocci, C., Soro, S., Delle Monache, G., (2002) Lipase-catalyzed regioselective acylation of resorcin[4]arenes. *J. Mol. Catal. B: Enzym.* 16, 241-247.
- Branneby, C., Carlqvist, P., Hult, K., Brinck, T., Berglund, P., (2004) Aldol additions with mutant lipase: analysis by experiments and theoretical calculations. *J. Mol. Catal. B: Enzym.* 31, 123-128.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M., (1983) CHARMM: A program for macromolecular energy minimization and dynamics calculations. *J. Comput. Chem.* 4, 187-217.
- Chebil, L., Anthoni, J., Humeau, C., Gerardin, C., Engasser, J.M., Ghoul, M., (2007) Enzymatic acylation of flavonoids: effect of the nature of the substrate, origin of lipase and operating conditions on conversion yield and regioselectivity. *J. Agric. Food Chem.* 53, 9496-9502.
- Chebil, L., Humeau, C., Falcimaigne, A., Engasser, J.M., Ghoul, M., (2006) Enzymatic acylation of flavonoids. *Process Biochem.* 41, 2237-2251.
- Chen, C., Zhou, J., Ji, C., (2010) Quercetin: a potential drug to reverse multidrug resistance. *Life Sci.* 87, 333-338.
- Cornard, J.P., Merlin, J.C., Boudet, A.C., Vrielynck, L., (1997) Structural study of quercetin by vibrational and electronic spectroscopies combined with semiempirical calculations. *Biospectroscopy* 3, 183-193.
- Da Cunha, E.F.F., De Castro Ramalho, T., De Alencastro, R.B., Maia, E.R., (2007) Docking simulations and QM/MM studies between isoniazid prodrug, catalase-peroxidase (KatG) and S315T mutant from *Mycobacterium tuberculosis*. *Comput. Math. Methods Med.* 8, 113-124.
- Danieli, B., Luisetti, M., Sampognaro, G., Carrea, G., Riva, S., (1997) Regioselective acylation of polyhydroxylated natural compounds catalyzed by *Candida antarctica* lipase B (Novozyme 435) in organic solvents. *J. Mol. Catal. B: Enzym.* 3, 193-201.
- Darden, T., York, D., Pedersen, L., (1993) Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089-10092.
- De Oliveira, E.B., Humeau, C., Chebil, L., Maia, E.R., Dehez, F., Maigret, B., Ghoul, M., Engasser, J.M., (2009) A molecular modelling study to rationalize the regioselectivity in acylation of flavonoid glycosides catalyzed by *Candida antarctica* lipase B. *J. Mol. Catal. B: Enzym.* 59, 96-105.
- Feher, M., (2006) Consensus scoring for protein-ligand interactions. *Drug Discov. Today* 11, 421-428.
- Gotor-Fernandez, V., Brieva, R., Gotor, V., (2006) Lipases: useful biocatalysts for the preparation of pharmaceuticals. *J. Mol. Catal. B: Enzym.* 40, 111-120.
- Guieysse, D., Salagnad, C., Monsan, P., Remaud-Simeon, M., Tran, V., (2003) Towards a novel explanation of *Pseudomonas cepacia* lipase enantioselectivity via molecular modelling of the enantiomer trajectory into the active site. *Tetrahedron: Asymmetry* 14, 1807-1817.
- Gusdinari, T., Herowati, R., Kartasasmita, R.E., Adnyana, I.K., (2011) Anti-inflammatory and antioxidant activity of quercetin-3, 3', 4'-triacetate. *J. Pharmacol. Toxicol.* 6, 182-188.

- Humphrey, W., Dalke, A., Shulten, K., (1996) VMD-Visual Molecular Dynamics. *J. Mol. Graphics* 14, 33-38.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926-935.
- Kim, K.K., Song, H.K., Shin, D.H., Hwang, K.Y., Suh, S.W., (1997) The crystal structure of a triacylglycerol lipase from *Pseudomonas cepacia* reveals a highly open conformation in the absence of a bound inhibitor. *Structure* 5, 173-185.
- Lafaquière, V., Barbe, S., Puech-Guenot, S., Guieysse, D., Cortès, J., Monsan, P., Siméon, T., André, I., Remaud-Simeon, M., (2009) Control of lipase enantioselectivity by engineering the substrate binding site and access channel. *ChemBioChem* 10, 2760-2771.
- Leopoldini, M., Russo, N., Toscano, M., (2006) Gas and liquid phase acidity of natural antioxidants. *J. Agric. Food Chem.* 54, 3078-3085.
- MacKerell Jr., A.D., Bashford, D., Bellot, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher III, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., Karplus, M., (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem.* 102, 3586-3616.
- Martins, H.F.P., Leal, J.P., Fernandez, M.T., Lopez V.H.C., Cordeiro, M.N.D.S., (2004) Towards the predictions of the activity of antioxidants: experimental and theoretical study of the gas-phase acidities of flavonoids. *J. Am. Soc. Mass Spectrom.* 15, 848-861.
- Mendonza-Wilson, A.M., Glossman-Mitnik, D., (2004) CHH-DFT determination of the molecular structure, infrared and ultraviolet spectra of the flavonoid quercetin. *J. Mol. Struct.: THEOCHEM* 681, 71-76.
- Montenegro, L., Carbone, C., Maniscalco, C., Lambusta, D., Nicolosi, G., Ventura, C.A., Puglisi, G., (2007) In vitro evaluation of quercetin-3-O-acyl esters as topical prodrugs. *Int. J. Pharm.* 336, 257-262.
- Morita, S., Fujii, A., Mikami, N., Tsuzuki, S., (2006) Origin of the attraction in aliphatic C-H/ π interactions: infrared spectroscopic and theoretical characterization of gas-phase clusters of aromatics with methane. *J. Phys. Chem. A* 110, 10583-10590.
- Murakami, A., Ashida, H., Terao, J., (2008) Multitargeted cancer prevention by quercetin. *Cancer Lett.* 269, 315-325.
- Musialik, M., Kuzmich, R., Pawlowski, T.S., Litwinienko, G., (2009) Acidity of hydroxyl groups: an overlooked influence on antiradical properties of flavonoids. *J. Org. Chem.* 74, 2699-2709.
- Palocci, C., Falconi, M., Alcaro, S., Tafi, A., Puglisi, R., Ortuso, F., Botta, M., Alberghina, L., Cernia, E., (2007) An approach to address *Candida rugosa* lipase regioselectivity in the acylation reactions of trytilated glucosides. *J. Biotechnol.* 128, 908-918.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., Schulten, K., (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781-1802.
- Ringer, A.L., Figgs, M.S., Sinnokrot, M.O., Sherrill, C.D., (2006) Aliphatic C-H/ π interactions: methane-benzene, methane-phenol, and methane-indole complexes. *J. Phys. Chem. A* 110, 10822-10828.
- Roccatano, D., (2008) Computer simulations study of biomolecules in non-aqueous or cosolvent/water mixture solutions. *Curr. Protein Peptide Sci.* 9, 407-426.
- Ryckaert, J.P., Ciccotti, G., Berendsen, J.C., (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327-341.
- Saija, A., Tomaino, A., Trombetta, D., Pellegrino, M.L., Tita, B., Messina, C., Bonina, F.P., Rocco, C., Nicolosi, G., Castelli, F., (2003) 'In vitro' antioxidant and photoprotective properties and interaction with model membranes of three new quercetin esters. *Eur. J. Pharm. Biopharm.* 56, 167-174.
- Soares, C.M., Teixeira, V.H., Baptista, A.M., (2003) Protein structure and dynamics in nonaqueous solvents: insights from molecular dynamics simulation studies. *Biophys. J.* 84, 1628-1641.
- Svedendahl, M., Carlqvist, P., Branneby, C., Allnér, O., Frise, A., Hult, K., Berglund, P., Brinck, T., (2008) Direct epoxidation in *Candida antarctica* lipase B studied by experiment and theory. *ChemBioChem* 9, 2443-2451.

Teng, R.-W., Bui, T.-K.-A., McManus, D., Armstrong, D., Mau, S.-L., Bacic, A., (2005) Regioselective acylation of several polyhydroxylated natural compounds by *Candida antarctica* lipase B. *Biocatalysis Biotransformation* 23, 109-116.

Trodler, P., Pleiss, J., (2008) Modeling structure and flexibility of *Candida antarctica* lipase B in organic solvents. *BMC Struct. Biol.* 8, art. n° 9.

Trodler, P., Schmid, R.D., Pleiss, J., (2009) Modelling of solvent-dependent conformational transitions in *Burkholderia cepacia* lipase. *BMC Struct. Biol.* 9, art. n° 38.

Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., Fujii, A., (2006) Magnitude and directionality of the interaction energy of the aliphatic CH/ π interaction: significant difference from hydrogen bond. *J. Phys. Chem. A* 110, 10163-10168.

Uppenberg, J., Hansen, M.T., Patkar, S., Jones, T.A., (1994) The sequence, crystal structure determination and refinement of two forms of lipase B from *Candida antarctica*. *Structure* 2, 293-308.

Venkatachalam, C.M., Jiang, X., Oldfield, T., Waldman, M., (2003) LigandFit : a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graphics Model.* 21, 289-307.

Xanthakis, E., Theodosiou, E., Magkouta, S., Stamatis, H., Loutrari, H., Roussos, C., Kollis, F., (2010) Enzymatic transformation of flavonoids and terpenoids: structural and functional diversity of the novel derivatives. *Pure Appl. Chem.* 82, 1-16.

3. Contribution de l'article

Dans cette étude, une explication sur la différence de spécificité des lipases CALB et PCL dans l'acétylation de la quercétine est proposée. Pour cela, un protocole associant docking et dynamique moléculaire a été utilisé. A l'issue des simulations de docking, deux familles conformationnelles correspondant à deux orientations majeures de la quercétine dans les sites actifs de CALB et PCL ont été identifiées : soit le cycle A, soit le cycle B du flavonoïde pointe vers le cœur de l'enzyme. Plusieurs complexes de chaque famille ont ensuite été soumis à des simulations de dynamique moléculaire afin d'étudier leur comportement dynamique. L'analyse des trajectoires a permis de mettre en évidence les événements moléculaires qui pourraient expliquer la différence de spécificité des deux lipases vis-à-vis de la quercétine.

Cette étude permet de parvenir aux faits suivants :

- Le positionnement du donneur d'acyle a été évalué par le nombre de liaison hydrogène établi entre le carbonyle de l'acétate lié à la sérine catalytique et les résidus du trou oxyanionique. Dans le cas de la CALB, le donneur d'acyle ne forme pas de liaisons hydrogène avec les résidus du trou oxyanionique. De plus, les groupements hydroxyles de la quercétine sont localisés à une distance supérieure à 4 Å. En revanche dans la PCL, deux liaisons hydrogène sont présentes sur une majorité des trajectoires. De ce fait, les distances entre les fonctions hydroxyles de la quercétine insérées dans la poche du site actif de la PCL et les résidus catalytiques ont été mesurées. Des valeurs inférieures à la distance seuil fixée (4 Å) ont été constatées pour tous les groupements hydroxyles de la quercétine.
- Des résidus spécifiques dans le site actif de la CALB et PCL ont montré leur implication particulière dans la stabilisation de la quercétine. Dans le cas de l'enzyme CALB, les résidus aliphatiques hydrophobes Ile189 et Ile285 encerclent étroitement le substrat tandis que dans l'enzyme PCL, les résidus polaires (Thr18, Tyr23, Tyr29) et aliphatiques hydrophobes tapissant le mur de la cavité permettent des fluctuations positionnelles marquées du substrat.
- Les interactions qui contribuent à stabiliser la quercétine dans la cavité catalytiques de la CALB et, qui participent à l'acheminement du substrat près des résidus catalytiques dans le cas de la PCL sont de type électrostatique (liaisons hydrogène) et hydrophobique ($\text{CH}_3\text{-}\pi$).
- Pour l'acétylation enzymatique de la quercétine avec le biocatalyseur CALB, les résultats expérimentaux sont en concordance avec les résultats de modélisation : aucune estérification de la quercétine n'est détectée. Lors de l'utilisation du biocatalyseur PCL, des produits estérifiés en 3', 4' et 7 de la quercétine ont été identifiés expérimentalement. D'après les critères de distance fixés, les résultats de modélisation indiquent que toutes les fonctions hydroxyles de la quercétine sont acylables.

D'après la méthodologie employée et nos connaissances actuelles sur le mécanisme catalytiques des lipases, cette étude permet d'établir les conclusions suivantes :

- L'absence d'activité catalytique observée expérimentalement avec la CALB dans l'acétylation de la quercétine est la conséquence d'une orientation inappropriée du donneur d'acyle liée à

la sérine catalytique émanant de la présence de la quercétine dans le site actif et, d'une proximité insuffisante des groupements hydroxyles de la quercétine des résidus catalytiques.

- Le positionnement éloigné de la quercétine vis-à-vis des résidus catalytiques provient d'une part, de la rigidité de la molécule de flavonoïde combinée à l'étroitesse du site actif et d'autre part, des interactions hydrophobiques avec les résidus Ile189 et Ile285 ainsi que des interactions électrostatiques avec les acides aminés Asp134 et Gln157.
- L'activité d'acétylation de la PCL est corroborée par les résultats de la modélisation moléculaire qui révèlent une orientation des deux substrats, acétate et quercétine, appropriée à l'attaque nucléophile par le flavonoïde.
 - Le positionnement adéquat de la quercétine est favorisé par la plus grande taille du site actif et par des interactions avec des résidus polaires et aliphatiques hydrophobes qui guident l'acheminement du substrat.
 - Les simulations de dynamique moléculaire indiquent que toutes les fonctions hydroxyles de la quercétine sont susceptibles d'être acylées. Elles devront être complétées par l'utilisation d'outils complémentaires de modélisations tenant compte de la réactivité des groupements hydroxyles de la quercétine afin d'être en accord avec les résultats expérimentaux montrant une sélectivité d'acylation en position 3', 4' et 7 de la quercétine.

4. Remarque complémentaire : les interactions CH₃- π

Au cours de cette étude, les interactions CH₃- π entre les cycles de la quercétine et les chaînes latérales des résidus aliphatiques ont été examinées. Bien que la liaison hydrogène classique soit considérée comme l'une des plus importantes interactions en biochimie, les interactions entre un groupe C-H et un système aromatique π sont également répandues. Ces dernières contribuent notamment à la reconnaissance moléculaire et à la liaison du ligand à la protéine. Contrairement à la liaison hydrogène qui est une interaction forte de type électrostatique (-4,80 kcal.mol⁻¹ entre deux molécules d'eau), l'attraction dans les interactions C-H/ π est principalement due aux interactions de dispersion et est plus faible (-1,45 kcal.mol⁻¹ entre une molécule de méthane et de benzène) (Tsuzuki et al., 2006). Plusieurs études de mécanique quantiques ont quantifié l'interaction C-H/ π en fonction de la distance et de l'angle entre une molécule de méthane et une molécule de benzène (Figure IV.7 (b)) (Ringer et al., 2006 ; Tsuzuki et al., 2006 ; Morita et al., 2006).

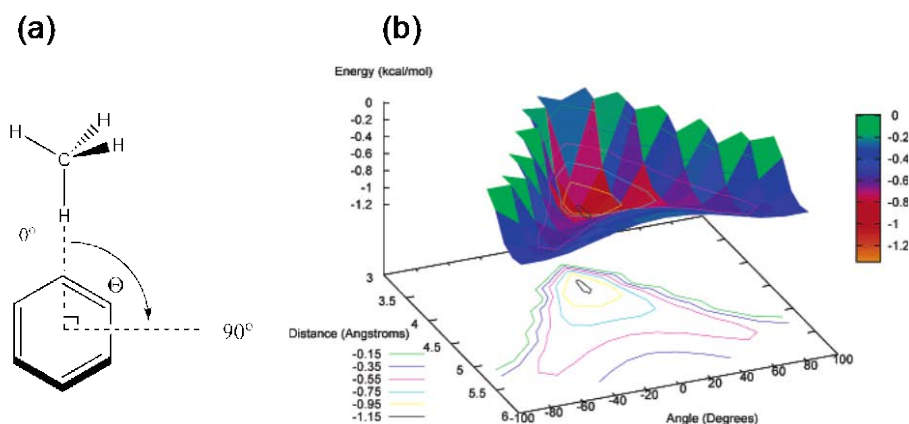


Figure IV.7. (a) Lorsque la liaison C-H du méthane est perpendiculaire au cycle benzénique, l'angle est par convention fixé à 0° tandis que l'angle est de 90° quand la liaison C-H est dans le plan du cycle aromatique ; (b) Surface d'énergie potentielle pour les interactions entre le méthane et le benzène ; la distance est mesurée entre le carbone du méthane et le centre du cycle aromatique et l'angle est mesuré entre la liaison C-H du méthane et la normale au cycle benzénique comme indiqué dans la figure 2 (a). (d'après Ringer et al., 2006)

L'interaction C-H/ π la plus forte correspond à la configuration dans laquelle la liaison C-H pointe vers le centre du noyau aromatique. Dans cette configuration, l'hydrogène partiellement positif a un meilleur accès pour interagir avec le système π . Lorsque la séparation inter-fragmentaire augmente, l'adaptation de l'angle décale la liaison C-H/ π du méthane de la perpendiculaire. La convention de notation des angles est indiquée dans la Figure IV.7 (a). A la distance inter-fragmentaire de 3.8 Å, un angle de 0° maximise l'interaction C-H/ π tandis qu'à une distance de 5.5 Å, l'interaction est plus intense à un angle de 90°. Dans un système méthane – naphthalène (Figure IV.8), le méthane est localisé au dessus de la frontière des deux cycles benzénique afin de maximiser l'interaction de dispersion (Morita et al., 2006).

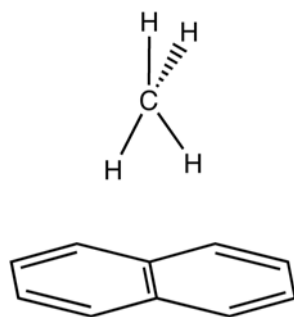


Figure IV.8. Représentation du système méthane – naphthalène. L'interaction CH - π est maximisée lorsque le carbone de la molécule de méthane se positionne au centre de la molécule de naphthalène.

Chapitre V. Étude par Modélisation
Moléculaire de l'Influence de la Mutation
des Résidus Hydrophobes Ile189 et Ile285
sur les Modes d'Interactions de la
Quercétine dans le Site Actif de la Lipase B
de *Candida antarctica*

1. Introduction

Dans l'étude précédente, les simulations de dynamique moléculaire des complexes associant la quercétine et la CALB-acétylée ont permis de mettre en évidence l'implication de deux résidus contraignant le flavonoïde dans une position inadéquate par rapport aux résidus catalytiques. Ces résidus sont les isoleucines 189 et 285. Ils sont localisés à l'entrée du site actif et leurs chaînes latérales se projettent dans la cavité catalytique (Figure V.1). Des mutations simples ont été réalisées *in silico*, qui consiste à remplacer ces deux résidus par les acides aminés aliphatiques, alanine et valine. L'objectif de ce travail est d'étudier l'impact de ces mutations sur :

- la structure et le comportement dynamique de la CALB
- le comportement dynamique de la quercétine dans le site actif des variants
- les interactions entre la quercétine et les variants
- l'affinité de liaison de la quercétine pour les sites actifs modifiés de la CALB

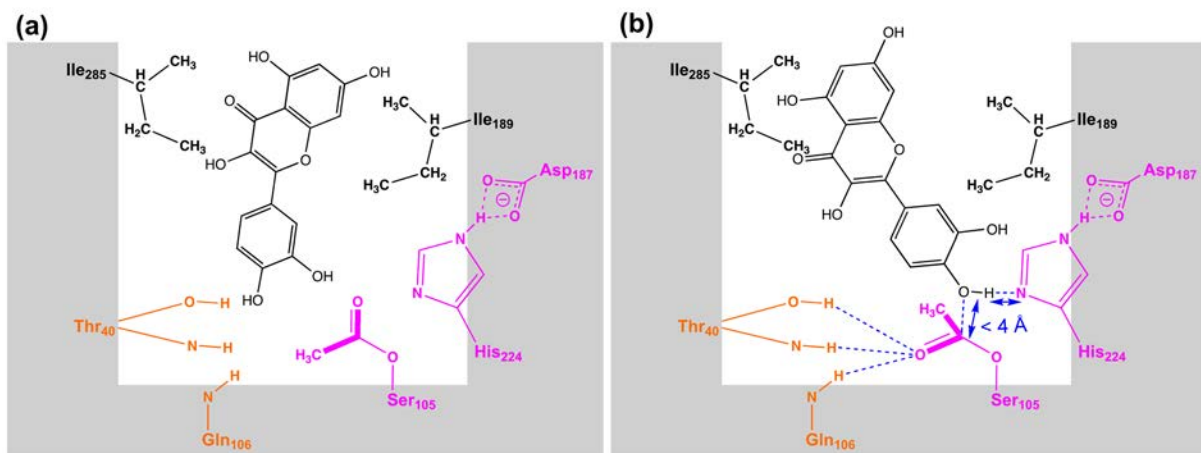


Figure V.1. (a) Schématisation simplifiée de l'encombrement stérique occasionné par les chaînes latérales des résidus Ile189 et Ile285 dans la cavité de la CALB. Les résidus de la triade catalytique sont indiqués en rose et ceux du trou oxyanionique sont en orange ; (b) schématisation du positionnement idéal de la quercétine pour l'attaque nucléophile de l'acyle-enzyme. Les critères géométriques fixés lors d'une étude précédente sont rappelés dans le schéma : la présence de liaisons hydrogène entre le carbonyle de l'acétate fixé sur la sérine catalytique (Ser105) et les résidus du trou oxyanionique, la présence de liaison hydrogène entre l'Asp187 et l'His224 et, le groupement hydroxyle du substrat localisé à une distance inférieure à 4 Å de l'acétate lié à la sérine catalytique (Ser105) et l'histidine (His224) catalytique.

Les quatre variants simples Ile189Ala, Ile189Val, Ile285Ala et Ile285Val ont été construits (Figure V.2) et soumis à des simulations de dynamique moléculaire de 20 ns. L'analyse de leur comportement conformationnel a été réalisée en exploitant le RMSD, le RMSF, le volume de la poche catalytique ainsi qu'en suivant les distances et les angles caractérisant les positions relatives des résidus de la triade catalytique. Pour chaque système muté, une image instantanée a été sélectionnée de façon à respecter les critères suivants : (i) une liaison hydrogène doit être établie entre les résidus Asp187 et His224 ; (ii) les liaisons amides des résidus Thr40 et Gln106 du trou oxyanionique doivent être orientées vers l'oxygène de la sérine catalytique. Un groupement acétate a ensuite été ajouté sur la sérine catalytique de manière à obtenir un modèle d'acyle-enzyme pouvant être utilisé comme cible

de docking. Une procédure similaire à celle présentée dans le chapitre précédent a été appliquée pour générer, classifier et sélectionner les complexes enzyme/substrats issus du docking. La stabilité temporelle et conformationnelle de la quercétine dans la cavité catalytique de CALB et ses variants a été examinée au cours de trajectoires de dynamique moléculaire de 10 ns. Finalement, l'affinité de la quercétine pour la CALB et ses variants a été évaluée par un calcul d'énergie libre de liaison en utilisant la méthode MM-PBSA.

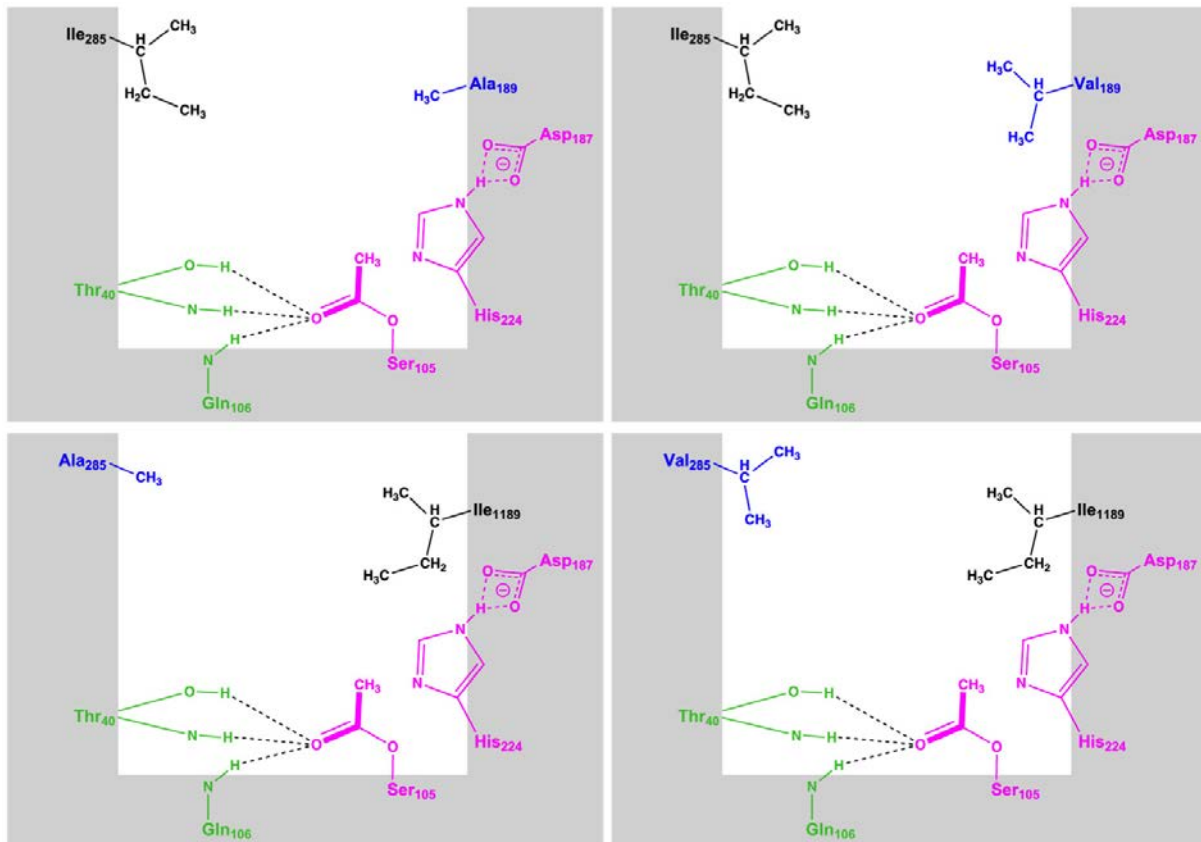


Figure V.2 Représentation schématisée des mutations effectuées sur les résidus Ile189 et Ile285. Ces acides aminés isoleucines ont été remplacés soit par une alanine soit par une valine pour raccourcir la longueur des chaînes latérales qui sont projetés dans la cavité de la CALB.

La rédaction de cette partie a été rédigée sous forme de projet de publication

2. Article

Interactions between quercetin and mutants of *Candida antarctica* lipase B : a theoretical study

Bidouil Christelle^a, Chebil Latifa^a, Maigret Bernard^b, Ghoul Mohamed^a, Engasser Jean-Marc^a & Humeau Catherine^{a*}

^a Laboratoire Ingénierie des Biomolécules, ENSAIA-INPL, Université Lorraine, 2 av. de la Forêt de Haye, 54500, Vandoeuvre-lès-Nancy, France.

^b Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), CNRS, Université Lorraine, BP 239, 54506, Vandoeuvre-lès-Nancy, France

*Corresponding author: Tel.: +33 3 83 59 57 84; fax: +33 3 83 59 57 78.

E-mail address: catherine.humeau@ensaia.inpl-nancy.fr

Abstract:

In order to assess the possibility of designing lipases that catalyze the acetylation of the aglycone flavonoid quercetin, molecular dynamics simulations were applied to *Candida antarctica* lipase B (CALB) mutants in which two isoleucines in the active site are replaced by the smaller hydrophobic alanine or valine. The four single variants mutants exhibit similar dynamic behavior as the CALB wild-type. Two of them have a slightly larger active site cavity. According to docking simulations, all the mutants can bind quercetin by its B cycle. In some of them binding can also occur by the flavonoid A or C cycle. From MM-PBSA free energy calculations, the lipase mutants were found to have a larger affinity for quercetin due to stronger electrostatic interactions. Molecular dynamics simulations, on a 10 ns time scale, of the quercetin-lipase complexes showed an increase of quercetin mobility in the CALB mutants and the formation of additional hydrogen bonds with the lipase cavity. Within the catalytic triad, the mutations on Ile285 have a failure effect on the establishment of the hydrogen bond between the catalytic aspartate and histidine. And, the variants do not maintain the necessary hydrogen bond between the serine-bound acetate and the oxyanion hole. Among all the mutants, three of quercetin hydroxyl groups can be stabilized near the acetate and histidine, but at distances unfavorable for efficient proton transfer.

Keywords: *Candida antarctica* lipase B, quercetin, mutation, molecular modeling, conformational study, binding free energy

1. INTRODUCTION

Candida antarctica lipase B (CALB) is one of the most used lipases to carry out either hydrolysis reactions or synthesis reactions in non-aqueous media on an industrial scale. This lipase displays high catalytic activity on a broad range of substrates, together with high chemo-, regio- and/or enantioselectivity [1].

In recent years, an increasing number of studies reported the use of lipases for flavonoid bioconversions [2]. These substrates are naturally occurring polyphenolic antioxidants and are abundantly found in fruits, vegetables, and beverages such as wine and tea. Flavonoids have a characteristic fifteen-carbon backbone structure consisting of two aromatic rings (A and B) joined by a three-carbon linked γ -pyrone ring (C). Many of them have glycoside moieties covalently bound to the flavan structure [3, 4]. Flavonoid compounds exhibit a wide range of biological, pharmacological and medicinal properties, such as anti-oxidant, anti-allergenic, anti-viral and cancer preventive effects, making them very attractive for use as functional ingredient in cosmetic, pharmaceutical and food products [5, 6]. Yet their poor solubility in both polar and non-polar media [7, 8] strongly restricts their incorporation in food and health formulations.

Higher solubilization of flavonoids can be achieved by using acylated derivatives. Acylation was also found to increase their stability and bioavailability and in some cases their biological activities [5, 9, 10]. Among enzymes that are able to catalyze that reaction, CALB was reported to be very active on glycosylated flavonoids. For example earlier a strong activity and regioselectivity of CALB was reported concerning the acylation of quercetin glycosylated derivatives as isoquercitrin and rutin. Surprisingly, no acylation activity was observed on the aglycon flavonoid quercetin [2, 11-13].

A computer-based molecular modeling of enzyme/substrate complexes involving CALB and flavonoid was recently proposed to explain the specificity and the regioselectivity of the lipase CALB towards polyphenolic compounds at the molecular level [14]. The models considered the binding pocket of the lipase that consists of an elliptical and steep funnel whose walls are predominantly covered by aliphatic hydrophobic amino acids [15]. The catalytic machinery, namely the triad Ser105...His224...Asp187, is located at the cavity bottom and participates to a more hydrophilic region [16]. The well-known mechanism of the lipase is based on the consecutive acylation and deacylation of the Ser105 residue. Two tetrahedral intermediates are produced throughout the catalytic pathway: the first one that results from the nucleophilic attack of the catalytic serine on the acyl donor substrate and the acyl-enzyme complex; and the second one intermediate results from the nucleophilic attack of the acyl acceptor substrate on the acyl-enzyme [15, 16]. A methodology combining docking and molecular dynamics simulations was applied to study the positions, the orientations and the interactions of the two substrates within the CALB active site together with the accessibility of the flavonoid hydroxyl groups to the catalytic residues.

Computational results showed that both glycosylated flavonoids isoquercitrin and rutin localized their aglycone part at the cavity entrance establishing hydrogen bonds and hydrophobic stabilizing interactions. The sugar moiety of the substrates was placed close to the cavity bottom. Models were shown to correctly predict the regioselectivity of the reaction. Only the primary 6''-OH of the isoquercitrin glucose and the secondary 4'''-OH of the rutin rhamnose were expected to

be acetylated, as these groups the only ones to stabilizing simultaneously near to the catalytic histidine and the acetate bound to the catalytic serine [14].

A similar methodology was applied to investigate the quercetin binding modes within CALB catalytic cavity. In that case, the presence of the flavonoid led to a bad orientation of the serine-bound acetate towards the oxyanion hole. Furthermore, quercetin hydroxyls groups remained far from the catalytic residues throughout molecular dynamics trajectories. This unfavorable positioning of both substrates within CALB catalytic pocket was partly attributed to the interactions between quercetin and hydrophobic residues. Especially, two residues were identified as potentially responsible for quercetin stabilization: the residue Ile189 that was shown to cause a steric hindrance of the catalytic histidine and the residue Ile285 that projected its side chain into the catalytic site [17].

The present work aims to further understand the role of these two hydrophobic residues in quercetin binding modes within CALB catalytic pocket. Four mutants of CALB were constructed *in silico* by replacing the residues Ile189 and Ile285 with smaller hydrophobic amino acids, alanine or valine, leading to the variants Ile189Ala, Ile189Val, Ile285Ala, and Ile285Val. The influence of these single-point mutations on the dynamic behavior of CALB and the conformations of enzyme/substrates complexes were studied. Particular attention was given to the effect of mutations on the localization, the orientations and the interactions of both substrates within CALB catalytic pocket. Results were expected to give general guidance for CALB redesign with the aim of developing a lipase that would be able to catalyze the regioselective acylation of aglycon flavonoids like quercetin.

2. COMPUTATIONAL METHODS

2.1. Construction of the *in silico* enzyme variants

Models of CALB mutants were constructed from wild type CALB X-ray structure with an ethylhexylphosphonate inhibitor (PDB ID: 1LBS, [16]), using the Discovery Studio 2.5. 'mutate tool' and replacing the residues Ile189 and/or Ile285 by new amino acids. Therefore, four simple variants (ile189ala, ile189val, ile285ala, ile285val) were built manually.

2.2. MD calculations

Structural study of CALB mutants was carried out through 20 ns trajectories. Complexes following docking simulations were subjected to 10 ns trajectories. All MD simulations were conducted using the CHARMM force field [18] and NAMD program [19]. The protonation state of protonable residues was adjusted at pH 7. MD simulations were performed in a truncated octahedral box with explicit TIP3 water [20]; one Na⁺ cation was added to neutralize the systems. Constant temperature and pressure were applied and were controlled *via* Langevin thermostat and Langevin piston, respectively. Electrostatic interactions were calculated by using the Particle-Mesh-Ewald method [21] with a non bonded cutoff of 10 Å. All bonds involving hydrogen atoms were constrained with the SHAKE algorithm [22]. A time step of 1 fs was used to integrate the equation of motion.

Systems were minimized in 6400 steps combining the conjugate gradient and the line minimizer methods implemented in NAMD. Then, a 500 ps equilibration step was applied, keeping the protein harmonically restrained. The force constant was initially fixed at 10 kcal.mol⁻¹.Å⁻² and was

progressively decreased ($0.5 \text{ kcal.mol}^{-1}.\text{Å}^{-2}/25 \text{ ps}$). Systems were equilibrated again for 500 ps by releasing all the restraints. Trajectories were conducted during 10 or 20 ns, and conformations were saved every 1 ps. Two or three simulations of each system were performed using different initial random velocity distributions. Dynamic behaviours were studied using 2000 frames from each simulation covering the production phase (one frame every 10 ps for CALB mutants and one frame every 5 ps for enzyme-substrates complexes).

Analysis of CALB mutants trajectories was carried out using the root mean square deviation (RMSD) of the $C\alpha$ atoms. Atomic positional fluctuations of $C\alpha$ atoms were calculated as well as a mass-weighted average value for each residue. These parameters are related to the B-factor through the following relationship (Eq. V.1). The simulated B-factors were calculated basing on the coordinates of the last 18 ns of trajectories. The volume of the catalytic cavity was estimated all over the 20 ns trajectories using the Fpocket tool [23], which is an open source pocket detection package based on Voronoi tessellation.

$$B_i = \frac{8\pi^2}{3} \langle \Delta r_i^2 \rangle \quad (\text{V.1})$$

2.3. Docking calculations

2.3.1. Selection and preparation of targets

The frames selected from molecular dynamics trajectories for further docking simulations were chosen according to the relative orientations of the catalytic triad residues. Distances and angles between Asp187 and His224, and between Ser105 and His224 were monitored to check the orientation of the residues and their interactions through hydrogen bonding. One snapshot was chosen among poses that presented a good orientation of the catalytic triad and was defined as target for further docking. According to the mechanism of acylation, the catalytic serine is firstly acylated by the acyl donor substrate, leading to the acyl-enzyme. In order to mime this intermediate, the Ser105 side chain hydroxyl was replaced by an acetate in the models [14, 24, 25]. Then, the energy of the systems was minimized executing 1000 Steepest Descent steps to adjust the position of the serine-bound acetate in the oxyanion hole. The resulting structures were taken as targets for the docking procedure.

2.3.2. Docking procedure

Docking of quercetin in acetyl-CALB targets was performed through Monte Carlo simulations, using LigandFit [26] (Discovery studio 2.5, Accelrys, Inc., San Diego, CA, USA). LigandFit docking procedure includes two major steps: (i) the specification of the receptor binding site thanks to a cavity detection algorithm that finds protein invaginations; (ii) docking of the ligand into the binding site using a Monte Carlo conformational search (15 000 trials) that generates poses consistent with the shape of the cavity. The top 20 best poses were allowed to be saved.

2.3.3. Scoring procedure

Poses were evaluated using a set of scoring functions including LigScore1, LigScore2, PLP1, PLP2, Jain and PMF, whereas the candidate ligands were prioritized according to the DockScore function. Ligand poses were clustered according to RMSD values between all pair-wise of conformations. Besides, poses were ranked considering a consensus of the six over mentioned scoring functions. The consensus score ranged from 0 to 6, the best score being 6. Among complexes exhibiting a consensus score of 5 or 6, one complex was selected in each cluster of poses in order to sample quercetin conformations within lipase catalytic pocket. All selected complexes were submitted to 10 ns molecular dynamics simulations.

2.4. MM/PBSA Calculations

For each system, the binding free energy between the ligand and the receptor was calculated using the MM/PBSA technic according to the following equation [27]:

$$\begin{aligned}\Delta G_{bind} &= \Delta G_{complex} - \Delta G_{protein} - \Delta G_{ligand} \\ &= \Delta E_{MM} + \Delta G_{PB} + \Delta G_{SASA} - T\Delta S\end{aligned}\quad (1)$$

where ΔE_{MM} is the molecular mechanic interaction energy between the receptor and the ligand; ΔG_{PB} and ΔG_{SASA} are respectively the polar and nonpolar contributions to desolvation upon ligand binding; and $-T\Delta S$ is the entropic contribution at temperature T.

The polar desolvation energy was calculated by solving the Poisson-Boltzmann (PB) equations using the CHARMM PBEQ module [28]. In PB calculations, the grid size used to solve the PB equation was 0.4 Å. Values for the solvent dielectric constant and the solute dielectric constant were set to 80 and 1, respectively. The nonpolar contribution was estimated to be proportional to the solvent accessible surface area (SASA) [29] with a solvent-probe radius of 1.4 Å: $\Delta G_{SA} = \gamma SASA + \beta$. γ and β values were set to 0.00542 kcal.mol⁻¹.Å⁻² and 0.92 kcal.mol⁻¹, respectively. Binding free energy calculations were conducted using CHARMM program. The average protein-ligand binding free energy was calculated including 500 snapshots taken from the 10 ns MD trajectories of complexes. The entropy contributions were not taken into account because calculations were computationally expensive.

3. RESULTS

3.1. Dynamics of wild-type CALB and its variants

In a first step the four simple variants (Ile189Ala, Ile189Val, Ile285Ala, Ile285Val) as well as CALB wild-type were subjected to 20 ns molecular dynamics simulations, in order to determine the influence of the mutations on the structural stability and flexibility of the enzyme, the cavity volume and on distances between residues of catalytic triad.

The enzyme stability was monitored using RMSD calculations applied to all C α atoms and referring to the starting structure. The obtained values are reported in Table V.1. Little variations were observed in RMSD trajectories for CALB or its variants throughout the simulations. Systems were shown to converge to stable states after the first 2 ns of simulation. The highest average value and the largest

variance were observed for the variant Ile189Val comparing with other systems. Indeed, RMSD evolution showed an increase after 10 ns of simulation, followed by a stabilization phase around $1.7 \pm 0.1 \text{ \AA}$, referring to the input structure.

Table V.1. Data relative to the structure and the flexibility of CALB wild type and its variants throughout molecular dynamics trajectories. RMSD mean values and standard deviation were calculated basing on C α atoms and referring to initial structures. Total B-factor was defined as the sum of independent B-factors calculated for each residue. Catalytic pocket volume was calculated for period of time of a 10 to 20 ns. The average value mentioned is calculated over the last 18 ns of each trajectory.

Parameters	CALB Wild-type	Ile189Ala	Ile189Val	Ile285Ala	Ile285Val
RMSD (Å)	1.3 ± 0.2	1.3 ± 0.2	1.5 ± 0.2	1.2 ± 0.1	1.3 ± 0.2
RMSD without residues 139 to 150 (Å)	1.2 ± 0.2	1.0 ± 0.1	1.2 ± 0.2	1.0 ± 0.1	1.1 ± 0.1
Total B-factor (Å 2)	5900	5200	5800	5400	5400
Catalytic pocket volume (Å 3)	210 ± 55^a	430 ± 90^a	290 ± 67^a	590 ± 120^a	190 ± 55^a

The regional motions in the protein structure were studied using B-factor calculations throughout MD trajectories (Figure V.3). Whatever the system, flexible regions were broadly similar and mainly concern a short α 5-helix flanked by short loops (residues 139 to 150) and another loop (residue 250 to 256) exposed to the solvent. However, different degrees of flexibility were noticed depending on the system under examination, the CALB wild-type followed by the variant Ile189Val and the variant Ile189Ala exhibiting the most and the less flexible structures, respectively (Table V.1). The flexibility of the region covering residues 139 to 150 was given specific consideration by calculating RMSD factor omitting this part of the systems (Table V.1). Results indicated a significant participation of the region under study in global RMSD values. In addition, the mobility of residues 187 to 189 was more pronounced for variants Ile285Ala and Ile285Val, suggesting some specific interactions between the regions that contain the residues Ile189 and Ile285

The volume of the catalytic cavity was determined both for CALB wild-type and variants aiming to assess space allowances for substrates. Comparing the average volumes calculated for CALB and its mutants (Table V.1) indicates a large increase of the cavity volume for the two mutants in which the isoleucine is replaced by alanine but no significant variation for the two mutants in which isoleucine is replaced by valine. From a dynamic point of view, variants that were obtained by substituting an Ile residue by an Ala residue were characterized by a large variability of volume with time. For instance, in the case of the variant Ile285Ala, a gradual increase of the volume was observed over the first 10 ns of the trajectory (Figure V.4). In the case of the variant Ile189Ala, the volume profile evolved cyclically throughout the trajectory between average values of 385 Å 3 and 465 Å 3 (Figure V.4).

Distances and angles between the aspartate, histidine and serine residues of the catalytic triad were also monitored throughout trajectories (Table V.2). As seen, similar average values and dynamics were obtained for the wild CALB and the four mutants. Interestingly, during the trajectories, large distance variations (between 2.4 and 6.7 Å) are observed between His224 and Asp187 for the five structures, reflecting the mobility of the catalytic histidine. This seems to indicate that hydrogen bond interactions are not maintained between the two residues all along the trajectories.

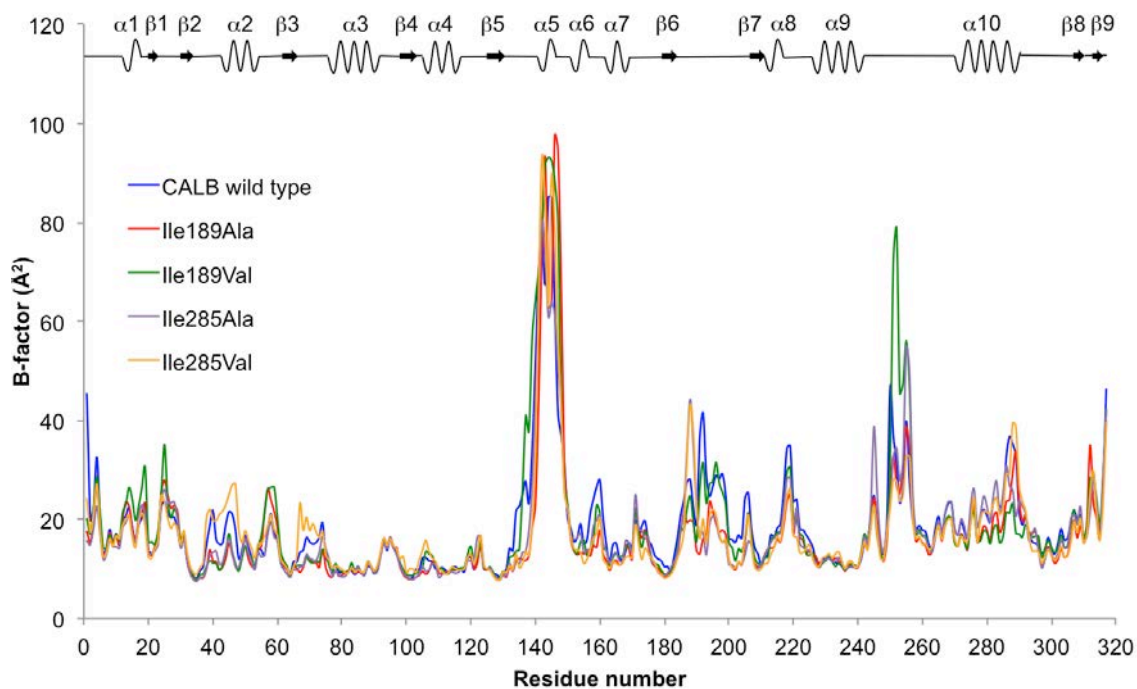


Figure V.3. B-factor values calculated from MD trajectories of CALB wild type and variants. Data relative to CALB wild-type, and the variants Ile189Ala, Ile189Val, Ile285Ala, Ile285Val are shown in blue, red, green, mauve and orange, respectively. The secondary structure of CALB is shown as a reference.

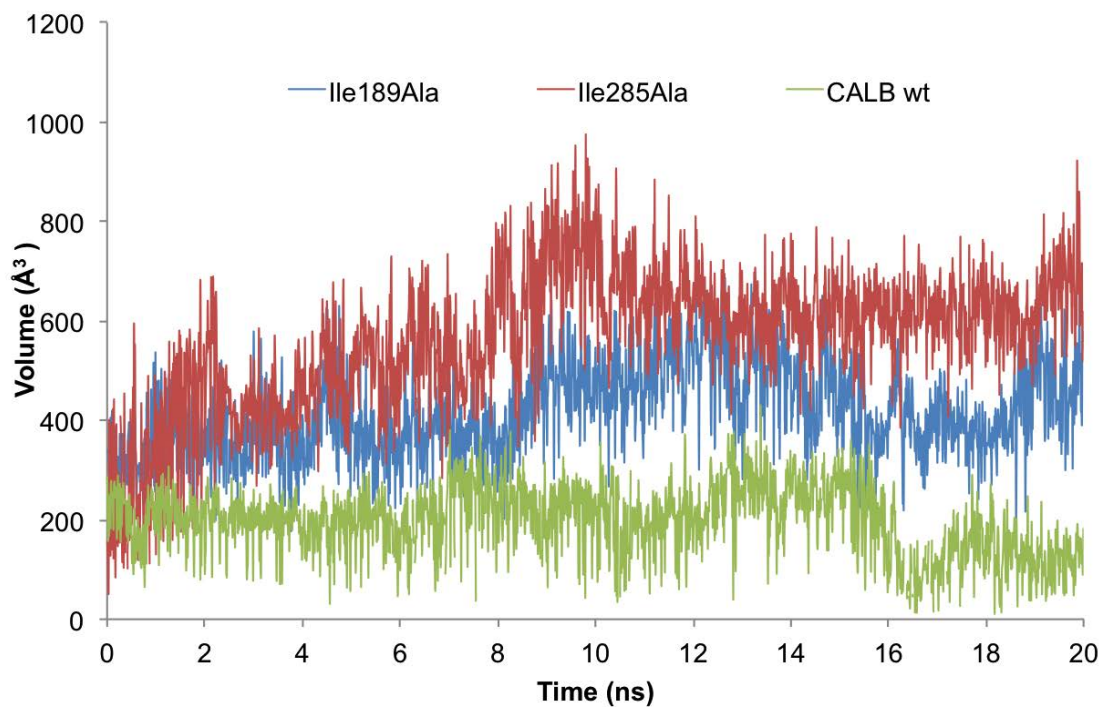


Figure V.4. Volume of catalytic pocket over the 20 ns simulation for CALB wild-type (green) and for variants Ile189Ala (blue) and Ile285Ala (red).

Table V.2. Average value and standard deviation of distances and angles between catalytic residues for CALB and its variants. Periods of time corresponding to specific phases of MD trajectories are specified as follows: ^a 2.0 to 5.5 ns, ^b 5.5 to 20 ns, ^c 2 to 17 ns, ^d 17 to 20 ns, ^e 2 to 10 ns, ^f 10 to 20 ns, ^g 2 to 14 ns, ^h 14 to 20 ns. When no period of time is specified, the average value calculated over the last 18 ns of each trajectory is mentioned.

Parameters	CALB Wild-type	Ile189Ala	Ile189Val	Ile285Ala	Ile285Val
Distance Asp187:OD1/His224:HD1 (Å)	2.4 ± 0.5 ^a 4.0 ± 1.4 ^b	2.5 ± 0.6 ^c 5.1 ± 0.9 ^d	2.7 ± 0.8 ^e 4.0 ± 0.5 ^f	2.9 ± 0.9 ^g 4.8 ± 1.3 ^h	2.9 ± 1.2 ^e 6.7 ± 1.0 ^f
Angle Asp187:OD1/His224:HD1/His224:ND1 (°)	135 ± 19 ^a 126 ± 29 ^b	144 ± 16 ^c 100 ± 17 ^d	145 ± 16 ^e 149 ± 12 ^f	137 ± 21 ^g 104 ± 24 ^h	135 ± 26 ^e 79 ± 16 ^f
Distance Asp187:OD2/His224:HD1 (Å)	2.5 ± 0.6 ^a 3.8 ± 1.2 ^b	2.6 ± 0.7 ^c 4.9 ± 1.0 ^d	2.7 ± 0.7 ^e 2.2 ± 0.5 ^f	2.4 ± 0.7 ^g 5.1 ± 1.1 ^h	2.9 ± 1.1 ^e 6.5 ± 1.2 ^f
Angle Asp187:OD2/His224:HD1/His224:ND1 (°)	140 ± 21 ^a 119 ± 27 ^b	141 ± 17 ^c 101 ± 18 ^d	145 ± 16 ^e 154 ± 13 ^f	139 ± 19 ^g 102 ± 23 ^h	134 ± 27 ^e 80 ± 16 ^f
Distance Ser105:OG/His224:NE2 (Å)	3.6 ± 1.3	3.6 ± 0.6	3.9 ± 0.5	3.7 ± 0.6	3.5 ± 0.5

3.2. Docking results

3.2.1. Selection of frames

For each system under study, a frame was chosen from molecular dynamics trajectories and designed as docking target for further simulations. Snapshots that matched the following criteria were retained (Figure V.5 (a)): (i) a hydrogen bond interaction between Asp187 oxygen atom and His224 HD1 hydrogen atom must be present (ii) hydrogen bonds between the catalytic serine OG atom and the HN atom of oxyanion hole residues. In the case of mutants Ile189Ala, Ile189Val, Ile285Ala and Ile285Val, frames were selected respectively during 2 to 17 ns, 10 to 20 ns, 2 to 14 ns and 2 to 10 ns time periods. Ser105 hydroxyl group was then substituted for acetate in order to generate acyl-enzyme models for all variants (Figure V.5 (b)). The structures thus obtained were optimized and then submitted to docking procedure.

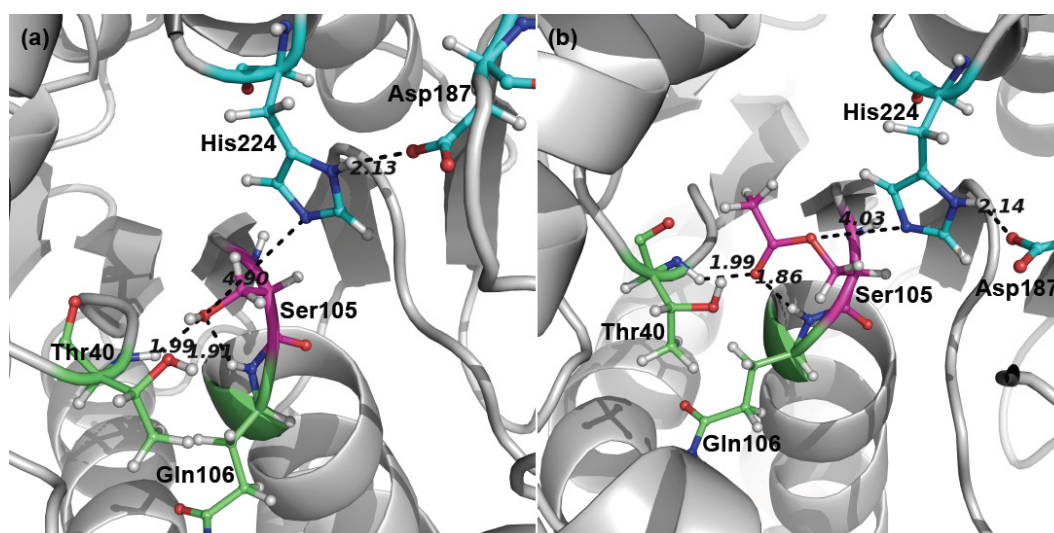


Figure V.5. Snapshot from CALB molecular dynamics trajectory showing catalytic distances (a). Optimized structure of acyl-enzyme system (b). Acyl-enzyme involving Ile285Ala mutant was taken as an

example. Hydrogen bonds between Asp187 (cyan) and His224 (cyan), and between Ser105 (pink) and oxyanion hole residue (green) are shown.

3.2.2. Selection of complex

Docking simulations were undertaken in order to determine the influence of CALB mutation on quercetin binding modes and interactions within enzyme-substrates complexes. The twenty poses retained from the scoring procedure were clustered according to their structural similarities (RMSD between all pairs of poses) and classified based on their consensus scoring. All systems taken together, poses fell into four major conformational families: family CA for which quercetin cycle A points towards the cavity bottom; family CB characterized in that quercetin cycle B points towards the cavity bottom; families CCA and CCB in which the three aromatic cycles of quercetin are spread across the cavity, with either the A-ring or the B-ring facing the helix $\alpha 5$. Major conformational families observed for each variant are reported in Table V.3. Complexes that were chosen for further analyses are given in the right section of the Table 3.

Table V.3. Major conformational families obtained from docking for CALB wild type and its variants. Selected complexes are given in the right section. Nomenclature was established as follows: CALB wt and q refer to CALB wild-type and quercetin, respectively.

CALB Variants	Family	Selected complexes		
		Name	Family	Consensus scoring
CALB wild-type*	CA and CB	CALB wt – qA	CA	5
		CALB wt – qB	CB	6
Ile189Ala	CA, CB and CCB	Ile189Ala – qA	CA	5
		Ile189Ala – qB	CB	5
Ile189Val	CB	Ile189Val – qB	CB	6
Ile285Ala	CA and CB	Ile285Ala – qA	CA	6
Ile285Val	CB and CCA	Ile285Val – qB	CB	5
		Ile285Val – qCA	CCA	6

* Results from [17].

Several observations can be made about quercetin positioning within the catalytic pocket. The first one is that the CB orientation was found in all systems, whatever the target. On the contrary the CA orientation was not observed in Ile189Val and Ile285Val mutants. A second point is that two new orientations (CCA and CCB) were found for the mutants Ile189Ala and Ile285Val. By contrast, only vertical positions were adopted by quercetin in CALB wild type cavity.

3.3. Dynamics of enzyme-substrates complexes

In order to complete the static representation of the systems, enzyme-substrates complexes were submitted to molecular dynamicssimulationsaiming to explore their conformational space.

3.3.1. Global structure behaviour

RMSD of both C α atoms and quercetin atoms were followed with time (Figure V.6). Two reference systems were also analysed for assessing the impact of mutations on quercetin dynamic behaviour: CALB wt–qA and CALB wt–qB. Whatever the complex, low RMSD values were obtained throughout MD simulations indicating that no distortion occurred in protein structure in the presence of

quercetin (Figure V.6 (a) and (c)). RMSD values for CALB variants were globally lower and more fluctuant than those of CALB wild-type. As shown in Figure V.6 (b) and (d), little spatial fluctuations were observed for quercetin in complexes CALB wt-qA, CALB wt-qB, Ile285Val-qCA and Ile189A-qA. In other complexes (Ile189Ala-qB, Ile189Val-qB and Ile285Val-qB), quercetin moved in such a way that a shift in orientation was made possible. In the complex Ile285Ala-qA, quercetin adopted a stable position during almost the entire trajectory, then get out of the cavity. This came firstly with a disruption of H-bond interactions between quercetin hydroxyl groups and Asp134 residue and secondly with a water molecules entry into the cavity. An overall observation of the results shows that quercetin movement within the catalytic pocket did not affect the protein structure. Indeed, no correlation was found between variations of both quercetin RMSD and protein RMSD.

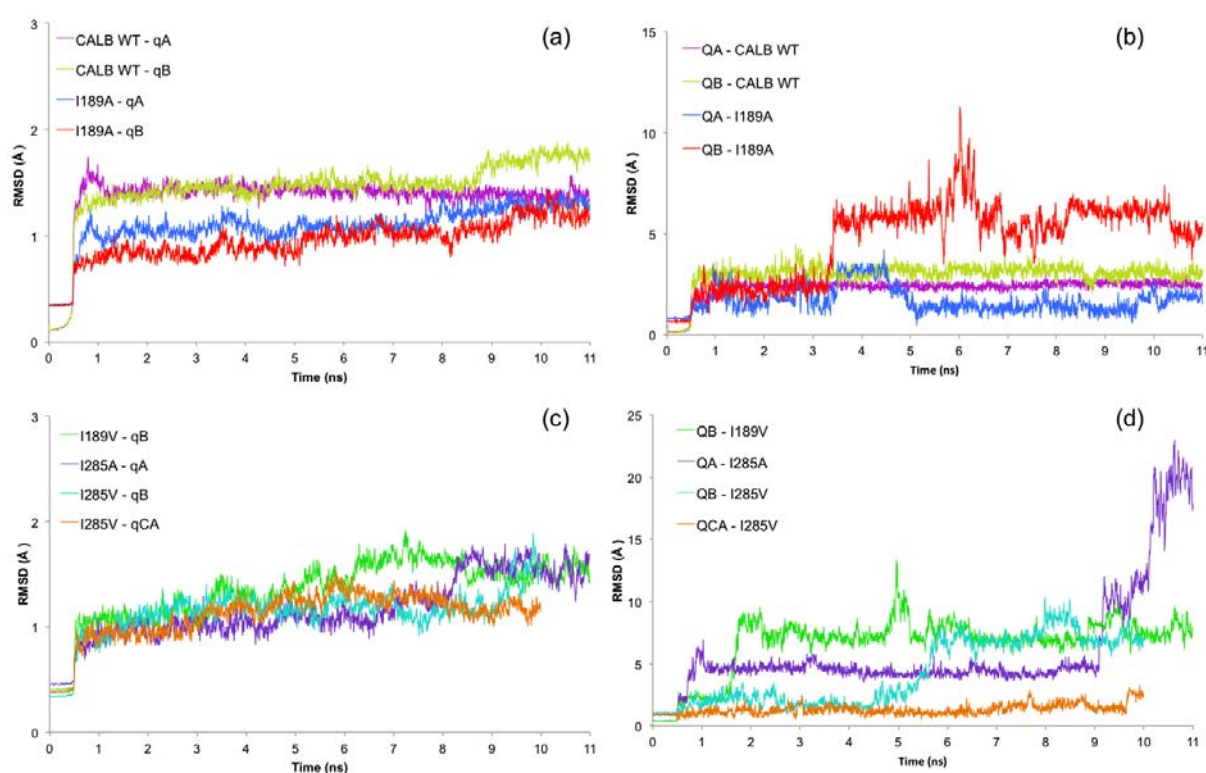


Figure V.6. Evolution of C α atoms RMSD with time for complexes involving quercetin and CALB wild type or CALB variants (a and c). Evolution of quercetin RMSD in the same systems (b and d). RMSD values were determined during the equilibration (0 to 1 ns) and the production (1 to 11 ns) phases, referring to the molecular dynamics input structures.

The flexibility of enzyme-substrates complexes was evaluated through the measurement of the B-factor of the C α (Figure V.7). Whatever the complex under study, two flexible regions were identified: helix α 5 flanked by short loops (residues 139 to 150) and a loop (residues 250 to 256) exposed to solvent. These moving structural elements had already been identified in systems without quercetin and are usually found in CALB dynamic studies (Figure V.3). Globally, variants exhibited similar or even higher flexibility than CALB wild-type. A particular behaviour was observed with the complex Ile285Val-qCA. An additional flexible region that includes the catalytic histidine (residues 223 to 225) was identified, indicating a distortion of the catalytic structure.

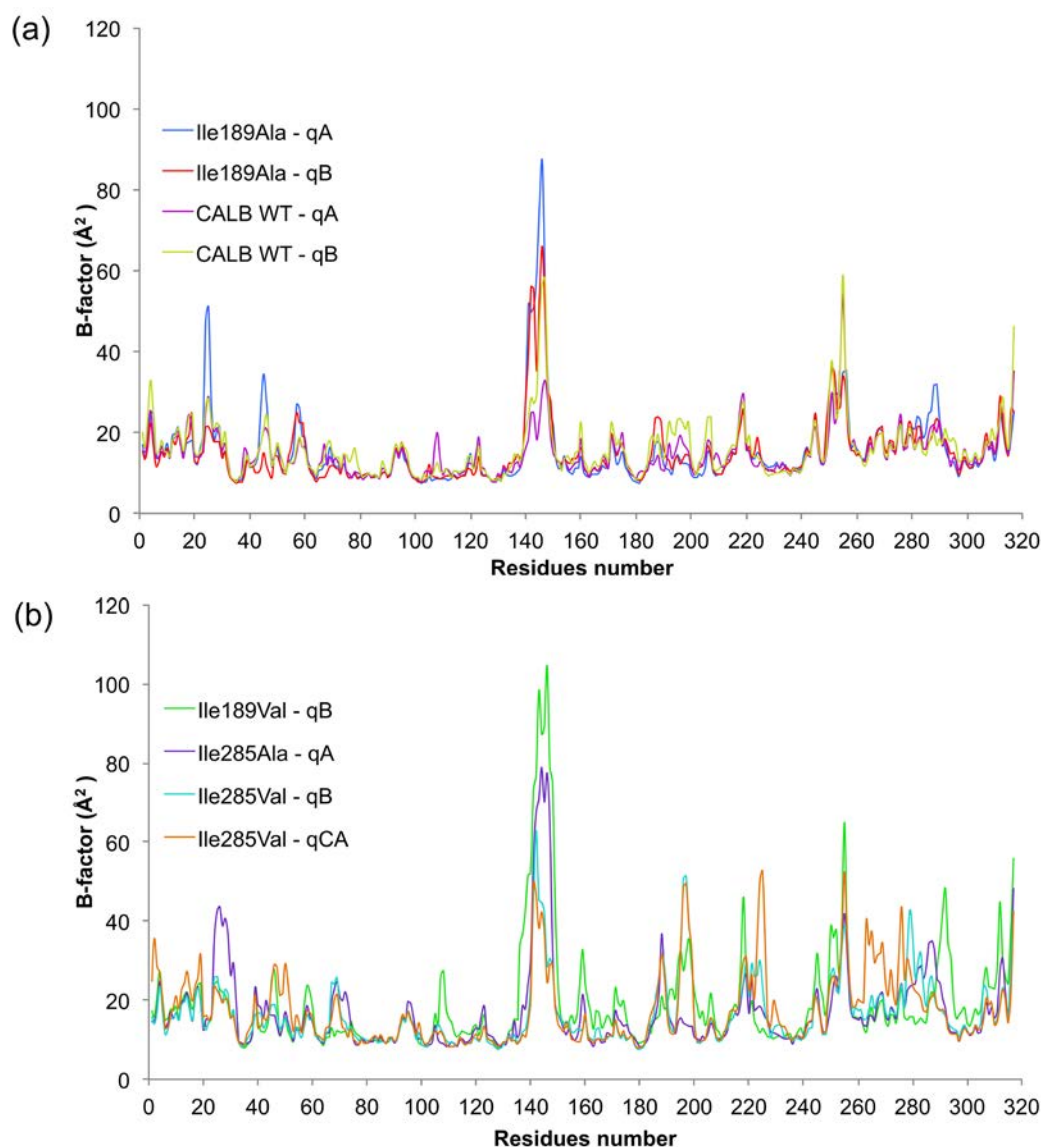


Figure V.7. Atomic fluctuations (B-Factor) of residues, based on C α atoms and calculated over 10 ns molecular dynamics trajectories. The different complexes under study involved either CALB variants or CALB wild and are given in the legend.

3.3.2. Study of quercetin interaction

One important objective of this study was to investigate the influence of CALB mutations on quercetin binding modes and interactions. Hydrogen bond interactions are often reported as important structural conditions stabilizing substrates. The occurrence of hydrogen bond interactions between quercetin hydroxyl groups and the residues forming the cavity was calculated throughout molecular dynamics trajectories (Table V.4). Thr40, Asp134, Thr138, Ser150, Gln157, Asp187 and Glu188 were identified as main residues interacting with quercetin (Figure V.8). CALB wt-qA, CALB wt-qB and Ile285Val-qCA complexes were characterized by strong anchor and stabilization of quercetin thanks to hydrogen bond interaction with atoms OD1 or OD2 of the residue Asp134 located at the cavity bottom. This result agrees with the low RMSD values observed for quercetin as indicated just before (Figure V.6 (b) and (d)). A similar interaction was observed in the complex Ile285Val-qB. However, in addition to previous observations, quercetin inserted between the catalytic residues Asp187 and His224, its 7-OH

group interacting with the residue Asp187 through a hydrogen bond. This repositioning of the substrate came with a rapid increase of quercetin RMSD, as shown in Figure V.6 (d). In the case of the complex Ile189Ala-qA, in addition to an interaction with the residue Asp134, small displacements of quercetin were observed that were attributed to an intermittent interaction of either its 3'-OH or its 4'-OH group with the residue Glu188 (Figure V.6 (b)). As suggested by quercetin RMSD values, complexes Ile189Ala-qB and Ile189Val-qB were characterized by wider movements of quercetin, which B-ring hydroxyl groups interacted alternatively with the residue Thr40 (Figure V.6 (b) and (d)). The weakest interactions were obtained in the complex Ile285A-qA. In that case, quercetin was observed to go out of the cavity after 9 ns of simulation.

Table V.4. Occurrence ratio as well as average value and standard deviation of distance and angle of H-bond interactions between CALB residues and quercetin hydroxyl groups throughout molecular dynamics trajectories (2000 frames).

Complexes	Donor	Acceptor	Occupied (%)	Average distance (Å)	Average angle (°)
CALB wt - qA	7-OH	Asp134:OD1	38	2.0 ± 0.2	145 ± 10
	7-OH	Asp134:OD2	58	2.0 ± 0.2	145 ± 10
	5-OH	Thr40:OG1	11	2.4 ± 0.3	140 ± 10
	3-OH	Thr40:O	82	2.0 ± 0.2	145 ± 10
CALB wt - qB	3-OH	Thr40:O	48	2.0 ± 0.2	135 ± 10
	4'-OH	Asp134:OD1	100	2.0 ± 0.2	155 ± 10
	Gln157:NE2	4'-OH	86	2.1 ± 0.3	150 ± 15
Ile189Ala - qA	5-OH	Ace:O	41	2.2 ± 0.3	130 ± 10
	Thr40:OG1	5-OH	28	2.3 ± 0.3	135 ± 10
	7-OH	Asp134:OD1	30	2.1 ± 0.3	160 ± 10
	7-OH	Asp134:OD2	75	2.1 ± 0.2	160 ± 10
	3'-OH	Glu188:OE1	18	2.1 ± 0.2	150 ± 15
	3'-OH	Glu188:OE2	17	2.1 ± 0.2	150 ± 15
	4'-OH	Glu188:OE1	36	2.1 ± 0.3	150 ± 15
Ile189Ala - qB	4'-OH	Glu188:OE2	36	2.1 ± 0.3	150 ± 15
	7-OH	Glu188:OE1	33	2.1 ± 0.3	150 ± 15
	7-OH	Glu188:OE2	32	2.1 ± 0.3	150 ± 15
	3'-OH	Ace:O	27	2.0 ± 0.2	140 ± 15
	3'-OH	Thr40:O	24	2.2 ± 0.3	150 ± 15
	4'-OH	Ace:O	23	2.0 ± 0.2	150 ± 15
Ile189Val - qB	4'-OH	Thr40:O	52	2.0 ± 0.2	155 ± 15
	7-OH	Glu188:OE2	25	2.1 ± 0.3	150 ± 15
	7-OH	Glu188:OE1	26	2.1 ± 0.3	150 ± 15
	3'-OH	Thr40:O	9	2.0 ± 0.2	155 ± 15
Ile285Ala - qA	4'-OH	Thr40:O	67	2.0 ± 0.2	150 ± 15
	Ser150:OG	4-O	24	2.1 ± 0.3	130 ± 10
	Thr40:OG1	7-OH	15	2.2 ± 0.3	130 ± 10
	3'-OH	Glu188:OE1	10	2.1 ± 0.3	150 ± 15
	3'-OH	Glu188:OE2	10	2.2 ± 0.4	150 ± 15
Ile285Val - qB	4'-OH	Glu188:OE1	12	2.1 ± 0.3	150 ± 15
	4'-OH	Glu188:OE2	16	2.2 ± 0.3	150 ± 15
	7-OH	Asp187:OD2	21	2.1 ± 0.3	150 ± 15
	7-OH	Asp187:OD1	18	2.1 ± 0.3	150 ± 15
	3'-OH	Asp134:OD1	91	2.1 ± 0.2	150 ± 15
	3'-OH	Asp134:OD2	84	2.2 ± 0.3	145 ± 15
Ile285Val - qCA	4'-OH	Asp134:OD1	83	2.1 ± 0.3	145 ± 15
	4'-OH	Asp134:OD2	75	2.1 ± 0.3	150 ± 15
	3-OH	Asp134:OD1	13	2.2 ± 0.3	135 ± 10
	3-OH	Asp134:OD2	43	2.1 ± 0.2	140 ± 10

Thr40:OG1	5-OH	20	2.2 ± 0.3	130 ± 5
3'-OH	Asp134:OD1	72	2.1 ± 0.2	150 ± 15
3'-OH	Asp134:OD2	23	2.1 ± 0.3	150 ± 15
Thr138:OG1	4'-OH	74	2.2 ± 0.2	130 ± 10

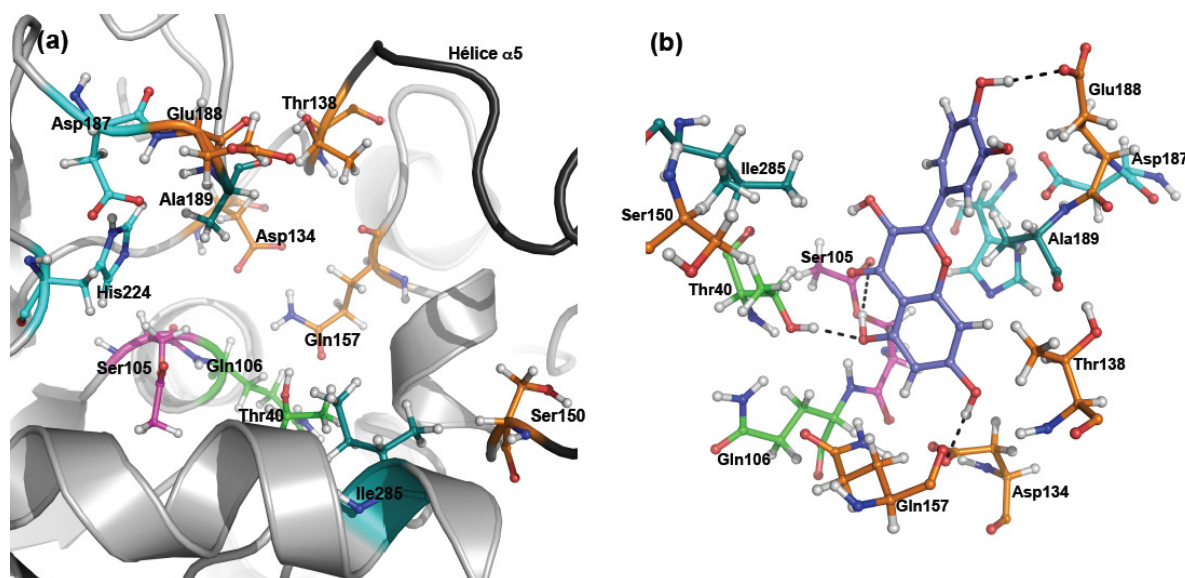


Figure V.8. (a) Major residues establishing hydrogen bond interactions with quercetin hydroxyl groups. (b) Hydrogen bond network stabilizing quercetin within Ile189Ala mutant (Ile189Ala-qA complex).

According to the previous observations, mutation of Ile189 and Ile285 residues forming the catalytic cavity resulted in an increase of quercetin mobility. CALB wild-type complexes were characterized by that quercetin can move restrictively due to strong and stable H-bond interactions and steric hindrance caused by residues Ile189 and Ile285. Replacing the residue Ile189 by an alanine or a valine amino acid led to shorter interaction periods thus to higher mobility of quercetin. This residue was identified as responsible for a steric hindrance that contributes to quercetin positioning within the catalytic cavity. Mutation of Ile285 residue decreased the interaction between quercetin and Thr40 residue while the strong interaction between Asp134 residue and embedded hydroxyl groups of quercetin was maintained during almost all of the trajectories.

3.3.3. Analysis of catalytic distances and interactions

For each enzyme-substrates complex, hydrogen bond pattern stabilizing, on one hand the catalytic triad, and on the other hand the serine-bound acetate within the oxyanion hole was analyzed. A conformation was considered catalytically efficient provided that the catalytic residues Asp187 and His224 establish a hydrogen bond interaction and the serine-bound acetate interacts with the oxyanion hole residues Thr40 and Gln106 [14, 24, 25]. Hydrogen bond interactions were considered as existing if the donor-acceptor(H) distance was below or equal to 3.0 Å while the donor-acceptor(H)-acceptor angle was greater or equal to 120° [30]. Results were expressed as the percentage of snapshots respecting above mentioned conditions (also called occurrence ratio) (Table V.5).

Most enzyme-substrates complexes obtained with CALB variants were characterized by low occurrence ratios whether for catalytic triad interactions or for serine-bound acetate stabilization. Two

exceptions could be the complexes Ile189Ala–qA and Ile189Val–qB, that provided similar or even better results than those obtained with CALB wild-type. These results suggest that only a small proportion of frames adopted an ideal conformation with regard to catalytic mechanism.

Table V.5. Percentage of frames exhibiting a good orientation of the serine-bound acetate towards the oxyanion hole and a correct conformation of the catalytic triad. The last column indicates the percentage of frames for which both above mentioned conditions were satisfied. Calculation was carried out on 2000 frames.

Systems	Correct conformation of the catalytic triad	Correct orientation of the serine-bound acetate	Correct orientation of both the catalytic triad and the serine-bound acetate
CALB wild type	49 %	56 %	26 %
CALB wild type – qA	99 %	0 %	0 %
CALB wild type – qB	94 %	27 %	21 %
Ile189Ala – qA	49 %	58 %	27 %
Ile189Ala – qB	51 %	4 %	1 %
Ile189Val – qB	86 %	33 %	32 %
Ile285Ala – qA	18 %	15 %	4 %
Ile285Val – qB	0 %	13 %	0 %
Ile285Val – qCA	0 %	11 %	0 %

Another condition that must be satisfied so that reaction could occur is the proximity of quercetin hydroxyl groups to both catalytic histidine and serine-bound acetate. Two main distances require special attention: distance separating His224:NE2 atom and acyl acceptor group and distance between serine-bound acetate (Ace:C) and acyl acceptor group. These two distances must be less or equal to 4 Å. Complexes satisfying all previously mentioned criteria can be designated as productive complexes, i.e. complexes that lead to the formation of an ester [14, 31, 32]. The proximity of buried quercetin hydroxyl groups to both catalytic histidine and serine-bound acetate was studied for the complexes CALB wt–qB, Ile189Ala–qA and Ile189Val–qB that exhibit at least 20% of frames satisfying a correct orientation of the catalytic triad (Table V.6).

Table V.6. Average and standard deviation distances between quercetin hydroxyl groups and both the catalytic residue His224 and the serine-bound acetate. Only periods during which buried hydroxyl groups came near the atoms Ace:C and His224:NE2 were taken into account.

Complexes	Buried OH-groups	Distance from Ace:C (Å)	Distance from His224:NE2 (Å)	Time (ns)
CALB wt – qB	3-OH	3.6 ± 0.3	4.6 ± 0.9	1 to 2.7 6.3 to 6.8
	3'-OH	4.4 ± 0.3	5.1 ± 0.5	1 to 2.7 6.3 to 6.8
Ile189Ala – qA	5-OH	3.8 ± 0.3	4.3 ± 0.6	6 to 8.5
Ile189Val – qB	3'-OH	4.0 ± 0.9	4.2 ± 0.9	3.9 to 5 6 to 8.7

Whatever the complex, quercetin 3-OH, 5-OH and 3'-OH hydroxyl groups were close to both His224 residue and serine-bound acetate. However, average distances exceeded the maximal values admitted for a proton transfer towards the catalytic histidine and a nucleophilic attack of the serine-bound acetate. In light of these results, simple mutation of residues Ile189 and Ile285 appear to be promising but insufficient to allow quercetin acylation.

3.3.4. Binding free energies predicted by MM-PBSA method

Binding free energy within enzyme-substrates complexes was calculated by the MM-PBSA method aiming to determine CALB affinity towards quercetin. Comparison was made with CALB variants in order to estimate the impact of mutations on substrate binding. Binding affinities and corresponding energy contributions calculated for the different systems under study are summarized in Table V.7. One first observation is that mutations led to an increase of CALB affinity towards quercetin. A second one is that Ile285Ala and Ile285Val variants exhibited higher affinity for quercetin than Ile189-based variants. Moreover, the orientation of quercetin within the catalytic cavity was shown to affect ΔG_{bind} . Indeed, for a given variant, binding free energy differed depending on quercetin orientation. No correlation could be found between quercetin orientation, target structure and binding energy. For instance, complexes involving Ile189-based variants gave the highest affinity when quercetin adopted a CA orientation, whereas the CB orientation was the most favourable in the case of CALB wild-type.

A closer look at individual energy terms (ΔE_{vdW} , ΔE_{elec} , ΔG_{PB} , ΔG_{SASA}) showed that both van der Waals (ΔE_{vdW}) and electrostatic (ΔE_{elec}) contributions were predominant in quercetin binding, whatever the target. Complexes involving quercetin in the CA orientation were characterized in that electrostatic contribution was higher than van der Waals contribution. A contrary tendency was observed for complexes including quercetin in CB or CCA orientation. In such cases, the main driving force for substrate binding appeared to be van der Waals interactions. Furthermore, in the case of CALB variants, electrostatic contribution was found to increase with hydrogen bond number and occurrence. For instance, electrostatic interactions were lower in the complexes Ile189Ala – qB and Ile189Val – qB than in other complexes due to fewer hydrogen bond interactions. Moreover, in such systems, no hydrogen bond was established with the residue Asp134. In any case, the electrostatic contribution appeared to be fairly representative of quercetin binding affinity.

Table V.7. Average values and standard deviation of binding free energies and individual energy terms. All energies were expressed in kcal.mol⁻¹.

Complexes	ΔE_{vdW}	ΔE_{elec}	ΔG_{PB}	ΔG_{SA}	ΔG_{bind}
CALB WT – qA	-8.7 ± 1.8	-11.7 ± 2.7	4.9 ± 0.9	-4.4 ± 0.5	-21 ± 3
CALB WT – qB	-28.5 ± 1.8	-9.8 ± 1.9	1.1 ± 0.5	-3.5 ± 0.0	-37 ± 3
Ile189Ala – qA	-23.8 ± 1.8	-25.5 ± 3.2	5.4 ± 1.5	-4.2 ± 0.0	-42 ± 3
Ile189Ala – qB	-20.7 ± 1.2	-9.0 ± 3.2	3.0 ± 1.7	-3.7 ± 0.0	-36 ± 4
Ile189Val – qB	-23.0 ± 1.4	-14.8 ± 0.9	3.6 ± 1.1	-3.7 ± 0.0	-38 ± 3
Ile285Ala – qA ^b	-19.6 ± 1.6	-23.2 ± 1.4	5.2 ± 1.4	-3.5 ± 0.7	-41 ± 2
Ile285Val – qB	-20.7 ± 1.5	-28.0 ± 0.6	3.6 ± 1.2	-4.1 ± 0.0	-49 ± 3
Ile285Val – qCA	-29.3 ± 1.7	-21.4 ± 1.8	4.4 ± 0.1	-4.4 ± 0.1	-51 ± 3

^a ΔE_{vdW} , van der Waals contribution; ΔE_{elec} , electrostatic contribution; ΔG_{PB} , polar contribution of desolvation; ΔG_{SA} , nonpolar contribution of desolvation

^b ΔG_{bind} calculated on 7 ns molecular dynamics trajectories; snapshots where quercetin came out of the catalytic cavity were not taken into account

4. DISCUSSION

In this study, computational modeling was performed on CALB mutants in order first to better understand, on a molecular basis, the absence of CALB activity for the acetylation of quercetin;

second to evaluate the perspectives of a rational enzyme engineering extending the catalytic activity of the lipase towards the aglycone flavonoid. Mutations were targeted at the two Ile189 and Ile285 residues that were previously identified as potentially contributing to the inadequate insertion of quercetin in the lipase cavity due to steric hindrances. The replacement of the isoleucine by either one of the smaller alanine or valine residues was expected to widen the lipase pocket, and thus facilitate the substrate access towards the catalytic triad.

The wild type CALB and the four constructed mutants were first submitted to a structural analysis on a 20 ns time scale, which is more extensive than in previous studies of CALB mutants stability based on short trajectories (< 1ns) or energy minimization protocols [31-38]. All the systems were found to converge to stable states after the first 2 ns of simulation. Globally the four single mutants exhibit similar dynamic behavior as the CALB wild-type. Mutations were also shown to have no effect on the protein flexibility evaluated by the B-factor of the C α . In all systems, the most mobile elements were identified as the helix α 5 flanked by short loops and another loop that is exposed to the solvent, including residues 250 to 256. These results are in agreement with previous studies reporting the dynamic behavior of CALB in explicit water [39, 40]. Conformational study of mutants suggested that replacing the residues Ile189 and Ile285 did not cause any disruption in the protein structure.

A major finding of the structural study is the important increase of the cavity volume induced by the replacement of either one or the other isoleucine residues by alanine. Replacement by valine, on the contrary, has no significant effect. Similar CALB mutations to expand the cavity were suggested in previous reports [34, 35, 41], but none of them evaluated the magnitude of the expected increase.

The study also analyzed the influence of mutations on the quercetin binding mode in the lipase cavity. For the wild type CALB, the docking simulation identified both CA and CB complexes in which the quercetin A or B cycle points towards the cavity bottom. In all the four mutants quercetin is found to bind through its B cycle. But it binds through its A cycle only in the variants where alanine replaces isoleucine. Interestingly two new orientations (CCA and CCB) were found for the mutants Ile189Ala and Ile285Val.

Another computational prediction of the simulation is the higher quercetin affinity, evaluated as the MM-PBSA free energy, for the mutants as compared with its affinity for the wild CALB. Decreasing the side chain of residues Ile189 or Ile285 allows quercetin to insert into the lipase active site and to establish several hydrogen bonds with amino acids of the active site. The Asp134 was determined as a particularly strong electrostatic anchor for quercetin hydroxyl groups.

Molecular dynamics simulations, on a 10 ns time scale, of the selected quercetin-lipase complexes demonstrated that in CALB wild-type quercetin movement is restricted by H-bond interactions and steric hindrance caused by residues Ile189 and Ile285. Replacing the residues by an alanine or a valine led to shorter interaction periods, thus to higher quercetin mobility. The improved flavonoid mobility results in stronger electrostatic interactions between its hydroxyl groups and the Thr40, Asp134, Thr138, Ser150, Gln157, Asp187 and Glu188 residues on the active site. The interactions are established all over the trajectory for mutants with modified Ile285, but restricted to shorter time periods for the variants having Ile189 replaced by valine or alanine. These stronger

electrostatic interactions may arise from additional hydrogen bond formation, potentially important for substrate stabilization in the enzyme cavity.

At the level of the catalytic triad, the CALB mutations were not found to have a beneficial effect on the formation of the hydrogen bond network necessary for catalytic activity. Hydrogen bonds between the catalytic residues Asp187 and His224 are established during half of the trajectory for the Ile189 modified variants, and a shorter period for the Ile285 mutants. In addition, the mutants do not maintain the hydrogen bond interactions between the serine-bound acetate and the oxyanion hole residues Thr40 and Gln106 which are not adequately established in the wild type CALB. The simulation also calculated the dynamics of the quercetin distances with both the catalytic histidine and serine-bound acetate. Among all complexes, quercetin 3-OH, 5-OH and 3'-OH groups come close to both His224 residue and serine-bound acetate. However, average distances exceeded the maximal values admitted for a proton transfer towards the catalytic histidine and a nucleophilic attack of the serine-bound acetate.

Several previous reports produced Ile189 and Ile285 mutants of CALB through targeted mutagenesis experiments. The catalytic activity of the mutant Ile189Ala was studied in acylation reactions involving different linear secondary alcohols and methyl propanoate as acyl donor, in solid-gas bioreactor [35]. Results showed that the mutant was less effective than CALB wild-type to catalyze the reactions. Similar results were obtained in transacylation reactions of acrylates, suggesting that the mutant Ile189Ala was deleterious for substrate binding [34]. Other mutants like Ile189Val and Ile285Trp were also studied. Transacylation reaction of acrylate substrates was not improved, whatever the mutant. Linder et al. [41] conducted a theoretical study about Diels-Alder reaction catalyzed by CALB. Once again, the mutants Ile189Ala and Ile285Ala were suggested to be insufficient to improve the reaction and were shown to be less effective than double variants as Ser105Ala/Ile189Ala or Ser105Ala/Ile285Ala.

In conclusion, the present simulation that compares the quercetin lipase interactions in CALB and its mutants with modified Ile189 and Ile285 confirms the steric hindrance effect of these two residues within the lipase cavity. Replacement of isoleucine by smaller side chain residues can indeed expand the cavity volume and increase the quercetin mobility, thus favoring the establishment of hydrogen bonds that stabilize the quercetin in the vicinity of the catalytic triad. However, these mutations alone are not capable to maintain the necessary hydrogen bonds network within the catalytic triad and to position quercetin close enough to both the catalytic histidine and the acetate substrate. In the future, additional mutations will be investigated within the perspective of designing CALB variants catalytically active for aglycone flavonoid acylation.

5. REFERENCES

1. Anderson, E.M., K.M. Larsson, and O. Kirk, *One biocatalyst - Many applications: the use of Candida antarctica B-lipase in organic synthesis*. Biocatalysis and Biotransformation, 1998. **16**: p. 181-204.
2. Viskupicova, J., M. Ondrejovic, and T. Maliar, *Enzyme-mediated preparation of flavonoid esters and their applications*. Biochemistry : p. 263-286.
3. Havsteen, B.H., *The biochemistry and the medical significance of flavonoids*. Pharmacology & Therapeutics, 2002. **96**: p. 67-202.
4. Boudet, A.M., *Evolution and current status of research in phenolic compounds*. Phytochemistry, 2007. **68**: p. 2722-2735.
5. Boots, A.W., G.R.M.M. Haenen, and A. Bast, *Health effects of quercetin: from antioxidant to nutraceutical*. European Journal of Pharmacology, 2008. **585**: p. 325-337.
6. Chen, C., J. Zhou, and C. Ji, *Quercetin: a potential drug to reverse multidrug resistance*. Life Sciences, 2010. **87**(11-12): p. 333-338.
7. Rothwell, J.A., A.J. Day, and M.R.A. Morgan, *Experimental determination of octanol-water partition coefficients of quercetin and related flavonoids*. Journal of Agricultural and Food Chemistry, 2005. **53**: p. 4355-4360.
8. Chebil, L., et al., *Solubility of flavonoids in organic solvents*. Journal of Chemical and Engineering Data, 2007. **52**: p. 1552-1556.
9. Mattarei, A., et al., *Regioselective O-derivatization of quercetin via ester intermediates. An improved synthesis of rhamnetin and development of a new mitochondriotropic derivative*. Molecules, 2010. **15**: p. 4722-4736.
10. Viskupicova, J., M. Ondrejovic, and E. Sturdik, *The potential and practical applications of acylated flavonoids*. Pharmazie, 2009. **64**: p. 355-360.
11. Ardhaoui, M., et al., *Acylation of natural flavonoids using lipase of Candida antarctica as biocatalyst*. Journal of Molecular Catalysis B: Enzymatic, 2004. **29**: p. 63-67.
12. Chebil, L., et al., *Enzymatic acylation of flavonoids : effect of the nature of the substrate, origin of lipase and operating conditions on conversion yield and regioselectivity*. Journal of Agricultural and Food Chemistry, 2007. **53**(23): p. 9496-9502.
13. Stevenson, D.E., et al., *Direct acylation of flavonoid glycosides with phenolic acids catalyzed by Candida antarctica lipase B (Novozyme 435)*. Enzyme and Microbial Technology, 2006. **39**: p. 1236-1241.
14. De Oliveira, E.B., et al., *A molecular modelling study to rationalize the regioselectivity in acylation of flavonoid glycosides catalyzed by Candida antarctica lipase B*. Journal of Molecular Catalysis B: Enzymatic, 2009. **59**: p. 96-105.
15. Pleiss, J., M. Fischer, and R.D. Schmid, *Anatomy of lipase binding sites : the scissile fatty acid binding site*. Chemistry and Physics of Lipids, 1998. **93**: p. 67-80.
16. Uppenberg, J., et al., *The sequence, crystal structure determination and refinement of two forms of lipase B from Candida antarctica*. Structure, 1994. **2**: p. 293-308.
17. Bidouil, C., et al., *Combined docking and molecular dynamics simulations to enlighten the capacity of Pseudomonas cepacia and Candida antarctica lipases to catalyze quercetin acetylation*. Journal of Biotechnology, 2011. **156**: p. 203-210.
18. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy minimization and dynamics calculations*. Journal of Computational Chemistry, 1983. **4**: p. 187-217.
19. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD*. Journal of Computational Chemistry, 2005. **26**(16): p. 1781-1802.

20. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water*. Journal of Chemical Physics, 1983. **79**(2): p. 926-935.
21. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems*. Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
22. Ryckaert, J.P., G. Ciccotti, and J.C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n -alkanes*. Journal of Computational Physics, 1977. **23**(3): p. 327-341.
23. Le Guilloux, V., P. Schmidtke, and P. Tuffery, *Fpocket: an open source platform for ligand pocket detection*. BMC Bioinformatics, 2009. **10**: p. art. 168.
24. Botta, B., et al., *Lipase-catalyzed regioselective acylation of resorcin[4]arenes*. Journal of Molecular Catalysis B: Enzymatic, 2002. **16**: p. 241-247.
25. Palocci, C., et al., *An approach to address *Candida rugosa* lipase regioselectivity in the acylation reactions of tritylated glucosides*. Journal of Biotechnology, 2007. **128**: p. 908-918.
26. Venkatachalam, C.M., et al., *LigandFit : a novel method for the shape-directed rapid docking of ligands to protein active sites*. Journal of Molecular Graphics and Modelling, 2003. **21**: p. 289-307.
27. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models*. Accounts of Chemical Research, 2000. **33**(12): p. 889-897.
28. Im, W., D. Beglov, and B. Roux, *Continuum solvation model: Computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation*. Computer Physics Communications, 1998. **111**(1-3): p. 59-75.
29. Sitkoff, D., K.A. Sharp, and B. Honig, *Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models*. The Journal of Physical Chemistry, 1994. **98**(7): p. 1978-1988.
30. Panigrahi, S.K. and G.R. Desiraju, *Strong and weak hydrogen bonds in the protein-ligand interface*. Proteins: Structure, Function, and Bioinformatics, 2007. **67**(1): p. 128-141.
31. Branneby, C., et al., *Aldol additions with mutant lipase : analysis by experiments and theoretical calculations*. Journal of Molecular Catalysis B: Enzymatic, 2004. **31**: p. 123-128.
32. Svedendahl, M., et al., *Direct epoxidation in *Candida antarctica* lipase B studied by experiment and theory*. ChemBioChem, 2008. **9**: p. 2443-2451.
33. Rotticci, D., et al., *Improved Enantioselectivity of a Lipase by Rational Protein Engineering*. ChemBioChem, 2001. **2**(10): p. 766-770.
34. Syrén, P.O., et al., *Increased activity of enzymatic transacylation of acrylates through rational design of lipases*. Journal of Molecular Catalysis B: Enzymatic, 2010. **65**(1-4): p. 3-10.
35. Marton, Z., et al., *Mutations in the stereospecificity pocket and at the entrance of the active site of *Candida antarctica* lipase B enhancing enzyme enantioselectivity*. Journal of Molecular Catalysis B: Enzymatic, 2010. **65**(1-4): p. 11-17.
36. Liu, D., et al., *Rational Design of *Pseudozyma antarctica* lipase B yielding a general esterification catalyst*. ChemBioChem, 2010. **11**: p. 789-795.
37. Magnusson, A., *Rational redesign of *Candida antarctica* lipase B*, in *Department of Biochemistry 2005*, Royal Institute of Technology: Stockholm. p. 54.
38. Branneby, C., et al., *Carbon-carbon bonds by hydrolytic enzymes*. Journal of American Chemical Society, 2003. **125**: p. 874-875.
39. Trodler, P. and J. Pleiss, *Modeling structure and flexibility of *Candida antarctica* lipase B in organic solvents*. BMC Structural Biology, 2008. **8**: p. art. 9.

40. Skjøt, M., et al., *Understanding the plasticity of the α/β -hydrolase fold: lid swapping on the *Candida antarctica* lipase B results in chimeras with interesting biocatalytic properties*. ChemBioChem, 2009. **10**: p. 520-527.
41. Linder, M., et al., *Computational design of a lipase for catalysis of the Diels-Alder reaction*. Journal of Molecular Modeling, 2011. **17**(4): p. 833-849.

3. Contribution de l'article

Dans cette partie du travail, une étude structurale de mutants de la CALB a été entreprise pour étudier l'effet de mutations ponctuelles des résidus Ile189 et Ile285 d'une part sur le comportement dynamique de la lipase, par des simulations de dynamique moléculaire, et d'autre part sur les modes d'interactions de la quercétine placée dans la cavité catalytique.

Les résultats de l'étude structurale effectuée sur les mutants Ile189Ala, Ile189Val, Ile285Ala et Ile285Val ont montré que :

- Le remplacement des résidus 189 et 285 par des acides aminés de type valine et alanine présentant une chaîne latérale plus courte n'entraîne pas de réarrangement conformationnel majeur de la structure tridimensionnelle de la CALB. Le RMSD et le RMSF des variants calculés sur des trajectoires de 20 ns sont similaires à ceux obtenus pour l'enzyme sauvage.
- Une augmentation du volume du site actif de la CALB est observée lorsque les résidus Ile189 et Ile285 sont substitués par une alanine ce qui est conforme au résultat escompté ; tandis que le remplacement des résidus isoleucine par une valine n'entraîne pas d'augmentation significative du volume du site actif.

A l'issue des trajectoires de dynamique moléculaire, un protocole de docking combiné à des simulations de dynamique moléculaire a permis d'étudier les modes d'interaction de la quercétine dans la cavité catalytique des mutants de la CALB. Quatre familles d'orientation de la quercétine ont pu être mises en évidence. Les familles CA et CB correspondent à une insertion de la quercétine présentant respectivement le cycle aromatique A et le cycle aromatique B vers le cœur de l'enzyme. Les familles CCA et CCB, quant à elles, désignent un positionnement horizontal de la quercétine dans le site actif de la lipase, avec soit le cycle A soit le cycle B du flavonoïde pointant vers l'hélice $\alpha 5$. Le comportement dynamique du flavonoïde aglycone dans les sites actifs des mutants de la CALB, les interactions avec les acides aminés environnants, l'affinité de liaison du substrat pour les poches catalytiques ont été évalués au cours des trajectoires de dynamique moléculaire. Les principaux constats sont :

- La présence de la quercétine n'affecte pas la structure protéique des mutants de la CALB.
- La mutation des résidus 189 ou 285 conduit à une mobilité accrue de la quercétine dans la cavité catalytique de la lipase mutée par rapport à celle observée dans le cas de l'enzyme sauvage. Ce résultat s'explique notamment par un élargissement de la cavité, induite par les mutations.
- Des interactions de type électrostatique ont été identifiées entre les groupements hydroxyles de la quercétine et les résidus Thr40, Asp134, Thr138, Ser150, Gln157, Asp187 et Glu188. Ces interactions sont observées pendant de courtes périodes dans le cas des mutants Ile189Ala et Ile189Val tandis qu'elles sont maintenues tout au long des trajectoires de dynamique moléculaire dans le cas des mutants impliquant le résidu Ile285.
- Les variants Ile189Ala, Ile189Val, Ile285Ala et Ile285Val présentent une affinité de liaison pour la quercétine supérieure à celle de l'enzyme sauvage. Cela peut s'expliquer par une

augmentation des interactions électrostatiques de type liaisons hydrogène entre le flavonoïde et les résidus du site actif dans le cas des mutants.

Par ailleurs, la présence d'une liaison hydrogène entre les résidus His224 et Asp187 de la triade catalytique, le positionnement de l'acétate lié à la sérine catalytique dans le trou oxyanionique et la proximité des groupements hydroxyles de la quercétine vis-à-vis des résidus catalytiques ont été examinés pour les différents complexes variant-quercétine. Ces différents éléments constituent des conditions essentielles pour l'obtention d'intermédiaires productifs, c'est-à-dire susceptibles de conduire à la formation d'un produit. Les résultats sont les suivants :

- La liaison hydrogène établie entre les résidus de la triade catalytique est maintenue pendant au moins la moitié des trajectoires, excepté pour les mutants impliquant le résidu Ile285.
- L'établissement de liaisons hydrogène entre le carbonyle de l'acétate et les résidus Thr40 et Gln106 du trou oxyanionique est effectif sur 86 % de la trajectoire pour le complexe Ile189Val-qB et 50 % de la trajectoire pour le complexe Ile189Ala-qA et Ile189Ala-qB. Quant aux autres complexes, ces liaisons hydrogène ne sont globalement pas établies.
- La combinaison des deux critères précédents (liaisons hydrogène entre les résidus de la triade catalytique et, entre l'acétate et les résidus du trou oxyanionique) indique que le mutant Ile189Ala-qA ainsi que le mutant Ile189Val-qB présentent entre 25 % et 30 % d'images instantanées avec une orientation correcte de la machinerie catalytique.
- La proximité des hydroxyles de la quercétine vis-à-vis des résidus catalytiques a été étudiée pour les complexes Ile189Ala-qA et Ile189Val-qB. Les distances mesurées entre les groupements hydroxyles et les résidus catalytiques sont en moyenne supérieures au critère de distance de 4 Å permettant d'envisager, d'une part un transfert de proton de la quercétine à l'histidine catalytique, d'autre part une attaque nucléophile de la quercétine sur l'acyl-enzyme.

D'après les résultats issus de la modélisation moléculaire, les mutants construits semblent insuffisants pour améliorer le potentiel catalytique de la CALB dans la réaction d'acylation de la quercétine. Une interaction électrostatique importante entre la quercétine et le résidu Asp134 a été mise en évidence dans plupart des complexes.

4. Étude complémentaire: étude structural des variants double Ile189Ala-Ile285Ala, Ile189Ala-Ile285Val, Ile189Val-Ile285Ala, Ile189Val-Ile285Val de la lipase B de *Candida antarctica*

Une étude complémentaire a été menée sur des mutations doubles impliquant les résidus Ile189 et Ile285 dans le but de vérifier leur impact sur la structure de la lipase B de *Candida antarctica*. Pour ce faire, des trajectoires de dynamique moléculaire de 20 ns ont été calculées pour les quatre variants Ile189Ala-Ile285Ala, Ile189Ala-Ile285Val, Ile189Val-Ile285Ala et Ile189Val-Ile285Val.

4.1. Stabilité structurale des mutants doubles

La stabilité des systèmes a été déterminée par l'analyse du RMSD des trajectoires obtenues à partir des simulations de DM. Il a été calculé en se référant à la structure initiale des systèmes. Les résultats sont représentés graphiquement dans la Figure V.9 et les valeurs moyennes sont données dans le Tableau V.1. Les valeurs moyennes et les écart-types des mutants Ile189Ala-Ile285Ala et Ile189Ala-Ile285Val sont similaires à celles obtenues pour la CALB sauvage, tandis que le variant Ile189Val-Ile285Ala se montre plus stable. Le suivi du RMSD au cours du temps (Figure V.9) indique que les trois systèmes Ile189Ala-Ile285Ala, Ile189Ala-Ile285Val et Ile189Val-Ile285Ala atteignent l'équilibre après 2 ns de trajectoire. De ce fait, l'analyse des trajectoires a été entreprise sur les 18 dernières nanosecondes des simulations. Pour le mutant Ile189Val-Ile285Val, le RMSD augmente progressivement au cours des 10 premières nanosecondes du système simulé jusqu'à atteindre un plateau dont la valeur moyenne est de 3.3 Å.

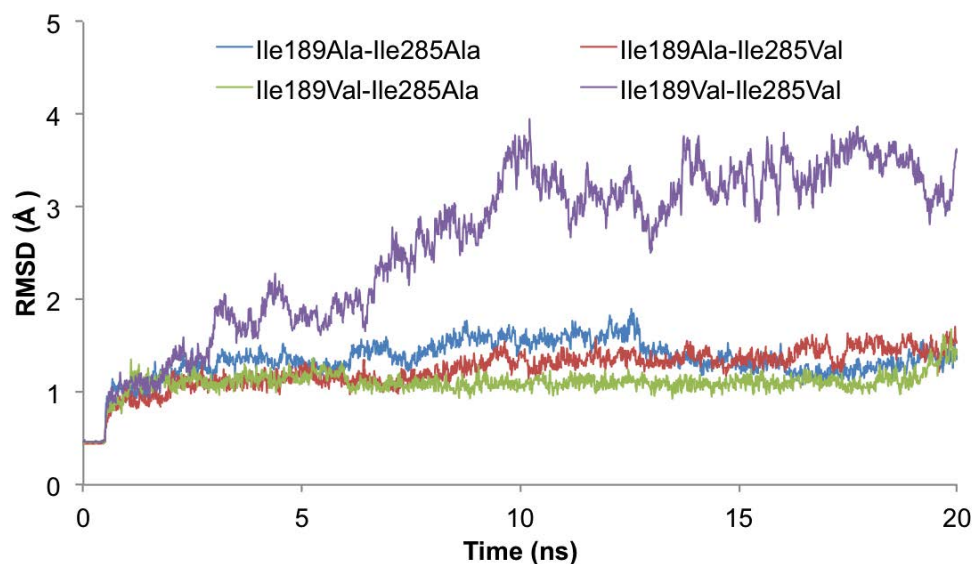


Figure V.9. Suivi du RMSD des C α au cours des trajectoires de DM (20 ns) pour les quatres variants doubles de la CALB.

Tableau V.1. Différents paramètres contrôlés au cours des trajectoires de DM des doubles mutants de la CALB. Les valeurs de la CALB sauvage sont également données à titre comparatif. Les périodes de temps sur les trajectoires de DM sont spécifiés comme suit : ^a2 à 5.5 ns, ^b5.5 à 20 ns, ^c(2 à 3 ns – 16 à 20 ns), ^d3 à 16 ns, ^e2 à 10 ns, ^f10 à 20 ns, ^g(2 à 5 ns - 15 à 20 ns), ^h5 à 15 ns. Les moyennes et les écart-types qui ne sont pas assignés à une période de temps sont calculés sur 18 ns de trajectoire, soit de 2 à 20 ns.

Paramètres	CALB sauvage	Ile189Ala-Ile285Ala	Ile189Ala-Ile285Val	Ile189Val – Ile285Ala	Ile189Val – Ile285Val
RMSD (Å)	1.32 ± 0.16	1.40 ± 0.15	1.30 ± 0.15	1.12 ± 0.09	3.29 ± 0.26
RMSD sans les résidus 139 à 150 (Å)	1.16 ± 0.15	1.06 ± 0.09	1.04 ± 0.07	0.97 ± 0.10	3.05 ± 0.32
B-factor total (Å ²)	5900	5900	5400	4500	10500
Distance Asp187:OD1/His224:HD1 (Å)	2.38 ± 0.49 ^d 3.97 ± 1.35 ^b	2.77 ± 0.65 ^c 4.34 ± 1.15 ^d	2.42 ± 0.54 ^e 5.92 ± 0.93 ^f	2.86 ± 0.90	2.37 ± 0.49 ^g 5.91 ± 0.77 ^h
Angle Asp187:OD1/His224:HD1/His224:ND1 (°)	134.69 ± 19.16 ^a 125.75 ± 28.77 ^b	126.88 ± 14.31 ^c 108.07 ± 28.18 ^d	140.79 ± 17.55 ^e 85.11 ± 13.79 ^f	153.41 ± 13.22	144.21 ± 16.77 ^g 69.16 ± 17.84 ^h
Distance Asp187:OD2/His224:HD1 (Å)	2.49 ± 0.61 ^a 3.80 ± 1.24 ^b	2.44 ± 0.73 ^c 4.48 ± 1.44 ^d	2.49 ± 0.62 ^e 5.86 ± 0.86 ^f	3.07 ± 0.91	2.57 ± 0.75 ^g 5.66 ± 0.59 ^h
Angle Asp187:OD2/His224:HD1/His224:ND1 (°)	140.71 ± 20.28 ^a 118.47 ± 27.31 ^b	150.92 ± 17.91 ^c 101.73 ± 31.78 ^d	138.93 ± 17.07 ^e 85.30 ± 12.86 ^f	154.32 ± 13.59	141.10 ± 18.06 ^g 50.03 ± 18.29 ^h
Distance Ser105:OG/His224:NE2 (Å)	3.55 ± 1.25	3.60 ± 0.55	3.32 ± 0.43	3.51 ± 0.55	6.61 ± 1.25

La flexibilité des doubles mutants a été évaluée par le B-factor dont le graphe est montré dans la Figure V.10. Les fluctuations du RMSD montrent une grande mobilité des atomes C α pour le mutant Ile189Val-Ile285Val. Le B-factor permet de préciser les résidus impliqués dans cette flexibilité importante. Les résidus 268 à 287 appartenant à l'hélice α 10 qui surplombe la cavité catalytique se déstructurent rapidement au cours de la trajectoire (Figure V.11). Les trois autres doubles mutants sont plus stables ; l'hélice α 5 flanquée de 2 courtes boucles (résidus 139 à 150) constitue l'élément de structure secondaire le plus mobile. Comme il a été observé précédemment, ce résultat est similaire aux observations faites pour la CALB sauvage et les mutants simples. Le RMSD a été calculé en absence des résidus 139 à 150 (hélice α 5) ; les valeurs sont données dans le Tableau V.1. La diminution des valeurs moyennes permet d'évaluer la contribution de l'hélice α 5 à la flexibilité globale des systèmes.

En outre, il est également observé que le B-factor total des variants Ile189Ala-Ile285Ala, Ile189Ala-Ile285Val et Ile189Val-Ile285Ala est inférieur à celui de la CALB sauvage. Cette faible flexibilité globale indique que les doubles mutations n'induisent pas de changement conformationnel important dans la structure de CALB, à l'exception de la mutation Ile189Val-Ile285Val.

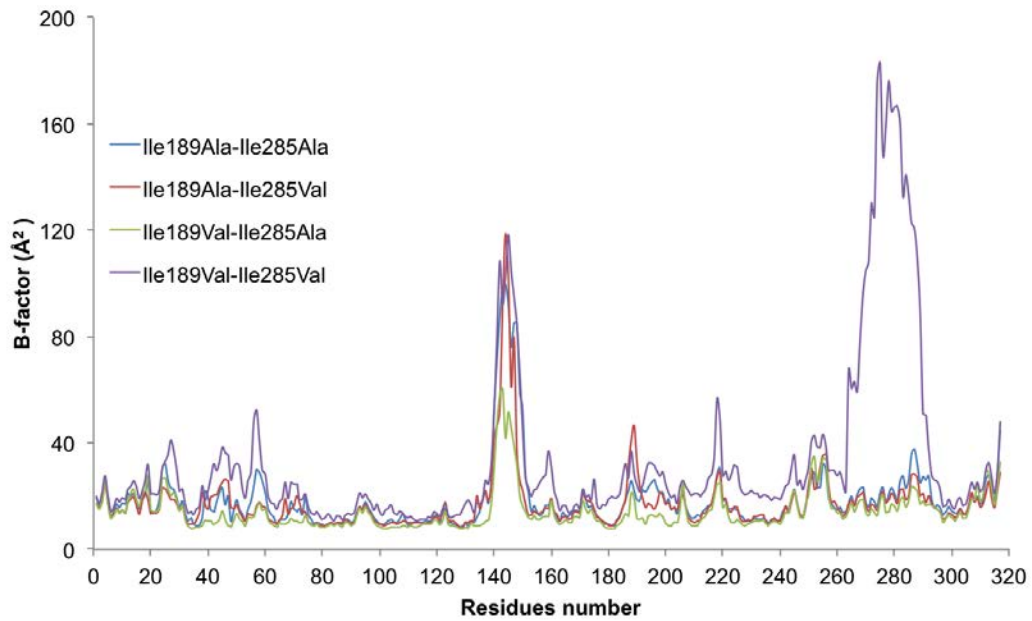


Figure V.10. B-factor des atomes C α calculé sur les 18 dernières nanosecondes de trajectoire pour les doubles variants de la CALB

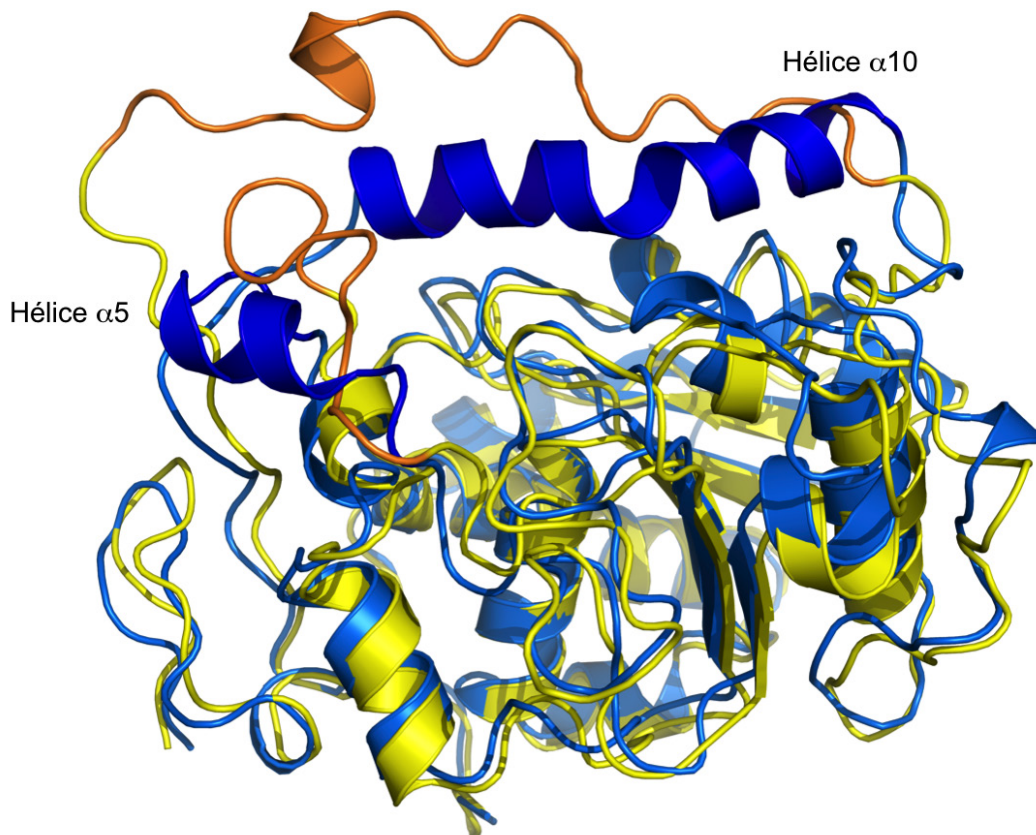


Figure V.11. Superposition de la frame initiale (bleu) et de la frame finale (jaune) pour le système Ile189Val-Ile285Val. L'hélice $\alpha 10$ (résidu 268 à 287) colorée en orange se détruit totalement au cours de la simulation de dynamique moléculaire. L'hélice $\alpha 5$ subit un déplacement latéral.

L'impact des doubles mutations sur le volume de la poche du site actif a également été étudié. Les valeurs moyennes sont reportées dans le Tableau V.2. Le volume de la poche du mutant Ile189Val-Ile285Val n'a pas pu être déterminé en raison de la déstructuration de l'hélice $\alpha 10$ qui contribue à la forme de la cavité. Pour les autres systèmes, les volumes sont au minimum le double de celui de la CALB sauvage ($209 \pm 56 \text{ \AA}^3$). Les doubles mutations occasionnent un élargissement prononcé de la cavité du site actif. Afin de souligner un éventuel effet synergétique des doubles mutations sur le volume, les données de volume relatives aux variants simples et doubles ont été comparées (Tableau V.2). Les volumes des mutants doubles mixtes, Ile189Ala-Ile285Val et Ile189Val-Ile285Ala, sont équivalents respectivement au mutant simple Ile189Ala et au mutant Ile285Ala. Pour le mutant Ile189Ala-Ile285Ala, une synergie négative est observée. En effet, l'espace dans la poche est inférieur au volume attendu avec des résidus alanine aux positions 189 et 285.

Tableau V.2. Valeurs moyennes et écart-types du volume de la poche du site actif (\AA^3) pour les mutants doubles et les mutants simples de la CALB.

Mutants doubles	Volume de la poche du site actif	Mutants simples	Volume de la poche du site actif
Ile189Ala-Ile285Ala	420 ± 87	Ile189Ala	432 ± 94
Ile189Ala-Ile285Val	437 ± 72	Ile189Val	287 ± 67
Ile189Val-Ile285Ala	569 ± 127	Ile285Ala	588 ± 116
Ile189Val-Ile285Ala	Non déterminé	Ile285Val	191 ± 54

L'influence des doubles mutations sur l'établissement des liaisons hydrogène entre les résidus constituant la triade catalytique de la CALB a également été étudiée. Les distances et les angles mesurés entre les résidus Asp187 et His224, et Ser105 et His224 sont donnés dans le Tableau V.1. Les trajectoires ont été découpées en différentes périodes de temps en fonction de l'établissement ou non de liaisons hydrogène. Quels que soient les systèmes, les liaisons hydrogène ne sont pas maintenues pendant les 20 ns des trajectoires. Certaines périodes de temps sont clairement marquées par l'absence de liaison hydrogène entre les résidus Asp187 et His224. Pour d'autres périodes de temps, la formation intermittente de la liaison hydrogène entre ces deux acides aminés est observée. Ceci est le cas pour la totalité de la trajectoire du variant Ile189Val-Ile285Ala et la période 'c' du mutant Ile189Ala-Ile285Ala. En revanche, les systèmes Ile189Ala-Ile285Val et Ile189Val-Ile285Val présentent au moins une liaison hydrogène bien établie sur une période de temps de plusieurs nanosecondes. Par ailleurs, la distance calculée entre la sérine et l'histidine catalytiques est stable pour les mutants Ile189Ala-Ile285Ala, Ile189Ala-Ile285Val et Ile189Val-Ile285Ala. Seul le système Ile189Val-Ile285Val ayant subi une déstructuration de l'hélice $\alpha 10$ présente une distance supérieure à celle observée pour l'enzyme sauvage, indiquant un éloignement des résidus catalytiques.

Globalement, à l'exception du variant Ile189Val-Ile285Val, les doubles mutants des acides aminés Ile189 et Ile285 ne sont le siège d'aucun réarrangement conformationnel spécifique au cours des simulations de dynamique moléculaire. En outre, les variants doubles n'induisent pas de flexibilité particulière dans la structure de la lipase. A l'exception du mutant Ile189Ala-Ile285Ala, les mutations doubles entraînent une augmentation du volume de la poche du site actif, en l'absence d'effet

synergique. L'accroissement du volume du site actif est causé par la présence d'un résidu alanine à la place des résidus Ile189 et/ou 285. Ces résultats suggèrent que la double substitution effectuée dans le site actif de la CALB ne perturbe pas la structure de la protéine.

4.2. Énergie d'interaction

La rapide déstructuration observée dans le cas du variant double Ile189Val-Ile285Val nous a conduit à calculer plus spécifiquement l'énergie d'interaction entre l'acide aminé en position 189 et l'hélice α_{10} qui induit le résidu 285. Ce calcul a été effectué pour déterminer l'impact des mutations sur la cohésion structurale de la protéine. L'énergie d'interaction fait référence à l'énergie d'interaction non liée qui est la somme de l'énergie de van der Waals et de l'énergie électrostatique, obtenue à partir de la mécanique moléculaire.

L'énergie d'interaction non liée, présentée dans la Figure V.12, révèle que la CALB sauvage est le siège d'interactions globalement plus fortes que celles calculées pour les variants. En outre, le remplacement d'une isoleucine par une alanine est plus déstabilisant pour la structure que la substitution par une valine. Cela indique qu'une diminution de la chaîne latérale conduit à une diminution de l'interaction. La diminution de l'énergie d'interaction provient principalement des interactions de van der Waals (Figure V.13). En effet, les variations de la contribution énergétique électrostatique sont faibles (Figure V.14).

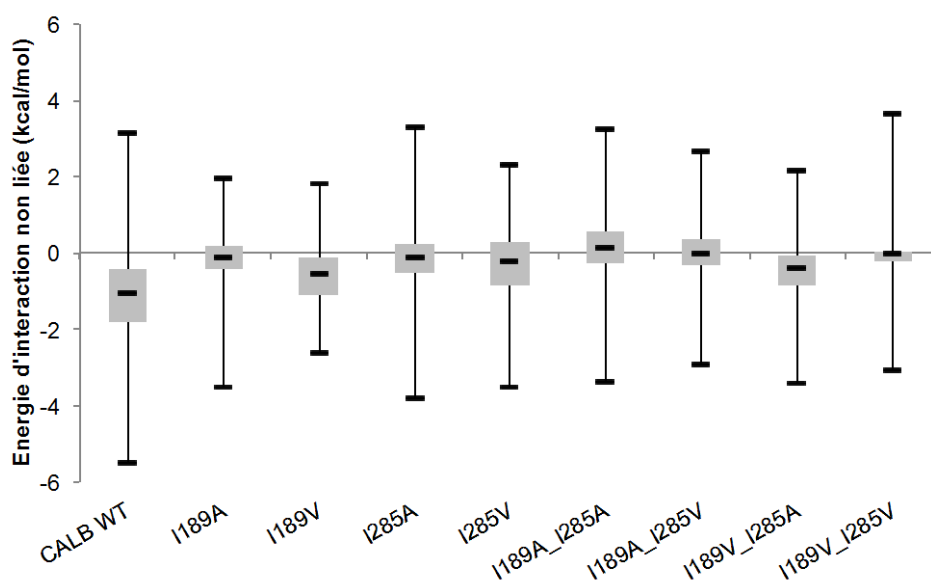


Figure V.12. Représentation en boîte à moustache de l'énergie d'interaction non liée calculée entre le résidu Ile189 et l'hélice α_{10} pour la CALB sauvage et ses variants simples et doubles pendant la phase de production de la trajectoire de DM. La boîte grisée correspond à 75% de l'information.

Pour les mutants doubles incluant une alanine à la position 189, la mutation a un effet synergétique déstabilisant entre le résidu Ala189 et l'hélice α_{10} . En effet, la double mutation donne une énergie de liaison qui excède la somme individuelle des variants simples. Dans le mutant Ile189Val-Ile285Ala, aucune synergie n'est constatée. Dans le cas du système Ile189Val-Ile285Val, la

substitution d'une isoleucine par une valine a une synergie déstabilisante conduisant à la perte totale d'interaction entre l'acide aminé Val189 et l'hélice $\alpha 10$.

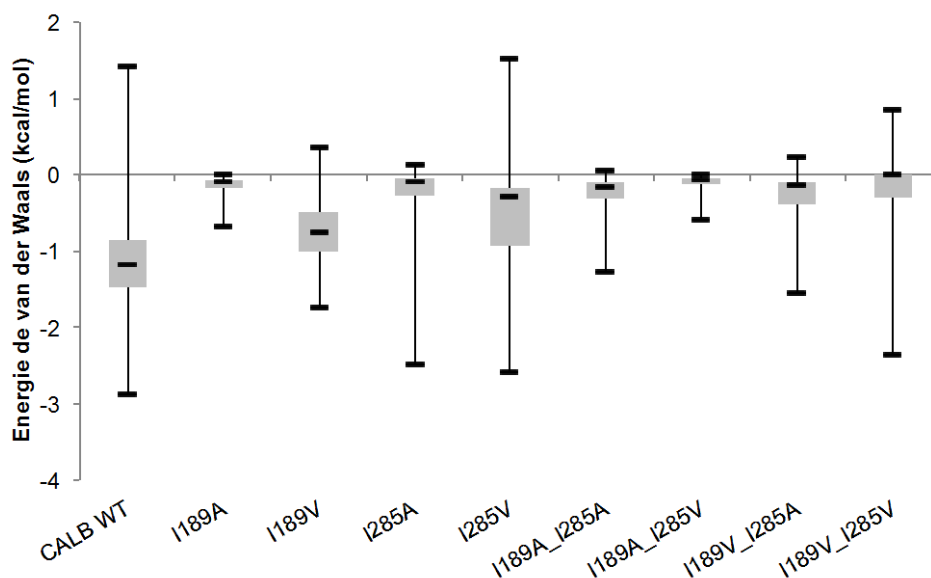


Figure V.13. Représentation en boîte à moustache de l'énergie d'interaction de van der Waals entre le résidu Ile189 et l'hélice $\alpha 10$ pour la CALB sauvage et ses variants simples et doubles.

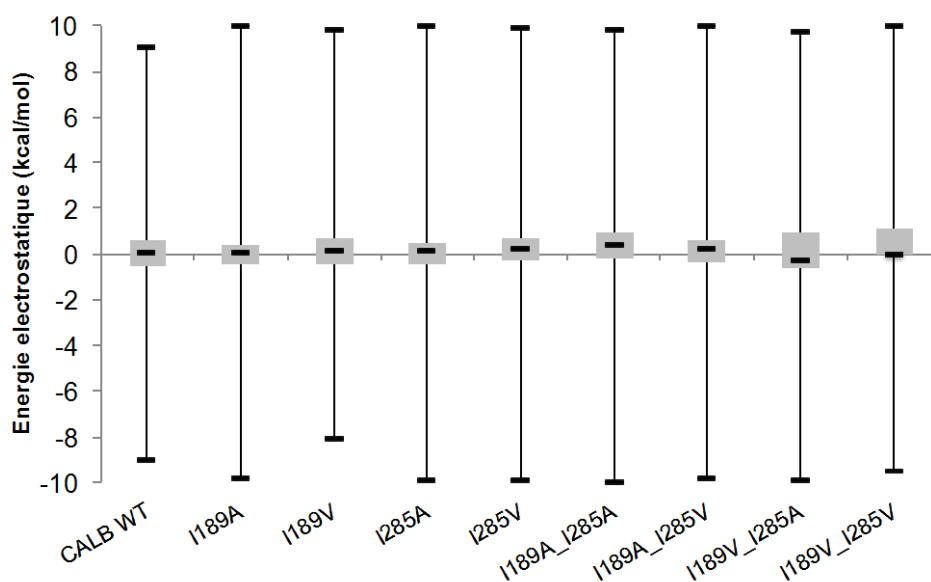


Figure V.14. Représentation en boîte à moustache de l'énergie d'interaction électrostatique entre le résidu Ile189 et l'hélice $\alpha 10$ pour la CALB et ses variants simples et doubles.

D'après ces résultats, les mutations simples et doubles des résidus Ile189 et Ile285 indiquent que ces acides aminés sont interconnectés. En effet, la substitution de l'une ou des deux isoleucines conduit à une diminution de l'interaction entre le résidu 189 et l'hélice $\alpha 10$. Les mutations doubles sont énergétiquement moins stables que les mutants simples. Les résultats suggèrent que ces acides aminés maintiennent l'intégrité de la forme de la cavité catalytique et participent à la cohésion interne de la protéine.

5. Étude complémentaire : étude structurale du variant Asp134Ala et des interactions avec la quercétine

Il a été démontré précédemment que dans les complexes impliquant la quercétine, le résidu Asp134 établit des liaisons hydrogène avec les fonctions hydroxyles de la quercétine. Cet acide aminé localisé au fond de la poche du site actif (Figure V.15) apparaît comme un point d'ancrage électrostatique pour le flavonoïde. Dans l'optique de vérifier cette hypothèse, ce résidu a été muté par un acide aminé aliphatique, l'alanine. Dans une première étape, une étude structurale par dynamique moléculaire (10 ns) du variant Asp134Ala a été réalisée afin de vérifier sa stabilité temporelle, de mettre en évidence d'éventuels réarrangements structuraux et de déterminer l'impact de cette mutation sur la structure de la CALB. Lors de la deuxième étape de cette étude, un protocole de docking a été appliqué afin de déterminer les interactions majeures établies entre la quercétine et la cavité catalytique et de conclure sur l'influence de cette substitution sur les modes d'interactions du flavonoïde.

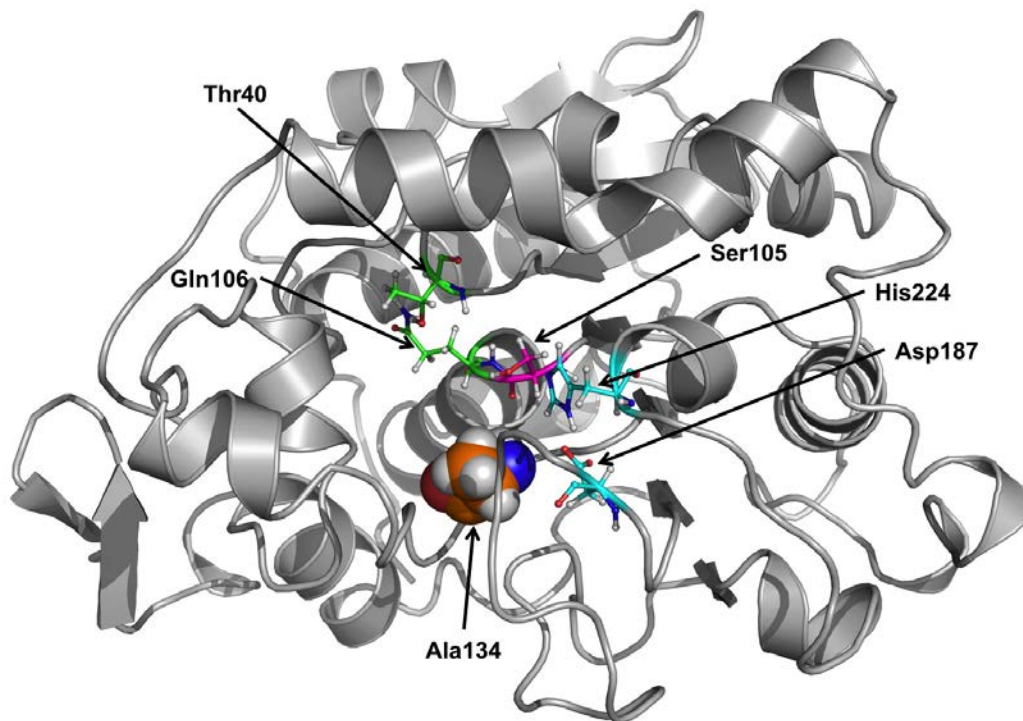


Figure V.15. Variant Asp134Ala de la lipase B de *Candida antarctica*. L'acide aminé muté, localisé au fond de la poche du site actif est représenté en sphère orange. Les résidus constituant la triade catalytique (Ser105, Asp187, His224) sont indiqués en rose et cyan, et ceux du trou oxyanionique sont colorés en vert.

5.1. Étude structurale du variant Asp134Ala

Le mutant Asp134Ala a été construit et soumis à une simulation de dynamique moléculaire d'une durée de 10 ns. Le suivi du RMSD au cours du temps est présenté dans la Figure V.16. La valeur moyenne calculée sur la phase de production (1 à 11 ns) est égale à 1.05 ± 0.07 Å. La stabilité temporelle du RMSD du variant Asp134Ala indique que la substitution de ce résidu n'a provoqué aucun changement conformationnel important dans la structure de la protéine.

Le B-factor a été calculé et comparé à celui de l'enzyme sauvage (Figure V.17). Les régions les plus mobiles de la structure du mutant sont l'hélice $\alpha 5$ flanquée de deux courtes boucles (résidus 139 à 150) et une boucle exposée au solvant (résidus 250 à 256). Ces régions de flexibilité sont également présentes dans l'enzyme sauvage. La présence d'une alanine en lieu et place de l'Asp134 n'est à l'origine d'aucun événement structural particulier.

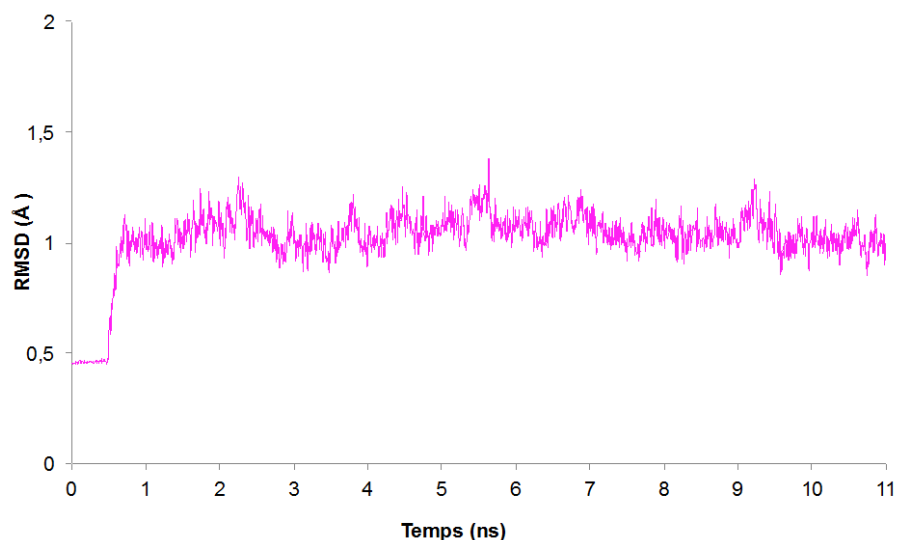


Figure V.16. Suivi du RMSD du C α au cours du temps pour le variant Asp134Ala de la CALB.

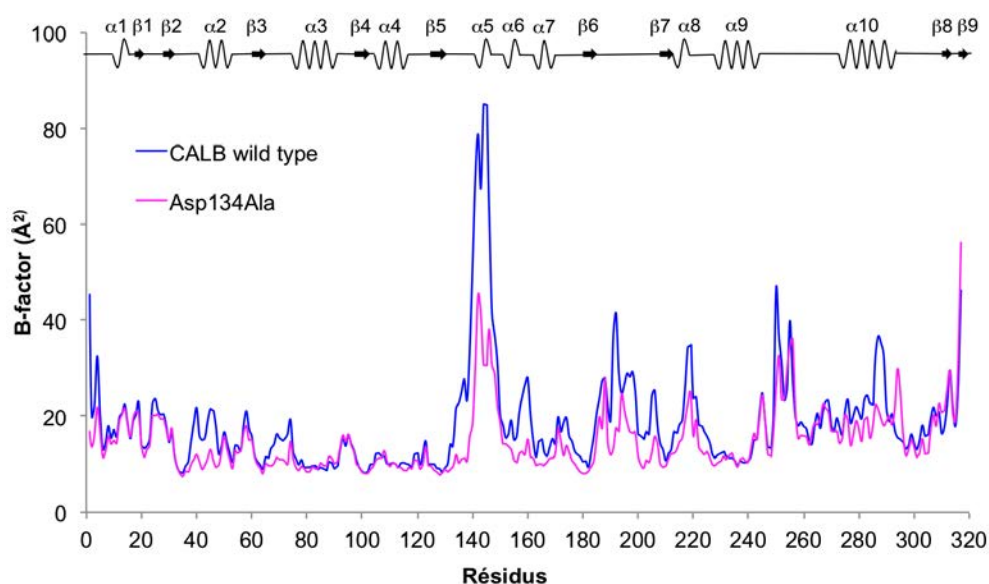


Figure V.17. B-factor en fonction des résidus du variant Asp134Ala de la CALB (rose) et de la CALB sauvage (bleu).

Le suivi du volume de la cavité catalytique au cours de la trajectoire pour le variant Asp134Ala et l'enzyme sauvage est donné dans la Figure V.18. Il apparaît clairement que le remplacement de l'aspartate 134 par une alanine n'affecte pas le volume de la cavité de la CALB contrairement aux résultats obtenus lors de la mutation des résidus Ile189 et Ile285.

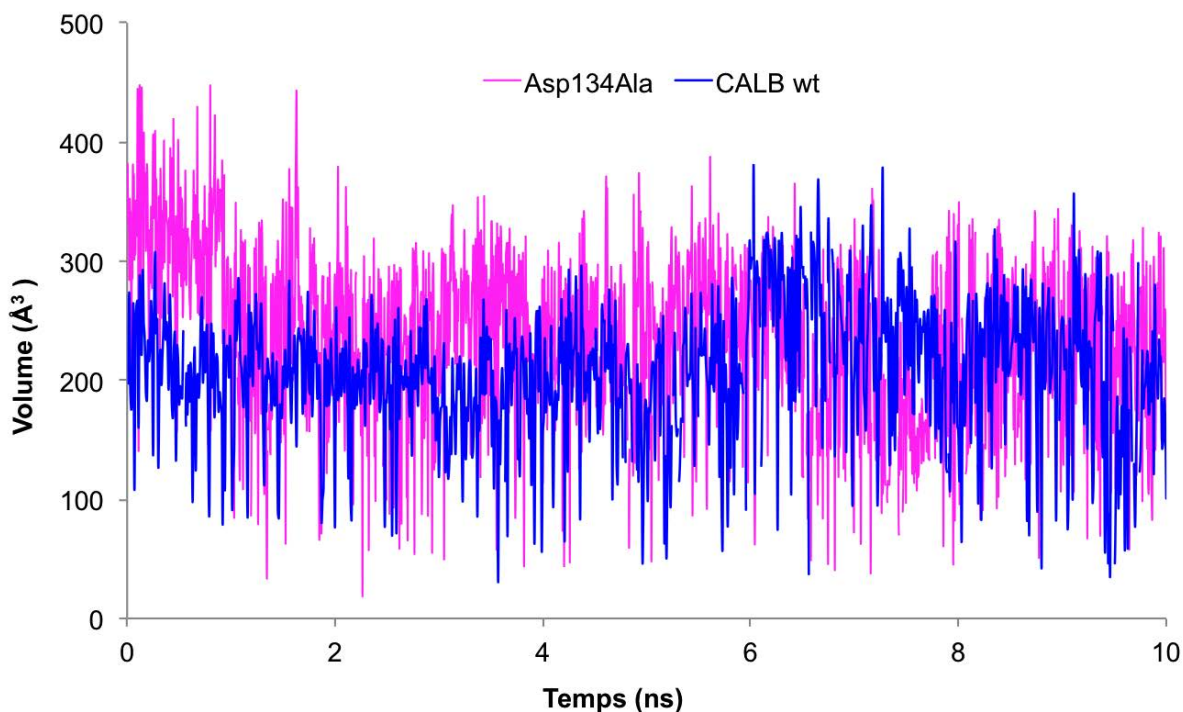


Figure V.18. Suivi du volume de la cavité catalytique du variant Asp134Ala au cours de la trajectoire de DM.

Une analyse des distances et des angles entre les acides aminés de la triade catalytique a été exécutée visant à évaluer la présence de liaisons hydrogène au cours de la simulation. Les valeurs moyennes et les écart-types sont regroupés dans le Tableau V.3. La liaison hydrogène entre l'Asp187 et l'His224 est présente pendant toute la trajectoire, oscillant entre les atomes OD1 et OD2 de l'acide aspartique. En outre, la distance entre la sérine catalytique et l'histidine catalytique est maintenue, indiquant une stabilité de ces résidus.

Tableau V.3. Distances et angles entre les résidus constituant la triade catalytique au cours des 10 ns de la phase de production de la trajectoire de DM.

Résidu:atome	Distance (Å)	Résidu:atome	Angle (°)
Asp187 :OD1/His224 :HD1	2.28 ± 0.40	Asp187 :OD1/His224 :HD1/His224 :ND1	145.65 ± 18.28
Asp187 :OD2/His224 :HD1	2.57 ± 0.59	Asp187 :OD2/His224 :HD1/His224 :ND1	137.95 ± 18.82
Ser105 :OG/His224 :NE2	3.14 ± 0.35	Ser105 :OG/Ser105 :HG1/His224 :NE2	109.26 ± 21.63

L'ensemble des résultats de modélisation moléculaire indique que la mutation ponctuelle du résidu Asp134 en alanine ne provoque aucune perturbation de la conformation protéique de la CALB. De plus, l'intégrité structurale de la triade catalytique est maintenue au cours des trajectoires de dynamique moléculaire.

5.2. Interaction de la quercétine avec le variant Asp134Ala

Une procédure de docking a été appliquée en utilisant une image instantanée issue de la trajectoire du variant Asp134Ala comme cible, Le but est d'évaluer l'impact de cette mutation sur les interactions entre la lipase et la quercétine. La sélection de l'image instantanée a été effectuée en se basant sur deux critères : la présence d'une liaison hydrogène entre les résidus Asp187 et His224 de la triade catalytique ; la présence de liaisons hydrogène entre le carbonyle de l'acétate lié à la sérine catalytique et les résidus du trou oxyanionique. Une image instantanée répondant à ces critères a été sélectionnée et soumise à la procédure de docking.

Les résultats révèlent que la quercétine adopte deux orientations préférentielles dans la cavité catalytique : une orientation CA, désignant l'insertion du cycle A vers le cœur de l'enzyme, et une orientation CCB, correspondant à un positionnement horizontal de la quercétine dans la cavité, le cycle B du substrat pointant vers l'hélice $\alpha 5$ (Figure V.19). Dans l'orientation CA, deux interactions électrostatiques sont observées. La première liaison hydrogène est établie entre le groupement 7-OH de la quercétine et le résidu Asp145 appartenant à l'hélice $\alpha 5$. La seconde liaison se forme entre le groupement 5-OH de la quercétine et le résidu Gln157 localisé au fond de la poche catalytique. Dans l'orientation CCB, les liaisons hydrogène se forment entre les fonctions hydroxyles 3'-OH et 4'-OH du flavonoïde et le résidu Asp145.

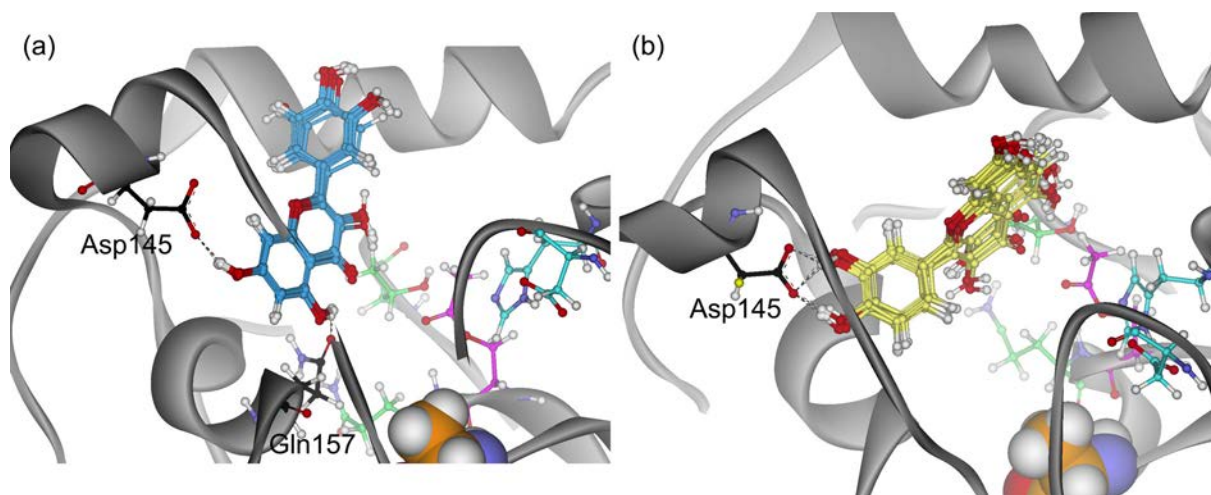


Figure V.19. Images montrant les deux orientations majoritaires de la quercétine, CA (a) et CCB (b), observées à l'issue de la procédure de docking dans la cavité catalytique du variant Asp134Ala. Les résidus de la triade catalytique sont représentés en stick de couleur cyan (Asp187, His224) et rose (Ser105) et ceux du trou oxyanionique sont en vert (Thr40 et Gln106). Le résidu Asp134 muté en alanine est représenté par une sphère orange. Les liaisons hydrogène entre les groupements hydroxyles de la quercétine et les résidus du site actif sont indiqués en traits discontinus.

Ces résultats de docking montre que le remplacement du résidu Asp134 en résidu aliphatique influence le positionnement de la quercétine dans le site actif. En effet, son orientation s'est adaptée de façon à interagir avec le résidu polaire Asp145. L'élimination du point d'ancrage électrostatique au fond de la poche est remplacée par un autre résidu hydrophile. D'après l'étude semi-empirique menée

par Codorniu-Hernandez et al., les acides aminés hydrophiles comprenant l'Asp, la Glu et la Lys ont une affinité élevée pour les flavonoïdes et plus particulièrement, si les résidus sont chargés au pH physiologique. La présence de tels résidus dans la poche du site actif constitue un élément attractif pour les groupements hydroxyles de la quercétine, qui la dévie des résidus catalytiques.

Chapitre VI. Étude Théorique des
Interactions entre la Quercétine et ses
Dérivés, et le Site Actif de la Lipase B de
Candida antarctica

1. Introduction

Dans les travaux précédents, les orientations et l'affinité de liaison de la quercétine pour la lipase B de *Candida antarctica* (CALB) ont été étudiées. Tout d'abord, il a été montré que la quercétine adopte deux orientations majeures dans la cavité de la CALB sauvage, elle s'insère soit avec son cycle A (CA) soit avec son cycle B (CB) vers le cœur de l'enzyme. Dans ce cas, le flavonoïde établit des liaisons hydrogène avec les résidus Thr40, Asp134, Gln157. Lors de la mutation des résidus Ile189 et Ile285 par des alanine ou valine, il en résulte quatre orientations de la quercétine dans les cavités de la CALB mutée. L'orientation CB est présente dans tous les mutants tandis que l'orientation CA est retrouvée lors de la mutation des isoleucines par une alanine. En outre, deux orientations complémentaires ont été mises en évidence. Celles-ci correspondent à un étalement de la quercétine dans les sites actifs des variants avec soit le cycle A (CCA) soit le cycle B (CCB) pointant vers l'hélice $\alpha 5$. D'autre part, le calcul d'énergie libre de liaison révèle un accroissement de l'affinité du flavonoïde aglycone pour les sites actifs de la CALB mutées. La diminution de la longueur des chaînes latérales des isoleucines permet à la quercétine de s'ancrer plus profondément dans le site actif. De ce fait, un renforcement du nombre de liaisons hydrogène établies avec la quercétine a été constaté. Les principaux résidus impliqués dans ces interactions sont les suivants : Thr40, Asp134, Thr138, Ser150, Gln157, Asp187 et Glu188. Il en résulte une intensification de la contribution électrostatique dans l'interaction globale de la quercétine avec le site actif de la CALB. Lors de la substitution du résidu Asp134 par un résidu aliphatique, il a été observé que la quercétine s'étale dans la cavité et adopte les orientations CA et CCB définies ci-dessus. Dans ce cas, les groupes hydroxyles de la quercétine établissent des liaisons hydrogène avec le résidu Asp145 localisé dans l'hélice $\alpha 5$. A partir de ces résultats de modélisation moléculaire, une étude est menée pour évaluer l'orientation et l'énergie libre de liaison de la quercétine et de ses dérivés dans le site actif de la CALB. Pour cela des systèmes correspondant à des réactions d'acétylation de la quercétine et d'hydrolyse de la quercétine acétylée ont été construits. L'étude des interactions enzyme-quercétine a été approfondie en observant l'effet de modifications structurales ponctuelles, localisées au niveau des différents groupements hydroxyles du flavonoïde aglycone ou de la chaîne latérale de la sérine catalytique. Une recherche de relations de type cause-effet a été menée afin de mieux comprendre les interactions entre la quercétine et la CALB. *In fine*, il s'agit d'identifier des pistes pour l'amélioration de la CALB et le design d'une lipase capable de catalyser l'acylation régiosélective de la quercétine.

Plus précisément, cette étude s'articule autour des points suivants :

- La comparaison des énergies libres de liaison de la quercétine et de son analogue glycosylé, l'isoquercitrine. La CALB étant capable de catalyser l'acylation de l'isoquercitrine, il paraissait intéressant de mettre en parallèle les systèmes impliquant d'une part la quercétine dont l'acylation n'est pas catalysée par la lipase, d'autre part l'isoquercitrine.
- L'étude de l'effet des interactions enzyme-substrat impliquant les groupes phénoliques de la quercétine sur ses modes de liaison dans la cavité catalytique de la CALB et l'affinité de la lipase pour le substrat. L'approche mise en œuvre consiste à supprimer localement les interactions électrostatiques impliquant les fonctions hydroxyle de la quercétine par le biais de

méthylations ponctuelles du substrat. L'effet sur les modes de liaison du substrat et l'affinité de la lipase est analysé.

- L'étude de l'influence de la chaîne acétate de l'acyle-enzyme sur les modes d'interactions de la quercétine ou ses dérivés avec la CALB. La quercétine ou ses dérivés méthylés est dockée dans la cavité catalytique de la CALB, cette dernière étant soit sous forme d'acyle-enzyme, soit sous forme d'enzyme native. L'influence de l'état libre ou acétylé de la sérine catalytique sur les positions préférentielles et les interactions de la quercétine ou ses dérivés méthylés est analysée. Une comparaison est faite avec des systèmes associant des dérivés acétylés de la quercétine et la CALB native (cas des réactions d'hydrolyse de la quercétine acétylée).

Une méthodologie combinant des simulations de docking et des calculs d'énergie libre de liaison a été appliquée aux différents systèmes enzyme-substrat. Dans un premier temps, cette procédure a été utilisée dans un modèle d'acétyle-enzyme produit au cours de la réaction d'acétylation de la quercétine. Pour construire ce modèle, la chaîne latérale de la sérine catalytique est acétylée. Les simulations de docking dans cette première cible ont été appliquées à deux substrats de référence, la quercétine et l'isoquercitrine et à différents dérivés méthylés de la quercétine : 7-O-méthyl, 3'-O-méthyl, 3,5-O-méthyl, 3,5,7-O-méthyl, 7,3',4'-O-méthyl quercétine. A l'issue du docking, les orientations et positionnements préférentiels de la quercétine ou ses dérivés ont été analysés et corrélés à la structure des substrats. Une démarche identique a été menée avec des systèmes associant la quercétine ou ses dérivés méthylés et la CALB native afin de mettre en évidence l'effet de la présence d'une chaîne acétate liée à la sérine catalytique sur les interactions entre l'enzyme et ses substrats. L'énergie libre de liaison des complexes générés dans l'étape précédente a été calculée en utilisant la méthode MM-PBSA. Les résultats d'énergie libre de liaison pour les différents systèmes ont été analysés en effectuant une répartition en quartile. De plus, les composantes de l'énergie libre de liaison ont été examinées de manière plus détaillée. Dans un second temps, la méthodologie exposée précédemment a été appliquée à des modèles associant la CALB native et des dérivés acétylés de la quercétine, l'objectif étant de mimer la réaction d'hydrolyse des esters de quercétine. Dans ce cas, des dérivés acétylés de la quercétine (7,3',4'-triacétate et 3,5,7,3',4'-pentaacétate de quercétine) ont été dockés dans la cavité catalytique de la CALB. La Figure VI.1 montre les différents systèmes étudiés ainsi que les objectifs recherchés.

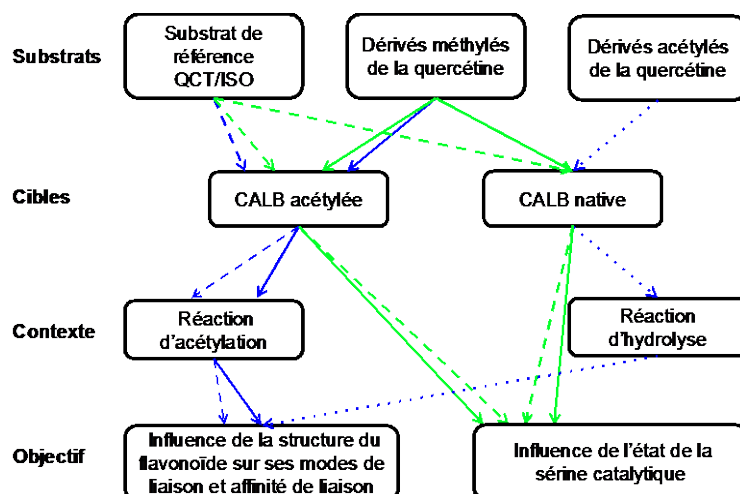


Figure VI.1. Présentation des différents systèmes étudiés. La couleur des flèches est reliée à l'objectif étudié et les traits pleins et discontinus des flèches correspondent aux différents substrats. QCT et ISO désignent la quercétine et l'isoquercitrine, respectivement.

Cette partie du travail a été rédigée sous la forme d'un projet de publication.

2. Article

Affinity of *Candida antarctica* lipase B towards quercetin and its derivatives: a theoretical study

Bidouil Christelle^a, Chebil Latifa^a, Maigret Bernard^b, Ronat-Heidt Evelyne^a, Ghoul Mohamed^a, Engasser Jean-Marc^a & Humeau Catherine^{a*}

^a Laboratoire Ingénierie des Biomolécules, ENSAIA-INPL, Université Lorraine, 2 av. de la Forêt de Haye, 54500, Vandoeuvre-lès-Nancy, France.

^b Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), CNRS, Université Lorraine, BP 239, 54506, Vandoeuvre-lès-Nancy, France

*Corresponding author: Tel.: +33 3 83 59 57 84; fax: +33 3 83 59 57 78.

E-mail address: catherine.humeau@ensaia.inpl-nancy.fr

Abstract:

Interactions involved both in quercetin or methyl-quercetin acetylation and acetyl-quercetin alcoholysis catalyzed by *Candida antarctica* lipase B (CALB) were studied through molecular modelling simulations. Docking of quercetin and its methylated derivatives against the catalytic pocket of CALB under an acetyl-enzyme state was performed aiming to study the influence of the substrate functional groups on its binding modes. For each enzyme-substrate complex, the binding affinity of the lipase was evaluated through free binding energy calculations using the MM-PBSA method. Methylation of quercetin hydroxyl groups was shown to influence the major orientations of the substrate within CALB pocket. Substrates exhibiting a free hydroxyl group on the 4' position adopted such an orientation that the aromatic B ring pointed down to the cavity bottom. A contrary tendency was observed with 4'-Me substrates that first slipped their aromatic A ring within the cavity. Otherwise, the binding affinity of CALB towards methyl-quercetin substrates decreased comparing with affinity towards quercetin, due to electrostatic contribution weakening. The catalytic state of the lipase (free enzyme or acetyl-enzyme) was shown to affect its affinity towards quercetin or methyl-quercetin, the highest binding affinities being obtained for complexes involving acetyl-enzyme systems. Concerning models miming acetyl-quercetin alcoholysis reaction, substrates exhibiting an acetyl group on the 3', 4' and 7 positions brought their hydrolysable groups near to the catalytic triad, suggesting that hydrolysis could occur. Binding affinities observed for these systems were globally lower than those obtained for models miming quercetin acetylation. An increase of apolar contribution together with a decrease of electrostatic contribution was observed.

Keywords: *Candida antarctica* lipase B, quercetin, electrostatic interactions, docking, MM-PBSA, binding free energy, alcoholysis, acetylation

1. INTRODUCTION

Flavonoids are polyphenolic compounds that are largely distributed in the plant kingdom. These compounds are subjected to intensive research due to their high applicative potential such as antioxidant, anti-allergenic, antiproliferative, antiviral, anti-inflammatory, and anticarcinogenic agents [1]. Most of these biological activities are related to their capacity to scavenge radical species that is strongly linked to their chemical structure. Several experimental and theoretical studies showed that the number of hydroxyl groups and their position on the benzopyrane skeleton play a pivotal role in their chemical and biological activities. Among flavonoids, quercetin (Q) has been widely investigated because of its numerous activities, especially its antioxidant power that appears very attractive for various areas of application as pharmaceuticals, functional foods and cosmetics [2, 3]. However, quercetin use remains limited because of unstable physicochemical properties (stability and solubility) that make its incorporation into various matrices difficult. One way to overcome these drawbacks is the preparation of lipophilic derivatives of quercetin that can be obtained by acylation with aliphatic molecules. The regioselectivity of the process to be used constitutes a key condition to maintain the properties of interest. That is the reason why enzymatic processes are preferred to chemical ones that generally lack selectivity.

Candida antarctica lipase B (CALB) is a versatile enzyme which efficiency to catalyze acylation reactions has been widely reported. Its catalytic triad consists of three amino acids: aspartic acid (D187), histidine (H224) and a nucleophilic serine (S105). These residues are part of an extensive hydrogen bond network that contributes to maintain the activity of the lipase and substrate positioning. Additional amino acids T40 and Q106 defined the oxyanion hole that stabilizes the catalytic intermediate through hydrogen bond interactions. CALB follows the well-known Bi-Bi ping-pong mechanism that is based on the successive acylation and deacylation of the catalytic serine and the formation of an acyl-enzyme intermediate whether in acylation or hydrolysis reactions. Several years ago, CALB was shown to be very efficient to catalyze the acylation of glycosylated flavonoids [4-6]. However, the direct acylation of aglycone flavonoids was reported to be much more difficult [4, 5, 7]. Such reactions need to be highly regioselective as flavonoids exhibit many hydroxyl groups that are susceptible to undergo acylation at the risk of losing the desired properties. One way to study the regioselectivity of these reactions consists in studying enzyme/substrates complexes through molecular modeling simulations [6, 8]. More specifically, substrate binding modes as well as interactions must be investigated to gain a better understanding of lipase selectivity properties towards such poly-functional compounds.

The present study aims to investigate the interactions between quercetin functional groups and CALB residues whether in quercetin acetylation reaction or in acetyl-quercetin alcoholysis. Quercetin hydroxyl groups were methylated either successively or simultaneously aiming to propose a cause-effect relationship between the structure of the substrate and its orientations and interactions within CALB catalytic cavity. Enzyme/substrates complexes were generated by docking then submitted to binding free energy calculations using the MM-PBSA implicit solvent method. Indeed, Rastelli et al. [9] reported a good correlation between binding free energies that were determined for a single energy-minimized structure in an implicit solvation model and those obtained from molecular

dynamics simulations that were carried out on explicit solvent systems. A similar approach was applied to acetyl-quercetin derivatives aiming to study interactions involved in alcoholysis reactions. For each system, the electrostatic and hydrophobic contributions were determined aiming to draw conclusions about the main force driving quercetin binding modes and interactions, whether in acylation or alcoholysis reactions.

2. COMPUTATIONAL METHODS

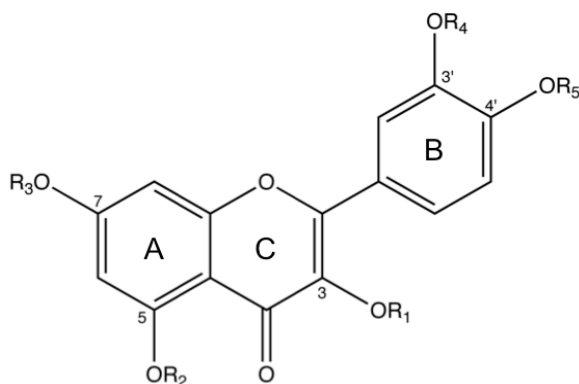
2.1. - Preparation of structures

2.1.1. Target preparation

The crystal structure of *Candida antarctica* lipase B (PDB entry: 1LBS [10]) was used as the starting system to prepare two models: (i) the free enzyme without any substrate into the catalytic cavity (CALB) was obtained by removing the ethylhexylphosphonate inhibitor (HEE) covalently bound to the Ser105; (ii) the acetyl-lipase that corresponds to the intermediate acyl-enzyme in the acetylation reaction (acetyl CALB) was constructed by replacing the HEE chain bounded to the OG of the catalytic serine (Ser105) by an acetate [11, 12]. For both models, all missing hydrogen atoms were added and all crystal water molecules were kept. Protonation states of residues were assumed to be those observed at pH 7.0. CALB and acetyl CALB systems were immersed in an orthorhombic box of TIP3P water molecules with a margin of 15 Å along each dimension. The systems were neutralized by the addition of one sodium counterion and submitted to molecular dynamics simulations using CHARMM forcefield and NAMD2.7 software. The systems were minimized for 6400 steps combining the conjugate gradient and the line minimizer methods implemented in NAMD. Then, an equilibration step of 500 ps was applied, keeping the protein harmonically restrained. The force constant was initially fixed at 10 kcal.mol⁻¹.Å⁻² and was progressively decreased (0.5 kcal.mol⁻¹.Å⁻²/25 ps). The systems were again equilibrated for 500 ps by releasing all the restraints. The particle mesh Ewald (PME) method [13] was used to compute the long-range electrostatic interactions. The SHAKE algorithm [14] was applied and the time step was set to 1.0 fs. Finally, trajectories of 10 ns were produced in NPT ensemble (1 bar, 300 K). Coordinates were saved every 1 ps during the sampling process. The frames selected from molecular dynamics trajectories for further docking simulations were chosen according to several criteria: (i) the relative conformational orientations of the catalytic residues [15]; (ii) the hydrogen bond network implying the oxyanion hole residues.

2.1.2. Ligand preparation

The three dimensional structures of substrates to be docked were manually constructed with Discovery Studio 2.5 (DS) (Discovery Studio 2.5 Accelrys Inc., USA), hydrogens were added, atoms types were adjusted by applying CHARMM force field, partial charges were assigned with CFF method and the raw structures were corrected using the 'clean geometry' functionality in the builder tool. Finally, all ligands were subjected to a single step minimization protocol using the steepest descent method for 200 steps and applying a generalized Born model with a Molecular Volume (GBMV) in order to consider the solvation effect [16]. Structures and nomenclatures of the 10 ligands under study are reported in Figure VI.2.



	R ₁	R ₂	R ₃	R ₄	R ₅
Native Quercetin (Q)	H	H	H	H	H
7- O - methyl Q	H	H	CH ₃	H	H
3' - O - methyl Q	H	H	H	CH ₃	H
3,5 - O - dimethyl Q	CH ₃	CH ₃	H	H	H
3,5,7 - O - trimethyl Q	CH ₃	CH ₃	CH ₃	H	H
7,3',4' - O - trimethyl Q	H	H	CH ₃	CH ₃	CH ₃
7,3',4' - O - triacetyl Q	H	H	CO-CH ₃	CO-CH ₃	CO-CH ₃
3,5,7,3',4' - O - pentaacetyl Q	CO-CH ₃	CO-CH ₃	CO-CH ₃	CO-CH ₃	CO-CH ₃
Isoquercitrin	H	H	glucose	H	H

Figure VI.2. Chemical structures and nomenclatures of the ligands under study

2.2. Molecular Docking

The favourable binding modes of the ligands within the catalytic cavity of the lipase were determined by a docking protocol using the molecular docking software LigandFit [17] included in DS 2.5 software. The targets CALB and acetyl CALB extracted from previous molecular dynamics simulations and the 10 ligands were used as inputs. LigandFit is a receptor-rigid docking algorithm that performs the conformation search of the ligand poses by Monte Carlo trial method and uses a shape-matching filter to compare ligand poses with binding site shape. Candidate ligand poses are then positioned into the binding site and submitted to a rigid body energy minimization. For every docked molecules, twenty poses were retained from each docking run. During the docking procedure, the ligand-enzyme complexes were minimized using 1000 steps of steepest descent algorithm. The receptor atoms were free to move within a sphere of 12 Å including the active site. The twenty poses were ranked according to their orientation in the active site and analysed to investigate the interactions. Moreover, the protein-ligand complexes were submitted to molecular mechanics (MM)/Poisson-Boltzmann surface area (PBSA) for relative quantification of binding affinity for substrates in active site of the enzyme.

2.3. Binding free energy calculations

The binding free energy (ΔG_{bind}) (eq. 1) of protein/ligand complexes was calculated using the MM/PBSA method [18].

$$\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - \Delta G_{\text{protein}} - \Delta G_{\text{ligand}} \quad (1)$$

$$\Delta G = E_{MM} + G_{solvation}^{polar} + G_{solvation}^{nonpolar} - TS \quad (2)$$

$$E_{MM} = E_{bond} + E_{angle} + E_{torsion} + E_{vdW} + E_{elec} \quad (3)$$

The free energy of complexes, protein and ligands were denoted as $G_{complex}$, $G_{protein}$ and G_{ligand} respectively. The molecular mechanic energy (E_{MM}) is the sum of intramolecular and intermolecular energies (Eq. 3). The polar contribution to the solvation free energy ($G_{solvation}^{polar}$) was calculated by solving the Poisson-Boltzmann equation. The grid spacing was set to 0.4 Å. The dielectric constants inside and outside the molecules were equal to 1.0 and 80.0, respectively. The nonpolar solvation contribution ($G_{solvation}^{nonpolar}$) was obtained from the solvent accessible surface area (SASA) using the following relationship: $\Delta G_{solvation}^{nonpolar} = \gamma SASA + \beta$. The values for γ and β were set to 0.00542 kcal.mol⁻¹.Å⁻² and 0.92 kcal.mol⁻¹, respectively. The probe radius was 1.4 Å. The entropy component $T \Delta S$ was evaluated by normal-mode analysis, at 298K.

A distribution of ΔG_{bind} values neglecting entropic contribution was performed for quercetin orientations including at least five poses. This distribution was represented as a box plot that reflected the values of the first quartile, median, third quartile as well as the minimum and maximum values of binding free energy.

3. RESULTS

3.1. Dynamic behaviour of CALB and acetyl-CALB targets

The conformational behaviour and the stability of CALB and acetyl-CALB all along molecular dynamics trajectories were studied in order to choose pertinent target for further docking. The RMSD of the C α atoms referring to the initial structure of the proteins, the 2D-RMSD and characteristic distances and angles implicated in the catalytic pocket were determined. During the production phase, RMSD values ranged from 0.96 to 1.60 Å for CALB and from 0.95 to 2.01 Å for acetyl-CALB. Main movements were observed on helix $\alpha 5$ residues (140 to 149) in both systems and on a loop including residues 23 to 30 in acetyl-CALB. When neglecting these residues, average RMSD on 10 ns simulation was 1.07 \pm 0.15 Å and 1.06 \pm 0.10 Å for CALB and acetyl CALB, respectively. These results are in accordance with a previous study that showed the high flexibility of helix $\alpha 5$ through molecular dynamics simulations [19]. The analysis of the 2D-RMSD graph indicated that CALB adopted three conformational families corresponding to three simulation periods (Figure VI.3 (a)). In the first one (from 0 ns to 4.8 ns), hydrogen bond network that favours correct orientation of catalytic residues was maintained (Ser105:HG1/His224:NE2, Asp187:OD1 (or Asp187:OD2)/His224:HD1) (Figure VI.4 (a)). After this period and up to 9.5 ns, a disruption of H-bond interactions between catalytic residues was observed, that was due to the insertion of a water molecule. During the last 0.5 ns of simulation, atoms Asp187:OD1 (or OD2) and His224:HD1 approached again and re-established an H-bond interaction. In the case of acetyl-CALB system, the 2D-RMSD showed that three major conformations (0 to 5.8 ns; 6.9 ns to 8.5 ns; 8.5 ns to 10 ns) as well as a transition phase (5.8 to 6.9 ns) were adopted during simulations (Figure VI.3 (b)). As for CALB, in the first conformational cluster,

hydrogen bond network between catalytic residues was maintained (Ser105:HG/His224:NE2, Asp187:OD1 (or Asp187:OD2) /His224:HD1) (Figure VI.4 (b)). The serine-bound acetate came alternatively within or without oxyanion hole (Figure VI.5). For the two other conformational clusters and the transition state, hydrogen bond interaction between Asp187 and His224 residues was broken while serine-bound acetate pointed towards the oxyanion hole.

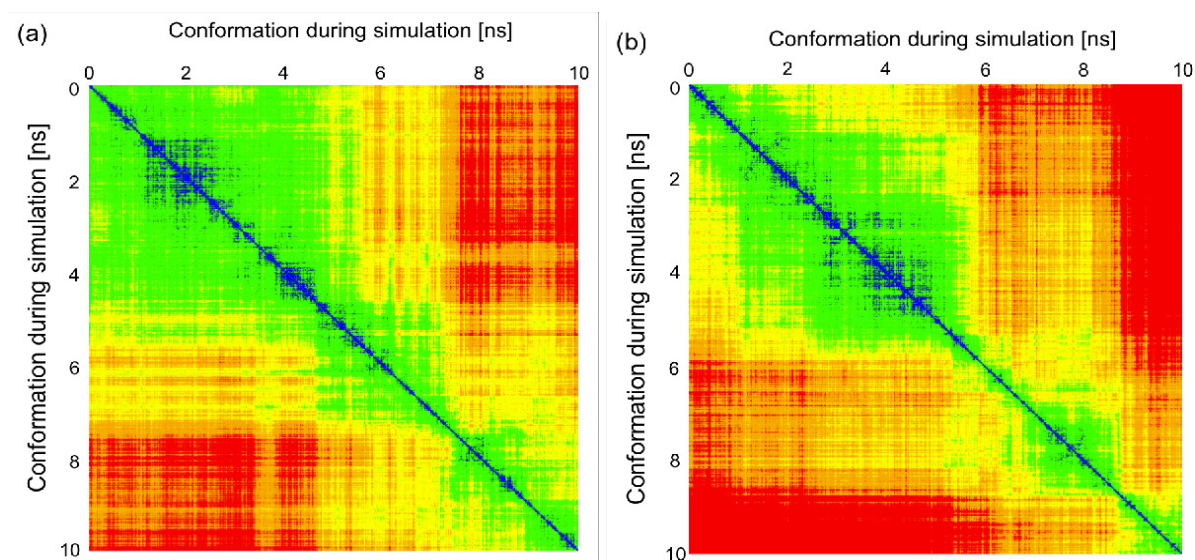


Figure VI.3. 2D-RMSD plots calculated from molecular dynamics trajectories of CALB (a) and acetyl CALB (b). The root mean square deviation of every conformation referring to all other conformations and basing on CALB C α atoms was given as a function of time.

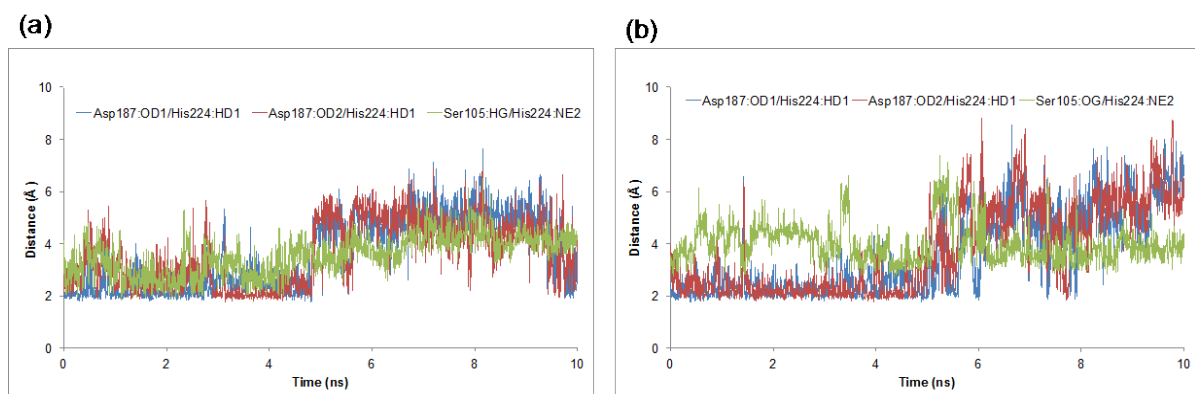


Figure VI.4. Distances (Å) within catalytic triad of CALB (a) and acetyl CALB (b) throughout molecular dynamics simulations.

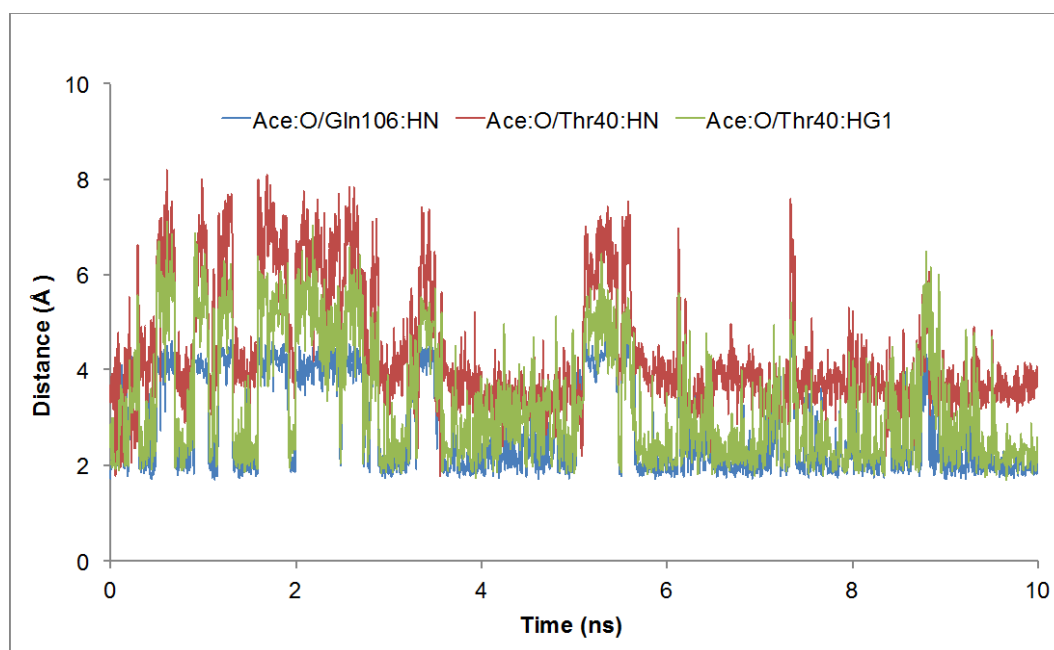


Figure VI.5. Distances (Å) between the serine-bound acetate and the oxyanion hole residues throughout molecular dynamics simulations of acetyl-CALB system.

For both systems CALB and acetyl-CALB, target for docking was chosen among frames of the first conformational family that respects criteria for lipase to be active (Figure VI.6).

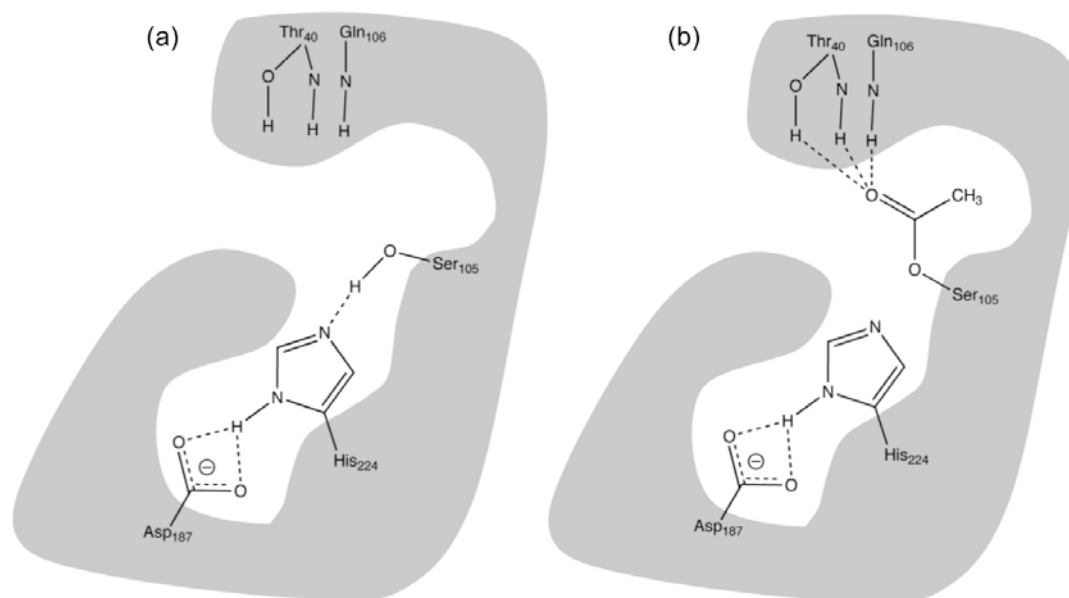


Figure VI.6. Schematic representation of the criteria retained for the selection of the frames subjected to further molecular docking. In CALB system (a), the Ser105 residue interacts with the His224 residue that also interacts with the Asp134 residue. These three amino acids are referred as the catalytic triad. In acetyl CALB system (b), additional interactions between the serine-bound acetate (Ace:O) and oxyanion hole residues (Thr40 and Gln106) are established.

In these target models, the RMSD values for CALB and acetyl-CALB referring to crystallographic data as starting structures were 1.11 Å and 0.97 Å, respectively. Characteristic distances and angles are

reported in Tables VI.1 and VI.2. For both systems, His224 imidazole (HD1) interacted with Asp187 (OD1) via a H-bond and Ser105:OG was turned to His224:NE2, showing a suitable orientation of the catalytic residues. In acetyl-CALB system, the carbonyl oxygen of the serine-bound acetate was located in the oxyanion hole with hydrogen bonding distances to the amide nitrogen of Thr40 and Gln106 (2.15 Å and 1.88 Å, respectively).

Table VI.1. Distances (Å) and angles (°) measured in CALB catalytic cavity.

Residues and atoms	Distances	Angles
Ser105:HG/His224:NE2	1.88	
Ser105:OG/Ser105:HG/His224:NE2		114.50
Asp187:OD1/His224:HD1	1.99	
Asp187:CG/Asp187:OD1/His224:HD1		123.55
Asp187:OD2/His224:HD1	3.15	
Asp187:CG/Asp187:OD2/His224:HD1		64.71
Ser105:OG/Gln106:HN	2.12	
Ser105:CB/Ser105:OG/Gln106:HN		75.99
Ser105:OG/Thr40:HN	2.00	
Ser105:CB/Ser105:OG/Thr40:HN		141.64
Ser105:OG/Thr40:HG1	5.34	
Ser105:CB/Ser105:OG/Thr40:HG1		97.75

Table VI.2. Distances (Å) and angles (°) measured in acetyl-CALB catalytic cavity.

Residues and atoms	Distances	Angles
Ser105:OG/His224:NE2	2.64	
Asp187:OD1/His224:HD1	2.75	
Asp187:CG/Asp187:OD1/His224:HD1		74.74
Asp187:OD2/His224:HD1	2.26	
Asp187:CG/Asp187:OD2/His224:HD1		97.02
Ace:O/Gln106:HN	1.88	
Ace:C/Ace:O/Gln106:HN		89.65
Ace:O/Thr40:HN	2.15	
Ace:C/Ace:O/Thr40:HN		122.51
Ace:O/Thr40:HG1	3.52	
Ace:C/Ace:O/Thr40:HG1		160.31

3.2. Molecular Docking

Quercetin, methyl-quercetin and acetyl-quercetin were docked against either acetyl-CALB or CALB, in order to study enzyme/substrates interactions during either quercetin acetylation or acetyl-quercetin alcoholysis. Electrostatic contribution was evaluated using various quercetin methylated forms that allowed suppressing locally interactions with hydroxyl groups. Docking of isoquercitrin (glycosylated form of quercetin) in acetyl-CALB was also performed in order to evaluate the effect of glycosylation on the orientations and the interactions of the flavonoid within the catalytic cavity. Results were compared with available experimental and theoretical data [4, 6]. For each docking run,

20 poses were retained; binding modes were examined and interactions between ligands and receptors were investigated.

3.2.1. Docking of quercetin and its methylated forms against acetyl CALB target

Quercetin was docked against acetyl-CALB in order to study interactions between the aglycone flavonoid and the acyl-enzyme during acetylation reaction. Results showed that quercetin adopted two orientations in the catalytic cavity, with either its A ring (CA conformation) or its B ring (CB conformation) pointing towards the enzyme core. The CB orientation was predominantly represented among docking poses. In this conformation, either 3'-OH or 4'-OH group established hydrogen bond interaction with the residue Asp134 located at the bottom of the cavity. 7-OH or 5-OH groups interacted with Glu188 residue, located at the cavity entrance. Whatever the pose, quercetin hydroxyl groups were quite far away from catalytic residues. In order to study the effect of the serine-bound acyl chain on the interactions between quercetin and CALB, docking was also performed against native CALB. In this case, only the CB orientation was observed. Interactions between quercetin hydroxyl groups and Asp134 and Glu188 residues were maintained.

The role of electrostatic interactions in lipase/quercetin binding modes was assessed by docking various quercetin methylated derivatives, that differ from each other by the number and the position of methyl and hydroxyl groups, against acetyl-CALB target. Depending on the number and the localization of methyl groups, either the CB or the CA orientation was mainly observed. The CB orientation was favoured in derivatives with 4'-OH group free and 7 or 3' or 3, 5, 7 positions locked by methyl groups. H-bond interaction was established between 4'-OH or 3'-OH group and Asp134 residue. In the case of 3'-methyl quercetin, the 4'-OH group was located between Ser105 and His224 residues while 7-OH and 5-OH groups interacted with Leu140 and Asp134, respectively. On the contrary, the CA orientation was predominant for the substrate 7, 3', 4'-methyl quercetin (Figure VI.7). When only the 3-OH and 5-OH groups were methylated, complexes corresponding to both orientations were obtained. In the CA orientation, two behaviours were observed depending on the presence or the absence of a free hydroxyl group at position 7. When the 7-OH group was free, an interaction was observed with the residue Asp134 while 3'-OH or 4'-OH groups interacted with Leu140 and Ser150 amino acids. When methylated, the 7 position was inserted between catalytic residues while the 5-OH group interacted with the residue Asp134. A same approach was applied to native CALB aiming to study the effect of the serine-bound acyl donor substrate on quercetin binding. In that case, the CB orientation was mainly observed (Figure VI.7). Interactions between free hydroxyl groups and Asp134 residue were formed as described before.

These results suggested that quercetin orientation strongly depended on the localization of free phenolic hydroxyl groups that can establish H-bond interactions with the polar residue Asp134 located at the cavity bottom. In order to verify if the nature of the flavonoid hydroxyl groups also influences its interactions with the residues constituting the catalytic pocket, the glycosylated form of quercetin, isoquercitrine, was docked against acetyl-CALB. In that case, the flavonoid first slipped its glucosidic part into the catalytic pocket (GLUC orientation) and formed H-bond interactions with the residue Asp134 and to a lesser extent with the residue Gln157. Hydroxyl groups of the flavan

backbone interacted with Leu140 located at the cavity entrance. In such a binding mode, 6'', 3'' or 4'' hydroxyl groups came near the catalytic residues. These results were partially in accordance with experimental data about isoquercitrin acetylation catalyzed by CALB. Indeed, Chebil et al. [4] reported the production of isoquercitrin 2'', 3'', 6''-triacetate, showing that only the sugar part of the flavonoid was acetylated. Docking simulations with CALB target led to a CA orientation of isoquercitrin, in which either the 7-OH or the 5-OH groups turned towards catalytic residues. Hydroxyl groups of the flavan part interacted with Asp134 residue while glycosidic hydroxyls interacted with Glu188 residue. These differences in substrate orientations depending on the target could be explained by changes in the active site polarity that depends on whether the catalytic serine is free or not.

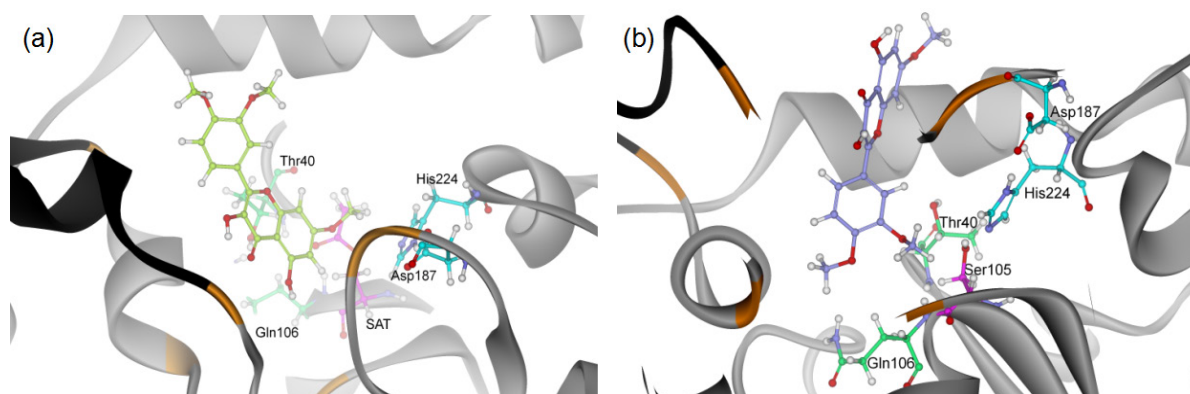


Figure VI.7. Models of complexes showing the substrate 7,3',4' – methyl quercetin in a CA orientation within acetyl-CALB target (a) and in a CB orientation within CALB target (b). (a): the 7-methyl group of the substrate pointed towards the catalytic residues while its 5-OH group interacts with Asp134 residue. (b): the 3'- and 4'-OH groups of the substrate pointed towards the cavity bottom while the 3- and 5-OH groups interact with the residues Ile189 and Glu188 located at the entrance cavity.

Docking of quercetin and isoquercitrin revealed the particular implication of the two polar residues Asp134 and Glu188 in substrate binding (Figure VI.8). More precisely, these residues interact with substrate hydroxyl groups through hydrogen bonds. Such interactions were shown to be independent on whether the catalytic serine is free or not.

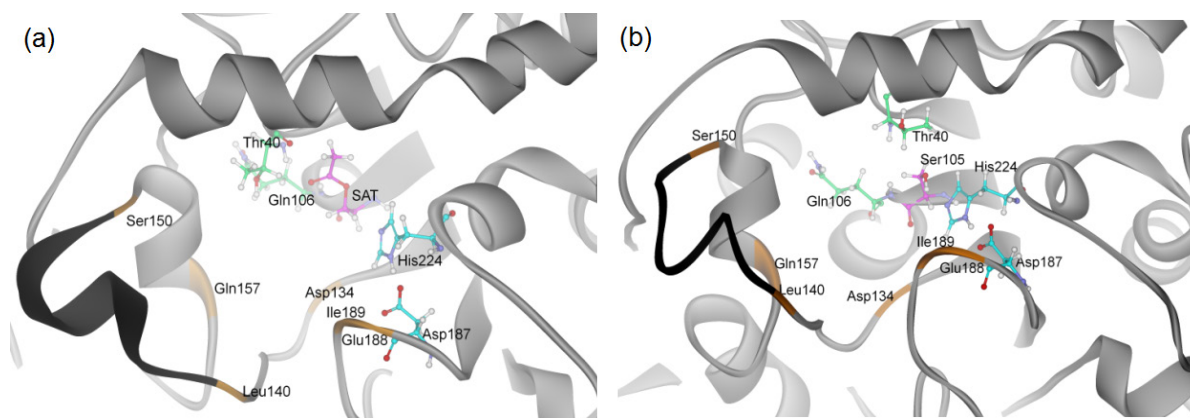


Figure VI.8. Models of acetyl-CALB (a) and CALB (b) targets. Residues constituting the oxyanion hole (Thr40 and Gln106) are shown in green; the catalytic triad is colored in blue (Asp187 and His224) and pink (SAT or Ser105); residues giving interactions with substrates are labelled and displayed in orange. The mobil $\alpha 5$ helix of CALB is shown in black.

3.2.2. Docking of quercetin acetyl derivatives against CALB target.

Triacetyl and pentaacetyl derivatives of quercetin were designed and docked against CALB target in order to mimic alcoholysis reaction. Docking of 7,3',4'-acetyl quercetin against CALB led to a major orientation with the B ring pointing towards the catalytic residues (CB orientation). Several poses were obtained with the flavonoid A ring facing the cavity bottom (CA orientation). When the flavonoid adopted the CB orientation, either the 4' position or the 3' position was turned to the oxyanion hole residues. 3-OH and 5-OH groups interacted with residues Glu188 and Ile189 located at the cavity entrance. In the CA orientation, 7-acetyl group was located at the bottom of the pocket and structural element 3-OH/4-keto/5-OH was inserted between Thr40 and Ser105 residues. An H-bond interaction between the carbonyl oxygen of the 3'-acetyl group and the residue Ser150 was observed. Orientation and proximity of quercetin acetyl groups towards oxyanion hole and catalytic residues suggested that the positions 4' and 3' were accessible and might be hydrolyzed. In the case of pentaacetyl-quercetin, docking led to an equitable repartition of the poses between CB and CA orientations. In the CB orientation, 4'- or 3'-acetyl group came near the oxyanion hole just as described before. In the CA orientation, 7-acetyl group came more or less close to the oxyanion hole residues Thr40 and Gln105 and the catalytic serine (Ser105) (Figure VI.8). Besides, little or no interactions were noticed in the presence of this substrate. These results are in accordance with experimental data reporting that alcoholysis of pentaacetyl-quercetin occurred on the 4'- and 3'-acetyl groups when using CALB [20].

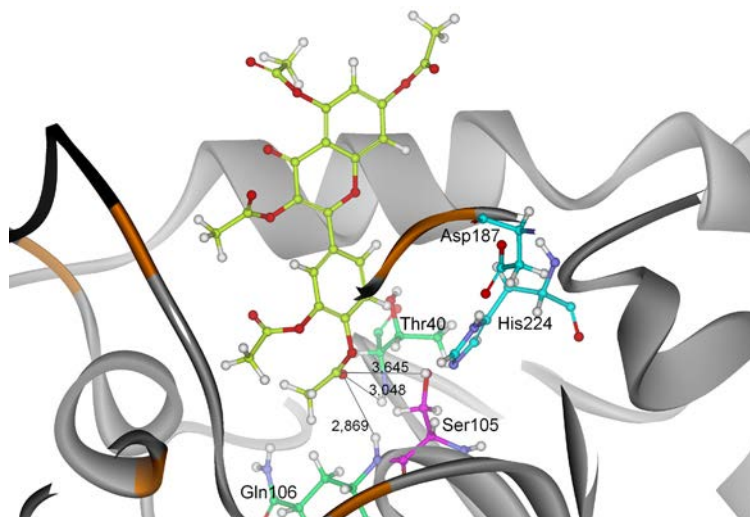


Figure VI.9. Example of pose resulting from the docking of pentaacetyl-quercetin against CALB and showing a model of complex involved in alcoholysis reaction. Catalytic distances (Å) are mentioned. Quercetin 4'-acetate group turns its carbonyl oxygen towards the oxyanion hole and places its electrophilic carbon atom near to the nucleophilic oxygen atom of the catalytic serine (Ser105).

3.3. Binding free energy calculations

In order to go deeper inside interaction understanding both in quercetin acetylation and acetyl-quercetin alcoholysis reactions, binding affinities of the flavonoid or its derivatives towards CALB or acetyl-CALB were studied. Relative binding free energies were determined and electrostatic and apolar contributions were examined more specifically. The same approach was applied to complexes formed with isoquercitrin aiming to compare interactions involving either phenolic hydroxyl groups or glycosidic hydroxyl groups.

3.3.1. Binding free energies of quercetin and methyl quercetin derivatives within acetyl-CALB target

Figure VI.9 shows the relative binding free energies (ΔG_{bind}) obtained for acetyl-CALB/quercetin complexes. Comparison was made in one hand with acetyl-CALB/isoquercitrin complexes and in the other hand with complexes obtained with CALB target.

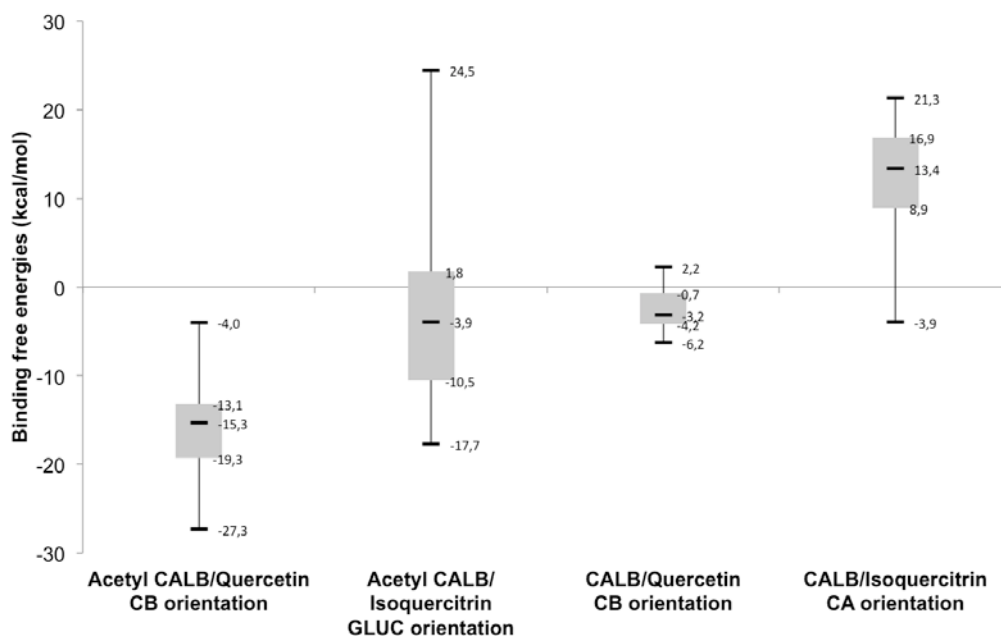


Figure VI.10. Distribution of binding free energies of acetyl-CALB/quercetin and acetyl-CALB/isoquercitrin complexes generated by docking. Comparison was made with binding free energies of complexes where the serine-bound acetate was removed. Binding free energies were calculated by the MM/PBSA method neglecting the entropic contribution.

Whatever the flavonoid docked within acetyl-CALB target, negative binding free energies were obtained, indicating favourable substrates/enzyme interactions. The highest binding affinities were observed with quercetin. Complexes involving isoquercitrin showed the greatest disparity in ΔG_{bind} values, due to the high flexibility of this molecule in comparison with quercetin.

Moving from acetyl-enzyme target to native enzyme target by removing the serine-bound acetate led to an increase in binding free energies both for quercetin and isoquercitrin complexes. Positive ΔG_{bind} values showed unfavourable binding between receptor and ligands. This result can be explained by the simultaneous presence of two competing nucleophilic centers facing one to the other (catalytic serine and flavonoid hydroxyl groups facing together).

According to data shown in Table VI.3, whatever the nature of the flavonoid, its orientation and the target, both intermolecular van der Waals (ΔE_{vdW}) and electrostatic (ΔE_{elec}) interactions provided major favourable contributions for substrate binding whereas polar solvation interactions were unfavourable. In both targets, ΔG_{APOLAR} contributions favoured ligand binding and were quite similar for one given ligand. Whatever the ligand, positive polar terms (ΔG_{POLAR}) weigh less in acetyl-CALB target than in CALB target showing a better affinity of ligands for the acetyl-enzyme than for the native lipase. This result could be explained by a change in the catalytic pocket polarisation when moving from the native

state to the acyl-enzyme state of the lipase. This is expected to help for the nucleophilic attack of the acyl-enzyme by the acyl-acceptor substrate. In the case of isoquercitrin, ΔG_{POLAR} contribution was much more important than in the case of quercetin. This difference was mainly due to the $\Delta G_{\text{solvation}}^{\text{polar}}$ contribution that is much higher for the glycosylated flavonoid than for the aglycone one. This result could be explained by the fact that quercetin was more buried within the cavity than isoquercitrin. Depending on the ligand and the target, the binding free energy entropic contribution showed no significant difference due to the structural similarity of the systems under study.

Table VI.3. Average and standard deviation of binding free energy contributions for complexes involving quercetin and isoquercitrin. Only major ligand orientations determined by previous docking were taken into account. ΔE_{vdW} , ΔE_{elec} , $\Delta G_{\text{solvation}}^{\text{polar}}$, $\Delta G_{\text{solvation}}^{\text{nonpolar}}$, $-\Delta S$, were intermolecular van der Waals contribution, intermolecular electrostatic contribution, polar solvation energy, nonpolar solvation energy, entropic contribution, respectively. All energies are expressed in kcal.mol⁻¹.

Complex								
Target	Ligand and major orientation	ΔE_{vdW}	$\Delta G_{\text{solvation}}^{\text{nonpolar}}$	ΔG_{APOLAR}	ΔE_{elec}	$\Delta G_{\text{solvation}}^{\text{polar}}$	ΔG_{POLAR}	$-\Delta S$
Acetyl-CALB	Quercetin	-23.8	-4.4	-28.2	-32.5	44.9	12.4	21.3
	<i>CB orientation</i>	± 4.2	± 0.1		± 6.1	± 6.2		± 0.0
	Isoquercitrin	-40.8	-5.6	-46.4	-31.5	74.5	43.0	22.4
	<i>GLUC orientation</i>	± 2.9	± 0.1		± 7.6	± 11.1		± 0.0
CALB	Quercetin	-22.7	-4.5	-27.5	-39.9	64.8	24.9	21.3
	<i>CB orientation</i>	± 0.9	± 0.0		± 4.5	± 4.6		± 0.0
	Isoquercitrin	-38.1	-5.7	-43.8	-45.9	102.1	56.2	22.4
	<i>CA orientation</i>	± 3.7	± 0.1		± 9.1	± 10.3		± 0.0

Quercetin methylated derivatives were designed and docked in acetyl-CALB and native CALB aiming to evaluate the role of phenolic OH groups in quercetin binding affinity. Binding free energies, polar and apolar contributions are reported in Figure VI.11 and Table VI.4, respectively. When docked within acetyl-CALB target, methyl-quercetin derivatives led to predominantly negative ΔG_{bind} values indicating favourable binding between the ligands and the acetyl-enzyme. Moreover, even if negative, ΔG_{bind} values were predominantly higher than those obtained with native quercetin, except for 7-methyl quercetin. This result suggests that methylation decreases the affinity of quercetin towards acetyl-CALB catalytic pocket. Highest values were observed for the 3,5-dimethyl quercetin in the CB orientation that established no specific interactions with acetyl-CALB residues according to docking results. The following affinity order could be established as 7-methyl > 3,5,7-methyl > 3,5-methyl *in CA orientation* > 7,3',4'-methyl > 3'-methyl > 3,5-methyl *in CB orientation*. Otherwise, no correlation could be found between binding affinity and the position or the number of quercetin free hydroxyl groups. ΔG_{APOLAR} contributions and more specifically van der Waals contributions were more important in the case of methyl quercetin derivatives than those determined for native quercetin. Conversely, electrostatic contribution decreased when replacing quercetin hydroxyl groups by methyl groups. This expected result can be explained by a decrease in hydrogen bond interactions in the case of methyl-quercetin comparing with quercetin. This trend was particularly observed for the 7,3',4'-trimethyl

quercetin where flavan structure was free of hydroxyl groups at its extremities. In this case, only one H-bond interaction between free hydroxyl groups and the residue Asp134 was noticed.

Just as shown before, positive values were obtained in the case of CALB target, whatever the ligand. Furthermore, ΔG_{bind} values were 2 to 4 times higher than those obtained with quercetin, showing an unfavourable effect of methylation on quercetin binding. Apolar interactions gained intensity when suppressing free hydroxyl groups but that tendency was compensated by the decrease of favourable electrostatic interactions.

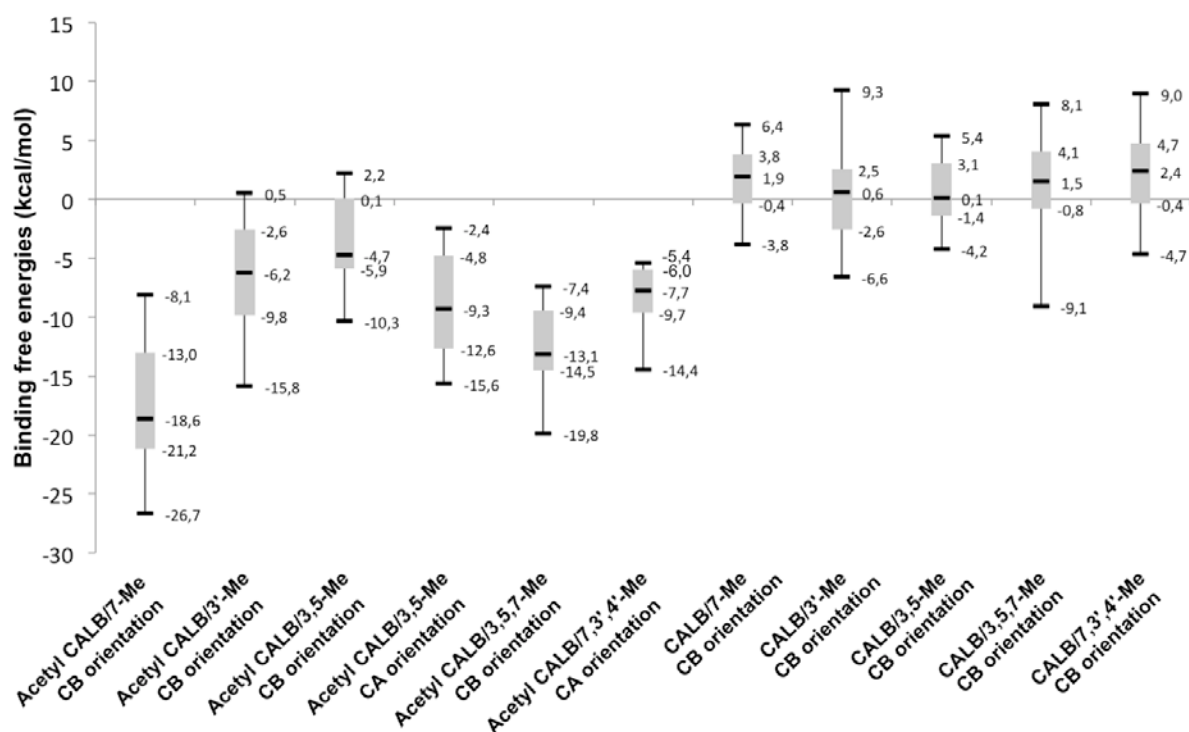


Figure VI.11. Distribution of binding free energies of complexes involving methyl quercetin derivatives within acetyl-CALB or CALB catalytic pocket. Binding free energies were calculated by the MMPBSA method neglecting the entropic contribution.

Table VI.4. Average and standard deviations of binding free energy contributions for complexes involving either acetyl-CALB or CALB target and methyl quercetin derivatives. Only major ligand orientations determined by previous docking were taken into account ΔE_{vdW} , ΔE_{elec} , $\Delta G_{\text{solvation}}^{\text{polar}}$, $\Delta G_{\text{solvation}}^{\text{nonpolar}}$, $-T\Delta S$, were intermolecular van der Waals contribution, intermolecular electrostatic contribution, polar solvation energy, nonpolar solvation energy, entropic contribution, respectively. All energies are expressed in $\text{kcal}\cdot\text{mol}^{-1}$.

Complex		ΔE_{vdW}	$\Delta G_{\text{solvation}}^{\text{nonpolar}}$	ΔG_{APOLAR}	ΔE_{elec}	$\Delta G_{\text{solvation}}^{\text{polar}}$	ΔG_{POLAR}	$-T\Delta S$
Target	Ligand and major orientation							
Acetyl-CALB	7-methyl quercetin CB orientation	-27.4 ± 3.7	-4.7 ± 0.1	-32.1	-24.1 ± 3.8	38.9 ± 5.1	14.8	21.5 ± 0.0
	3'-methyl quercetin CB orientation	-35.5 ± 1.0	-4.9 ± 0.0	-40.9	-21.0 ± 3.1	54.4 ± 5.2	33.4	21.5 ± 0.0
	3,5-methyl quercetin CA orientation	-29.1 ± 2.2	-5.1 ± 0.0	-34.2	-31.5 ± 4.4	56.8 ± 4.9	25.3	21.5 ± 0.0
	3,5-methyl quercetin CB orientation	-26.8 ± 2.7	-4.9 ± 0.2	-31.7	-27.4 ± 4.9	55.8 ± 5.7	28.4	21.5 ± 0.0

	3,5,7-methyl quercetin <i>CB orientation</i>	-29.1 ± 2.8	-5.1 ± 0.2	-34.2	-25.6 ± 2.4	47.0 ± 4.5	21.4	21.7 ± 0.0
	7,3',4'-methyl quercetin <i>CA orientation</i>	-34.2 ± 1.5	-5.2 ± 0.1	-39.4	-13.5 ± 2.0	44.5 ± 2.3	31.0	21.7 ± 0.0
CALB	7-methyl quercetin <i>CB orientation</i>	-29.5 ± 2.0	-4.8 ± 0.0	-34.3	-26.3 ± 5.2	62.6 ± 4.6	36.3	21.5 ± 0.0
	3'-methyl quercetin <i>CB orientation</i>	-28.9 ± 2.4	-4.7 ± 0.1	-33.6	-31.9 ± 2.8	65.8 ± 3.4	33.9	21.5 ± 0.0
	3,5-methyl quercetin <i>CB orientation</i>	-28.0 ± 1.6	-5.0 ± 0.1	-33.0	-35.2 ± 7.9	68.9 ± 6.0	33.7	21.5 ± 0.0
	3,5,7-methyl quercetin <i>CB orientation</i>	-33.2 ± 2.8	-5.2 ± 0.1	-38.4	-17.6 ± 7.0	56.9 ± 3.9	39.3	21.7 ± 0.0
	7,3',4'-methyl quercetin <i>CB orientation</i>	-32.7 ± 4.2	-5.1 ± 0.0	-37.8	-20.5 ± 2.8	60.3 ± 7.0	39.8	21.7 ± 0.0

3.3.2. Binding free energies of acetyl quercetin derivatives within CALB target

Binding free energies were evaluated in complexes miming acetyl quercetin alcoholysis catalyzed by CALB. Two ligands were studied corresponding to 7,3',4'-triacyl quercetin and 3,5,7,3',4'-pentaacyl quercetin. Relative binding free energies calculated by the MM-PBSA method, polar and apolar contributions are shown in Figure VI.12 and Table VI.5. Whatever the ligand and its orientation in CALB catalytic pocket, negative ΔG_{bind} values were obtained, indicating favourable binding affinity. Similar values were obtained for the two ligands under study. Substitution of quercetin hydroxyl groups by acetyl groups led to a decrease in ΔG_{bind} values when comparing results obtained for quercetin/CALB complexes. In addition, the presence of acetate groups on quercetin structure greatly decreased favourable electrostatic contribution whereas van der Waals contribution was about 2 times stronger.

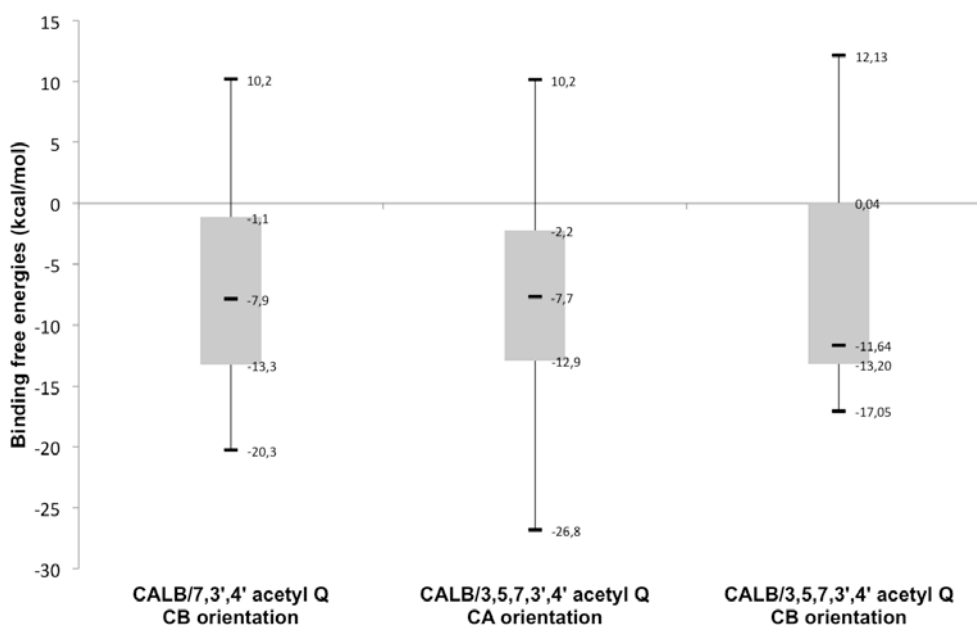


Figure VI.12. Distribution of binding free energies of complexes associating acetyl quercetin derivatives and CALB catalytic pocket. Energies were calculated by the MM/PBSA method neglecting the entropic contribution.

Table VI.5. Average and standard deviation of binding free energy contributions for acetyl quercetin/CALB complexes. Only major orientations of the ligands were taken into account ΔE_{vdW} , ΔE_{elec} , $\Delta G_{solvation}^{polar}$, $\Delta G_{solvation}^{nonpolar}$, $-\Delta S$, were intermolecular van der Waals contribution, intermolecular electrostatic contribution, polar solvation energy, nonpolar solvation energy, entropic contribution, respectively. All energies are expressed in kcal.mol⁻¹.

Complex								
Target	Ligand and major orientation	ΔE_{vdW}	$\Delta G_{solvation}^{nonpolar}$	ΔG_{APOLAR}	ΔE_{elec}	$\Delta G_{solvation}^{polar}$	ΔG_{POLAR}	$-\Delta S$
CALB	7,3',4'-acetyl CB orientation	-38.7 ±1.6	-5.6 ±0.2	-44.3	-18.4 ±5.3	56.1 ±6.1	37.7	22.3 ±0.0
	3,5,7,3',4'-acetyl CA orientation	-40.3 ±5.0	-6.2 ±0.2	-46.5	-8.5 ±10.2	47.1 ±10.4	38.6	22.8 ±0.0
	3,5,7,3',4'-acetyl CB orientation	-42.4 ±2.8	-6.3 ±0.1	-48.7	-16.8 ±6.5	59.2 ±12.5	42.4	22.8 ±0.0

4. DISCUSSION

In this study, quercetin and its methylated or acetylated derivatives were docked against CALB and acetyl-CALB target aiming to investigate the interactions between phenolic groups of the flavonoid and the residues constituting the active site. Moreover, the influence of the catalytic state of the catalytic serine on docking results as well as on binding free energy was investigated. Methylation of quercetin was shown to influence quercetin positioning within acetyl-CALB catalytic pocket. The presence of a free 4'-OH group mainly favoured the CB orientation of quercetin whereas a reversal position of the substrate was observed for 4'-O-methyl derivatives. Electrostatic terms (ΔE_{elec}) are directly reliable to hydrogen bond establishment in ligand/receptor binding [21]. In this study, substitution of quercetin -OH groups by -CH₃ groups led to a decrease of intermolecular electrostatic interactions that could be explained by a decrease in the number of hydrogen bonds between quercetin and surrounding residues. This trend was particularly marked in the case of the derivative 7, 3', 4'-methyl quercetin, showing the high contribution of the 7, 3', and 4'-OH groups to the electrostatic stabilization of quercetin. A similar result was reported by Khuntawee et al. [22] who reported that the 3'- and 4'-OH groups of flavonoids increase the binding efficiency of cyclin dependent kinase. Moreover, electrostatic and vdW interactions were shown to strongly contribute to substrate binding. In the present work, predominant H-bond interaction between quercetin free hydroxyl groups and the residue Asp134 was noticed. Similar observations were reported by Codomiu-Hernandez et al. [23] in a semi-empirical study about flavonoid-protein interactions, A theoretical affinity order was proposed, depending on flavonoid structure and amino acid residues. The authors concluded that the residues Lys, Glu and Asp have a strong affinity towards flavonoid compounds, especially if they are under a charged form.

In the case of models miming acetyl quercetin alcoholysis, electrostatic interactions were found to be reduced while van der Waals energies increased comparing with models miming acylation reactions. Only weak H-bonds interactions were established between the carbonyl oxygen atom of quercetin acetate groups and surrounding residues. The following affinity order was observed: $\Delta G_{bind} 4'Ac < \Delta G_{bind} 7Ac < \Delta G_{bind} 3'Ac$ for pentaacetyl Q and $\Delta G_{bind} 4'Ac < \Delta G_{bind} 3'Ac$ for triacetyl Q. These results are partly consistent with experimental data. In fact, Lambusta et al. [20] reported the

hydrolysis of pentaacetyl quercetin catalyzed by CALB and showed that the 3'Ac and the 4'Ac positions were preferentially hydrolyzed.

Finally, this work allowed showing a qualitative relationship between the chemical reactivity of the systems and binding free energies. Correct associations between targets and substrates appeared to be favoured whether for acylation or alcoholysis reactions. Thus, highest free binding energies were observed for complexes involving acetyl-CALB target and substrates that are expected to be acylated. In the same way, association of CALB target with substrates that are subjected to hydrolysis seemed to be favoured.

5. REFERENCES

1. Xiao, Z.-P., et al., *Flavonoids health benefits and their molecular mechanism* Mini-Review in Medicinal Chemistry, 2011. **11**: p. 169-177.
2. Boots, A.W., G.R.M.M. Haenen, and A. Bast, *Health effects of quercetin: from antioxidant to nutraceutical*. European Journal of Pharmacology, 2008. **585**: p. 325-337.
3. Mendoza, E.E. and R. Burd, *Quercetin as a systemic chemopreventative agent: structural and functional mechanisms*. Mini-Review in Medicinal Chemistry, 2011. **11**: p. 1216-1221.
4. Chebil, L., et al., *Enzymatic acylation of flavonoids : effect of the nature of the substrate, origin of lipase and operating conditions on conversion yield and regioselectivity*. Journal of Agricultural and Food Chemistry, 2007. **53**(23): p. 9496-9502.
5. Stevenson, D.E., et al., *Direct acylation of flavonoid glycosides with phenolic acids catalyzed by *Candida antarctica* lipase B (Novozyme 435)*. Enzyme and Microbial Technology, 2006. **39**: p. 1236-1241.
6. De Oliveira, E.B., et al., *A molecular modelling study to rationalize the regioselectivity in acylation of flavonoid glycosides catalyzed by *Candida antarctica* lipase B*. Journal of Molecular Catalysis B: Enzymatic, 2009. **59**: p. 96-105.
7. De Oliveira, E.B., et al., *An approach based on density functional theory (DFT) calculations to assess the *Candida antarctica* lipase B selectivity in rutin, isoquercitrin and quercetin acetylation*. Journal of Molecular Catalysis B: Enzymatic, 2010. **66**: p. 325-331.
8. Syrén, P.O., et al., *Increased activity of enzymatic transacylation of acrylates through rational design of lipases*. Journal of Molecular Catalysis B: Enzymatic, 2010. **65**(1-4): p. 3-10.
9. Rastelli, G., et al., *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA*. Journal of Computational Chemistry, 2010. **31**(4): p. 797-810.
10. Uppenberg, J., et al., *The sequence, crystal structure determination and refinement of two forms of lipase B from *Candida antarctica**. Structure, 1994. **2**: p. 293-308.
11. Botta, B., et al., *Lipase-catalyzed regioselective acylation of resorcin[4]arenes*. Journal of Molecular Catalysis B: Enzymatic, 2002. **16**: p. 241-247.
12. Palocci, C., et al., *An approach to address *Candida rugosa* lipase regioselectivity in the acylation reactions of trytilated glucosides*. Journal of Biotechnology, 2007. **128**: p. 908-918.
13. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems*. Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
14. Ryckaert, J.P., G. Ciccotti, and J.C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes*. Journal of Computational Physics, 1977. **23**(3): p. 327-341.
15. Gupta, R., N. Gupta, and P. Rathi, *Bacterial lipases: an overview of production, purification and biochemical properties*. Applied Microbiology and Biotechnology, 2004. **64**(6): p. 763-781.
16. Lee, M.S., F.R. Salsbury, and C.L. Brooks III, *Novel generalized Born methods*. Journal of Chemical Physics, 2002. **116**(24): p. 10606-10614.
17. Venkatachalam, C.M., et al., *LigandFit : a novel method for the shape-directed rapid docking of ligands to protein active sites*. Journal of Molecular Graphics and Modelling, 2003. **21**: p. 289-307.
18. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models*. Accounts of Chemical Research, 2000. **33**(12): p. 889-897.
19. Skjøt, M., et al., *Understanding the plasticity of the a/b-hydrolase fold: lid swapping on the *Candida antarctica* lipase B results in chimeras with interesting biocatalytic properties*. ChemBioChem, 2009. **10**: p. 520-527.

20. Lambusta, D., et al., *Application of lipase catalysis in organic solvent for selective protection-deprotection of bioactive compounds*. Journal of Molecular Catalysis B: Enzymatic, 2003. **22**: p. 271-277.
21. Panigrahi, S.K. and G.R. Desiraju, *Strong and weak hydrogen bonds in the protein–ligand interface*. Proteins: Structure, Function, and Bioinformatics, 2007. **67**(1): p. 128-141.
22. Khuntawee, W., T. Rungrotmongkol, and S. Hannongbua, *Molecular Dynamic Behavior and Binding Affinity of Flavonoid Analogues to the Cyclin Dependent Kinase 6/cyclin D Complex*. Journal of Chemical Information and Modeling, 2011. **52**(1): p. 76-83.
23. Codorniu-Hernández, E., et al., *Essential amino acids interacting with flavonoids: A theoretical approach*. International Journal of Quantum Chemistry, 2005. **103**(1): p. 82-104.

3. Contribution de l'article

Dans cette étude, la compréhension de la contribution des groupements hydroxyle de la quercétine dans les interactions du flavonoïde avec le site actif de la CALB a été améliorée grâce à une approche de docking et de calcul d'énergie libre de liaison. Pour la quercétine et ses dérivés méthylés ou acétylés, deux orientations majeures du flavonoïde ont été observées, correspondant à une insertion soit par le cycle A, soit par le cycle B dans le site actif de la CALB. Ces orientations sont désignées respectivement par CA et CB. L'isoquercitrine, quant à elle, introduit préférentiellement sa partie glucosidique dans la cavité catalytique (orientation GLUC). L'analyse des complexes issus du docking révèle que:

- la méthylation de la quercétine influence l'orientation du flavonoïde dans le site actif de l'acyl-CALB. Plus précisément, l'orientation CB de la quercétine est favorisée lorsque l'hydroxyle en position 4' est libre, tandis qu'une orientation inverse est observée pour les dérivés méthylés dans cette position ; la présence de groupements méthyles sur les positions 3 et 5 conduit aux deux orientations CA et CB de la quercétine. Une analyse plus poussée des interactions au sein des complexes a permis d'identifier l'Asp134 localisé au fond de la poche comme un résidu clé établissant des liaisons hydrogène avec tous les substrats.
- l'état libre ou acylé de la chaîne latérale de la sérine catalytique influence les positions et orientations des substrats. Ainsi, les dérivés méthylés de la quercétine adoptent une unique orientation CB lorsqu'ils sont dockés dans le site actif de la CALB native (sérine catalytique libre). Dans le cas de l'isoquercitrine, différentes orientations sont mises en évidence dépendant de la cible de docking utilisée. Une orientation GLUC est observée dans la cavité de l'acétyl-CALB tandis qu'une orientation CA est adoptée dans la cible CALB.
- dans le cas des substrats acétylés de la quercétine, les deux orientations CA et CB sont obtenues à l'issue du docking dans le site actif de la CALB. Dans ce cas, les groupements acétate des dérivés de la quercétine se positionnent à proximité des résidus du trou oxyanionique et de la sérine catalytique. Très peu d'interactions (voire aucune interaction pour certains substrats) sont établies entre les groupements acétate susceptibles d'être hydrolysés et les résidus du site actif. Seuls les groupes hydroxyles présents dans la structure des dérivés interagissent avec les résidus de la cavité catalytique de la CALB.

Les poses générées lors de la procédure de docking ont été soumises à un calcul d'énergie libre de liaison. Les résultats obtenus montrent que :

- la quercétine a une affinité de liaison supérieure pour la cavité de l'acétyl-CALB par rapport à l'isoquercitrine.
- la méthylation de la quercétine diminue l'affinité de liaison de l'acétyl-CALB pour le flavonoïde. Toutefois, aucune corrélation entre la position ou le nombre de groupements hydroxyle libres de la quercétine et l'affinité de liaison n'a pu être mise en évidence. Par ailleurs, la méthylation de la quercétine diminue la contribution apolaire et augmente la contribution électrostatique de l'énergie de liaison.

- dans le cas de la quercétine, ses dérivés méthylés et l'isoquercitrine, la forme acyl-enzyme de CALB conduit à des affinités de liaison supérieures par rapport à la CALB libre.
- l'acétylation de la quercétine accroît l'affinité de liaison de la CALB native pour le flavonoïde. En outre, le remplacement de groupements hydroxyle de la quercétine par des groupements acétate conduit à une réduction des interactions électrostatiques et à un renforcement de la contribution apolaire de l'énergie de liaison.

L'ensemble de ces résultats apporte un niveau de compréhension supplémentaire sur les interactions établies entre la quercétine et la cavité catalytique de la CALB :

- les interactions électrostatiques intermoléculaires entre la quercétine et le site actif de la CALB prédominent sur les interactions hydrophobes.
- dans les modèles mimant l'alcoolyse de dérivés acétylés de la quercétine, il apparaît que les groupements acétate des positions 7, 3' et 4' se positionnent idéalement vis-à-vis du trou oxyanionique et des résidus catalytiques, laissant présager la possibilité d'une hydrolyse. Dans ce cas, les groupements carbonyle des dérivés établissent des liaisons hydrogène avec les résidus Thr40 et Gln106 constituant le trou oxyanionique. Par ailleurs, le carbone électrophile des groupements acétate s'approchent à une distance inférieure à 4 Å de la sérine catalytique. Les conditions permettant l'alcoolyse des dérivés acétylés de la quercétine semblent donc réunies. Ces résultats sont en accord avec les données expérimentales extraites de la littérature.

Conclusion Générale – Perspectives

Au cours de ce travail, le potentiel catalytique de la lipase B de *Candida antarctica* (CALB) a été investigué dans le cadre de la réaction d'acétylation de la quercétine. Tout d'abord, une clarification expérimentale a été apportée sur l'activité enzymatique de la CALB envers le flavonoïde aglycone. Afin de comprendre la spécificité de la CALB dans cette réaction, le travail s'est poursuivi à une échelle moléculaire en combinant un protocole de docking et de dynamique moléculaire (DM). En se basant sur le mécanisme catalytique admis pour les lipases dans la réaction de transestérification, les résultats des simulations ont été analysés afin de trouver d'éventuelles corrélations entre les positions, orientations, interactions et stabilité des substrats, et les résultats expérimentaux. En outre, une procédure similaire a été exécutée pour déterminer l'impact de résidus particuliers sur la structure de cette lipase et la réaction ciblée. Dans ce cas, les résultats collectés lors des simulations de DM de l'enzyme mutée, en présence et en absence de la quercétine, ont permis d'analyser la structure protéique des mutants et le comportement dynamique du flavonoïde dans le site actif des variants. Finalement, l'orientation et l'affinité de liaison de la quercétine ainsi que de ses dérivés ont été évaluées par simulation de docking et par calcul d'énergie libre de liaison.

Les principaux résultats obtenus sont synthétisés sous forme de réponses aux questions soulevées au début de ce travail :

(1) La lipase B de *Candida antarctica* (CALB) catalyse-t-elle ou non l'acétylation de la quercétine avec l'acétate de vinyle comme donneur d'acyle (DA) ?

Les données expérimentales obtenues confirment les résultats antérieurs du Laboratoire sur l'inactivité catalytique de la CALB dans la réaction d'acétylation de la quercétine (QCT).

Une conversion de 90 % de la quercétine est observée après 72 h de réaction lorsque l'acétylation de celle-ci est menée en absence de la CALB dans un milieu réactionnel, constitué d'acétonitrile et d'un fort excès d'acétate de vinyle (DA/QCT = 330), maintenu à une température de 60°C. Cette acétylation chimique de la quercétine est relativement régiosélective. En effet, le produit majoritaire identifié est le monoacétate de quercétine. En plus, une faible quantité de diacétate de quercétine et des traces de tri- et de tétra-acétate de quercétine sont également détectées dans le milieu.

En présence de la lipase dans le milieu réactionnel, de faibles différences sont constatées après 72 h de réaction. Le pourcentage de quercétine convertie s'élève à 80 %, ce qui est légèrement inférieure à la réaction précédente. Le produit prédominant reste le monoacétate de quercétine tandis que les traces de tri- et de tétra-acétate de quercétine ont disparu.

A l'issue de ce travail expérimental, on en conclut que la réaction d'acétylation de la quercétine observée en présence d'un large excès d'acétate de vinyle est essentiellement de nature chimique. La CALB ne présente aucune activité catalytique d'acétylation vis-à-vis de la quercétine mais une faible activité d'hydrolyse en utilisant les traces d'eau présentes dans le milieu.

- (2) Comment les modes d'interactions enzyme-substrat peuvent-ils expliquer la spécificité de la lipase B de *Candida antarctica* et de la lipase de *Pseudomonas cepacia* (PCL) dans la réaction d'acétylation enzymatique du flavonoïde quercétine ? Peut-on expliquer au niveau moléculaire les différences d'activité d'acétylation entre la CALB et la PCL ? Quels sont les résidus étroitement impliqués dans le positionnement de la quercétine dans la CALB ?**

Les modèles obtenus dans cette étude indiquent que les différences structurales observées entre la PCL et la CALB au niveau des acides aminés tapissant leurs cavités catalytiques, donnent lieu à des interactions intermoléculaires enzyme-quercétine différentes selon la lipase employée. Le flavonoïde aglycone est stabilisé par des interactions formées avec les résidus de la cavité de la CALB. En revanche dans le cas de la PCL, le substrat est soumis à des fluctuations positionnelles par les interactions établies avec les acides aminés du site actif. Ces différences de stabilisation des substrats seraient à l'origine de la différence de spécificité de ces deux lipases vis-à-vis de la quercétine.

Les modèles montrent que la différence d'activité constatée expérimentalement entre la CALB et la PCL s'explique (au niveau moléculaire) par l'orientation de l'acétate et la proximité des groupements hydroxyles de la quercétine face aux résidus catalytiques. L'origine de l'inactivité de la CALB dans la réaction d'acétylation de la quercétine proviendrait du positionnement inadéquat des deux substrats, donneur d'acyle et flavonoïde. Au contraire, l'activité de la PCL est corrélée à une orientation appropriée de l'acétate et à un rapprochement suffisant des fonctions hydroxyles du flavonoïde près des résidus catalytiques.

Les éléments structuraux de la cavité catalytique de la CALB qui empêchent le flavonoïde aglycone d'accéder aux résidus catalytiques sont d'une part, les chaînes latérales de deux résidus hydrophobes, l'isoleucine 189 et l'isoleucine 285 et d'autre part, deux résidus polaires, l'aspartate 134 et la glutamine 157.

Des complexes enzymes-substrat ont été produits à partir d'une procédure de docking et de scoring appliquée au flavonoïde aglycone dans les sites actifs de la CALB et de la PCL acétylées. Ces complexes ont été soumis à des simulations de DM pour évaluer leur stabilité structurale sur une trajectoire de 10 ns. L'analyse de ces simulations a été basée sur le mécanisme réactionnel admis pour les lipases, en prenant en compte des critères structuraux établis dans des travaux similaires. Pour que ces complexes soient considérés comme productifs, c'est-à-dire susceptibles de mener à la formation d'un produit réactionnel, deux conditions doivent être respectées : (i) le carbonyle (-C=O) de l'acétate doit être correctement orienté vers les résidus du trou oxyanionique de la lipase pour favoriser l'attaque nucléophile du flavonoïde sur l'atome de carbone carbonyle ; (ii) un hydroxyle (-OH) de la quercétine doit être positionné entre les résidus catalytiques. Dans cette position, le proton de l'hydroxyle peut être capté par l'histidine, et l'oxygène ainsi déprotoné peut exécuter l'attaque nucléophile sur l'acétate de l'acyle enzyme.

Pour les deux lipases CALB et PCL, deux orientations possibles de la quercétine ont été trouvées à partir des résultats de docking : soit le cycle A du flavonoïde aglycone s'introduit vers le fond des cavités catalytiques des enzymes, soit c'est le cycle B. La production des trajectoires pour les complexes lipase-substrat ont permis d'observer que :

- dans le cas de la CALB, l'acétate n'est globalement pas orienté vers les résidus du trou oxyanionique. Cela se traduit par l'absence de liaisons hydrogène entre le carbonyle ($-C=O$) et, les résidus Thr40 et Gln106 qui constituent le trou oxyanionique de cette lipase. De plus, les groupements hydroxyles sont trop éloignés des résidus catalytiques Ser105 et His224 pour mener à la formation d'un produit réactionnel. Cela est occasionné d'une part, par les résidus Ile189 et Ile285 formant des interactions hydrophobes de type $CH_3-\pi$ avec les cycles aromatiques de la quercétine ; et d'autre part, par les résidus Asp134 et Gln157 interagissant avec les groupements hydroxyles de la quercétine via des liaisons hydrogène. De plus, la rigidité de la molécule de quercétine et l'étroitesse du site actif de la CALB sont deux obstacles complémentaires qui empêchent le rapprochement du flavonoïde des résidus catalytiques.
- dans le cas de la PCL, l'acétate adopte une orientation adéquate vers les résidus du trou oxyanionique (Leu17 et Gln88) avec l'établissement des liaisons hydrogène requises pour maintenir cette position. En outre, tous les groupements hydroxyles de la quercétine se rapprochent suffisamment des résidus catalytiques, Ser87 et His286, pour former des complexes dits productifs. Les différents éléments favorisant un positionnement approprié de la quercétine sont : (i) des interactions électrostatiques avec les résidus polaires Thr18, Tyr23, Tyr29 ; (ii) des interactions hydrophobes avec les résidus aliphatiques tapissant le site actif qui guident l'acheminement du substrat aux résidus catalytiques ; et (iii) l'accès plus large de la cavité aux résidus catalytiques pour le substrat.

(3) Quelle est l'influence de la mutation des résidus Ile189 et Ile285 sur la structure de la CALB, et sur le positionnement des substrats acétate et quercétine dans son site actif ?

Les mutations ponctuelles des résidus Ile189 et Ile285 de la CALB par les résidus alanine et valine n'entraînent pas de réarrangement majeur de la structure protéique. En outre, une augmentation significative du volume est constatée lors de la substitution des isoleucines par une alanine. Dans les modèles de la CALB mutée, l'orientation du substrat acétate n'est pas appropriée et le rapprochement des groupements hydroxyles face aux résidus catalytiques est insuffisant. Par ailleurs, une mobilité accrue de la quercétine est observée ainsi qu'un renforcement du nombre de liaisons hydrogène.

La structure protéique des quatre mutants simples de la CALB, Ile189Ala, Ile189Val, Ile285Ala et Ile285Val, a été étudiée sur des trajectoires de DM de 20 ns. Le comportement dynamique de ces variants est similaire à celui de l'enzyme sauvage. De plus, le remplacement des isoleucines par une alanine entraîne un accroissement significatif du volume du site actif. L'application d'un protocole de docking/scoring sur les structures mutées, issues des trajectoires de DM, a permis de générer et de

sélectionner des complexes variant-quercétine. Dans les mutants de la CALB, la quercétine adopte quatre types d'orientation : soit elle s'insère avec le cycle A ou le cycle B vers le cœur de l'enzyme, soit elle s'introduit horizontalement avec son cycle A ou son cycle B pointant vers l'hélice $\alpha 5$.

Les dynamiques moléculaires de 10 ns, qui ont ensuite été calculées pour les complexes lipase mutée-quercétine, indiquent que la présence du flavonoïde aglycone dans les sites actifs des mutants n'affecte pas la structure protéique. La réduction des chaînes latérales des isoleucines entraîne d'une part, des fluctuations positionnelles de la quercétine et d'autre part, une augmentation du nombre de liaisons hydrogène avec les résidus Thr40, Asp134, Thr138, Ser150, Gln157, Asp187 et Glu188. En outre, le calcul d'énergie libre de liaison effectué via la méthode MM-PBSA a montré que la quercétine a une affinité de liaison pour les mutants supérieure à celle de la lipase sauvage. Ce renforcement d'affinité est lié à un accroissement de la contribution électrostatique (liaison hydrogène) entre le flavonoïde et les résidus du site actif des mutants.

Dans les complexes mutés à la position 189, les liaisons hydrogène d'une part, entre les résidus de la triade catalytique et d'autre part, entre l'acétate et les résidus du trou oxyanionique sont formées sur au moins 50 % de la trajectoire. Lorsque la substitution est réalisée sur le résidu Ile285, ces liaisons hydrogène sont maintenues sur un faible pourcentage de la trajectoire. En associant ces deux critères (liaisons hydrogène entre les résidus de la triade catalytique et, entre l'acétate et les résidus du trou oxyanionique), deux complexes mutés ont sur un quart de leur trajectoire une orientation correcte de la machinerie catalytique. Pour ces deux complexes, les groupements hydroxyles de la quercétine restent en moyenne à une distance supérieure à 4 Å (distance seuil fixée) des résidus catalytiques. Cela implique que le transfert de proton de la quercétine à l'histidine catalytique et l'attaque nucléophile de la quercétine sur l'acyle-enzyme ne peuvent pas être envisagés.

D'après les résultats issus de la modélisation moléculaire, les mutants construits semblent insuffisants pour améliorer le potentiel catalytique de la CALB dans la réaction d'acylation de la quercétine.

(4) Comment la méthylation et l'acétylation de la quercétine influencent-elles son orientation dans la cavité de la CALB native et/ou acétylée et, son affinité pour le site actif ? Comment l'acétate lié de façon covalente à la sérine catalytique intervient-il sur le positionnement de la quercétine et de ses dérivés ainsi que sur l'affinité de liaison ?

Les dérivés méthylés de la quercétine introduits dans la CALB acétylée imitent la réaction d'acétylation. Le positionnement des substrats dans le site actif dépend de la localisation des groupements méthyles. Les substrats avec un groupe hydroxyle libre en position 4' s'insèrent par le cycle B du flavonoïde tandis que l'orientation des substrats avec un méthyl en position 4' est inversée (CA). De plus, la méthylation de la quercétine diminue l'affinité du flavonoïde pour la cavité.

Les substrats acétylés de la quercétine disposés dans la CALB miment la réaction d'acétylation. Dans ce cas, les acétates fixés au flavonoïde aglycone adoptent une position appropriée pour subir une hydrolyse. L'acétylation de la quercétine augmente son affinité de liaison pour la cavité de la CALB.

L'acétylation de la sérine catalytique de la CALB augmente l'affinité de liaison de la quercétine et de ses dérivés méthylés pour le site actif. De plus, le positionnement des substrats dans la cavité est influencé par la présence de l'acétate.

Le docking de dérivés méthylés et acétylés de la quercétine a été réalisé dans la CALB native et/ou acétylée. Ces dérivés s'insèrent par le cycle A ou le cycle B de la quercétine ce qui correspond respectivement aux deux grandes familles de conformères CA et CB. Faisant suite à la procédure de docking, un calcul d'énergie libre de liaison a été entrepris.

Dans la CALB acétylée, la méthylation de la quercétine intervient sur son positionnement dans le site actif: en présence d'un 4'-OH libre, le cycle B pointe vers le fond de la cavité de l'enzyme tandis qu'avec un 4'-O-méthyl, le substrat s'insère via son cycle A. De plus, une interaction de liaison hydrogène avec l'Asp134 est omniprésente quelque soient les substrats. Par ailleurs, l'affinité de liaison des substrats méthylés pour la CALB acétylée est réduite. La diminution du nombre d'hydroxyles disponibles sur la quercétine affaiblit la contribution électrostatique et renforce la contribution apolaire. Les groupements hydroxyles de la quercétine sont attirés par les résidus chargés négativement dans le site actif et de ce fait, les interactions électrostatiques entre la CALB et la quercétine sont importantes.

La chaîne acétate liée de façon covalente à la sérine catalytique affecte les positions et les orientations des conformères mais pas les interactions enzyme-substrat. L'orientation CB est prédominante pour les dérivés méthylés. L'affinité de liaison des substrats pour le site actif est supérieure en présence de la chaîne latérale. Dans les procédures de docking, le modèle de l'acyle enzyme fournit des résultats qui sont plus cohérents avec les résultats expérimentaux.

Dans la CALB native, l'acétylation de la quercétine donne lieu aux deux orientations CA et CB. Le carbonyle des acétates fixé à la quercétine se positionne à proximité des résidus du trou oxyanionique et de la sérine catalytique. Les interactions entre la quercétine et les résidus du site actif sont peu nombreuses. L'affinité de liaison des substrats acétylés pour la CALB est accentuée. Dans ce cas, les interactions électrostatiques intermoléculaires sont affaiblies et les contributions apolaires sont renforcées. Ce modèle d'alcoololyse indique que les acétates en position 7, 3' et 4' sont susceptibles d'être hydrolysés. Cela est en adéquation partielle avec les résultats expérimentaux dans lesquels la CALB hydrolyse les groupements hydroxyles du cycle B.

Dans l'ensemble, les résultats obtenus au cours de ce travail constituent une contribution pour la réaction d'acétylation de la quercétine et pour la compréhension de l'inactivité de la CALB envers le flavonoïde aglycone. Une réaction de nature chimique entre la quercétine et un large excès d'acétate de vinyle est identifiée pour la première fois. De plus, des événements moléculaires qui dépendent à la fois des structures de la lipase et du flavonoïde, expliquent l'absence d'activité catalytique de la CALB. En effet, les structures sont déterminantes pour les critères suivants: le positionnement et l'orientation du flavonoïde dans la cavité catalytique de l'acétyle-enzyme, la formation d'interactions intermoléculaires stabilisant le complexe enzyme-substrat et enfin, les groupements hydroxyles du flavonoïde atteignant les résidus catalytiques de l'enzyme.

Dans la continuité de ce travail, plusieurs perspectives peuvent être envisagées :

Approfondir l'étude expérimentale de l'acétylation chimique et chimio-enzymatique de la quercétine.

Dans la première partie de ce travail, l'étude expérimentale a permis de mettre en évidence l'acétylation chimique de la quercétine en présence d'un excès d'acétate de vinyle (ratio molaire donneur d'acyle (DA)/flavonoïde (F) = 330). Cette expérience peut être complétée en réalisant les différents points suivants :

- déterminer le ratio molaire minimal pour obtenir l'acétylation chimique. Généralement, dans la littérature des ratios molaires DA/F maximaux de 40 ont été testés et aucun produit n'a été détecté dans le milieu réactionnel [1-3]
- déterminer la régiosélectivité de l'acétylation chimique de la quercétine par résonance magnétique nucléaire (RMN). Cela contribuerait à fournir des informations sur la réactivité des groupements hydroxyles de la quercétine
- établir une cinétique complète de la réaction pour évaluer la vitesse de la réaction chimique
- vérifier le potentiel d'alcoololyse de la CALB en contrôlant la teneur en eau du milieu et confirmer la réaction chimio-enzymatique
- contrôler la régiosélectivité de la réaction chimio-enzymatique simultanée et séquentielle en présence de la CALB.

Simulation de dynamique moléculaire en présence de molécules de solvant

La réaction d'acétylation de flavonoïdes est menée expérimentalement dans des solvants organiques tels que l'acétone, l'acétonitrile ou le *tert*-amyl alcool. Cependant, les simulations de dynamique moléculaire ont été réalisées dans des boîtes d'eau. Ce choix a été motivé pour différentes raisons : (i) l'eau est le solvant le plus commun dans les simulations, (ii) les enzymes sont entourées d'une couche minimale de molécule d'eau, (iii) des simulations conduites dans différents solvants montrent que la CALB ne subit pas de changement conformationnel [4], ceci a été confirmé par une étude de dichroïsme circulaire indiquant que le contenu en structure secondaire de la CALB n'est pas modifié dans les solvants organiques [5]. Les modèles générés pour nos simulations ont permis d'expliquer la sélectivité envers la quercétine des biocatalyseurs CALB et PCL mais pas la régiosélectivité de la PCL. Il faudrait vérifier si l'utilisation de molécules de solvant organique dans les simulations peut apporter une explication sur la régiosélectivité de l'enzyme. En effet, aucune étude par modélisation moléculaire ne mentionne l'influence du solvant sur les interactions enzyme-substrat. Pourtant, une étude de modélisation moléculaire a démontré que des molécules de solvant peuvent s'insérer dans le mécanisme catalytique de la lipase [6]. Une autre étude sur la quercétine a montré que les molécules de solvant se répartissent différemment autour du flavonoïde aglycone selon leur hydrophobicité [7].

Élargir les stratégies de modélisation moléculaire pour les mutations

Malgré la modification rationnelle de la CALB, les choix des points de mutation étudiés par modélisation moléculaire se sont avérés insuffisants pour accroître le potentiel catalytique de la CALB. De plus, la méthodologie employée, c'est-à-dire l'étude structurale de la stabilité des mutants ainsi que l'étude du comportement dynamique des complexes enzymes-substrat, a limité le nombre de mutants testés. Une méthode alternative qui peut être utilisée pour examiner un plus grand nombre de mutations est la construction automatique d'une librairie combinatoire. Dans cette technique, des acides aminés sélectionnés dans la structure protéique sont remplacés par d'autres résidus afin d'obtenir toutes les combinaisons possibles de variants. Ensuite, l'application d'un protocole de docking et de scoring permet d'identifier les mutations d'intérêt.

Vérifications expérimentales des mutations

Lors de la conception des modèles et de la vérification de la stabilité structurale, on a supposé que les mutations sélectionnées n'engendraient pas de modifications dans le repliement de la protéine. De plus, la littérature mentionne plusieurs cas où une divergence est constatée entre les résultats de modélisation moléculaire et l'expérimentale [8, 9]. De ce fait, une vérification expérimentale des mutants permettrait de valider les modèles.

Modifier la méthodologie de docking

Dans le chapitre 5, une dynamique moléculaire de 20 ns a été réalisée sur l'enzyme mutée. Ensuite, des complexes enzyme-substrat ont été générés par docking en exploitant une *frame* de la trajectoire. Cet échantillonnage de *frames* pour le docking pourrait être élargi en réalisant par exemple un *clustering* des conformations au cours du temps (RMSD 2D) et en sélectionnant une *frame* par cluster identifié. Ces structures pourraient être ensuite employées pour appliquer un protocole de docking permettant d'entrer plusieurs structures initiales, et fournissant au final la conformation de la protéine ainsi que la pose du substrat qui donne le meilleur score. Ce type de procédure est réalisable avec le logiciel GOLD. Une autre alternative serait d'utiliser une méthode de docking considérant le récepteur comme flexible (par exemple, logiciel FlexX).

Continuer d'explorer l'ingénierie de la CALB

Les résultats préliminaires obtenus lors de la mutation de l'Asp134 par une alanine sont un point de départ intéressant. En effet, la modification du point d'ancrage électrostatique vers l'Asp145, localisée en face des résidus catalytiques, semble adéquate pour stabiliser les hydroxyles de la quercétine près des résidus catalytiques. Des dynamiques moléculaires sur des complexes lipases mutées-quercétine permettraient d'obtenir un complément d'informations sur le potentiel catalytique de cette mutation.

Une autre stratégie pour identifier d'autres cibles potentielles de mutations serait de fixer la quercétine de façon covalente aux résidus catalytiques pour mimer le second intermédiaire tétraédrique de la réaction d'acylation et ainsi, observer les déformations occasionnées dans le site actif.

Simulations de docking en modifiant le pH

Les études de modélisation moléculaire ont été réalisées à pH 7. A ce pH, les résidus aspartate et glutamate sont déprotonés. Des simulations de docking en modifiant l'état de protonation de ces résidus permettraient de voir l'influence d'une modification du pH sur les interactions CALB-quercétine. Une diminution du pH réduirait le potentiel électrostatique de la CALB ainsi que les interactions électrostatiques entre le flavonoïde aglycone et la CALB. Cependant, la déprotonation de l'aspartate catalytique devrait être maintenue car ce résidu stabilise l'histidine catalytique via une liaison hydrogène.

Modifier la régiosélectivité de la lipase de *Pseudomonas cepacia*

La lipase de *Pseudomonas cepacia* (PCL) catalyse la réaction d'acétylation de la quercétine. Expérimentalement, il a été montré que les groupements hydroxyles en position 7, 3' et 4' de la quercétine sont acylés. L'ingénierie de la régiosélectivité de cette lipase constitue une solution alternative au biocatalyseur CALB. Dans ce cas, deux pistes peuvent être explorées :

- le remplacement d'un des acides aminés suivants Thr18, Tyr23 et Tyr29 par des résidus polaires tels que l'acide aspartique renforcerait la contribution électrostatique. Ce point d'ancrage électrostatique pourrait attirer les groupes hydroxyles du cycle B de la quercétine et de ce fait, pourrait moduler la régiosélectivité de cette enzyme vis-à-vis du cycle A. En effet, l'étude de modélisation moléculaire du chapitre 4 a mis en évidence que ces trois résidus contribuent au maintien ou au déplacement de la quercétine dans le site actif de cette enzyme.
- le contrôle de la forme de la poche car la PCL possède un site actif plus large et une accessibilité plus aisée aux résidus catalytiques, en conséquence sa régiosélectivité pourrait être modulée.

Élargir la modélisation à des calculs quantiques

La réactivité de la quercétine et du donneur d'acyle n'a pas été prise en compte dans les modèles de cette étude. De ce fait, la modélisation utilisée dans ce travail pourrait être complétée par des calculs quantiques type DFT pour pouvoir prédire la réactivité des complexes formés et ainsi, compléter le protocole d'ingénierie *in silico* d'enzymes.

1. Chebil, L., et al., *Enzymatic acylation of flavonoids : effect of the nature of the substrate, origin of lipase and operating conditions on conversion yield and regioselectivity*. Journal of Agricultural and Food Chemistry, 2007. **53**(23): p. 9496-9502.
2. Xie, X.-N., et al., *Acylation of quercetin with a novel thermophilic esterase as biocatalyst*. Chemical Research in Chinese Universities, 2012. **28**(2): p. 225-229.
3. Stevenson, D.E., et al., *Direct acylation of flavonoid glycosides with phenolic acids catalyzed by Candida antarctica lipase B (Novozyme 435)*. Enzyme and Microbial Technology, 2006. **39**: p. 1236-1241.
4. Trodler, P. and J. Pleiss, *Modeling structure and flexibility of Candida antarctica lipase B in organic solvents*. BMC Structural Biology, 2008. **8**: p. art. 9.
5. McCabe, R.W., A. Rodger, and A. Taylor, *A study of the secondary structure of Candida antarctica lipase B using synchrotron radiation circular dichroism measurements*. Enzyme and Microbial Technology, 2005. **36**(1): p. 70-74.
6. Li, C., et al., *Analysis of the conformational stability and activity of Candida antarctica lipase B in organic solvents insight from molecular dynamics and quantum mechanics/simulations*. The Journal of Biological Chemistry, 2010. **285**(37): p. 28434-28441.
7. Chebil, L., et al., *Solubilities inferred from the combination of experiment and simulation. Case study of quercetin in a variety of Solvents*. Journal of Physical Chemistry B, 2010. **114**: p. 12308-12313.
8. Juhl, P.B., et al., *Engineering of Candida antarctica lipase B for hydrolysis of bulky carboxylic acid esters*. Journal of Biotechnology, 2010. **150**(4): p. 474-480.
9. Marton, Z., et al., *Mutations in the stereospecificity pocket and at the entrance of the active site of Candida antarctica lipase B enhancing enzyme enantioselectivity*. Journal of Molecular Catalysis B: Enzymatic, 2010. **65**(1-4): p. 11-17.

Annexe

Annexe A

Le champ de forces CHARMM

Le champ de forces CHARMM emploie l'expression suivante pour l'énergie potentielle[1]:

$$\begin{aligned}
 U(R) = & \sum_{\text{liaisons}} K_b (b - b_0)^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 \\
 & + \sum_{\text{dièdres}} K_\varphi (1 + \cos(n\varphi - \delta)) \\
 & + \sum_{\text{impropres}} K_\omega (\omega - \omega_0)^2 + \sum_{\text{Urey-Bradley}} K_{UB} (S - S_0)^2 \\
 & + \sum_{\text{paires non-liées}} \left\{ \varepsilon_{ij}^{\min} \left[\left(\frac{R_{ij}^{\min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{\min}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 \epsilon r_{ij}} \right\}
 \end{aligned} \tag{A.1}$$

K_b , K_θ , K_φ , K_ω , K_{UB} désignent respectivement les constantes de forces de liaison, d'angle, d'angle dièdre, d'angle de torsion impropre et d'Urey-Bradley. Les constantes b , θ , φ , ω , S représentent respectivement la longueur de liaison, l'angle de liaison, l'angle dièdre, l'angle de torsion impropre et la distance 1, 3 Urey-Bradley. Les indices zéro indiquent des valeurs d'équilibre correspondants aux termes individuels. Les constantes n et δ désignent respectivement la multiplicité ou la périodicité, et le déphasage pour le terme d'énergie de l'angle dièdres. Pour les termes d'énergie non liée, r_{ij} est la distance interatomique. Les paramètres ε_{ij}^{\min} et R_{ij}^{\min} désignent respectivement la profondeur du puits du potentiel de Lennard-Jones (LJ) et la distance à laquelle le terme LJ est minimum. Ces paramètres sont calculés avec les règles de combinaison de Lorentz-Berthelot :

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \tag{A.2}$$

$$R_{\min,ij} = \frac{R_{\min,i} + R_{\min,j}}{2} \tag{A.3}$$

Les paramètres q_i et q_j désignent les charges partielles des atomes i et j . La constante diélectrique relative, ϵ , est fixé à un dans les systèmes solvatés explicitement ce qui correspond à la permittivité du vide, ϵ_0 .

1. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy minimization and dynamics calculations*. Journal of Computational Chemistry, 1983. 4: p. 187-217.

**AUTORISATION DE SOUTENANCE
DU DOCTORAT DE L'UNIVERSITE DE LORRAINE**

o0o

VU LES RAPPORTS ETABLIS PAR :

**Monsieur VINH Tran, Professeur, Université de Nantes,
Monsieur BOULARD Yves, Chercheur HDR, CEA de Saclay**

Le Président de l'Université de Lorraine, autorise :

Mademoiselle BIDOUIL Christelle

à soutenir devant un jury de l'UNIVERSITE DE LORRAINE, une thèse intitulée :

"Modélisation moléculaire de l'acétylation de la quercétine par des lipases : étude des interactions enzyme-substat"

en vue de l'obtention du titre de :

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Intitulé du doctorat : **"Procédés Biotechnologiques & Alimentaires"**

Fait à Vandoeuvre, le **26 Octobre 2012**
Le Président de l'Université de Lorraine,
Pierre MUTZENHARDT



RÉSUMÉ

La quercétine (QCT) est un composé polyphénolique d'origine végétale connu pour ses activités antioxydantes et ses effets bénéfiques sur la santé. Sa solubilité, sa stabilité, sa biodisponibilité et ses activités biologiques peuvent être améliorées par une acylation sélective de ses groupements hydroxylés. Ce travail vise à étudier la possibilité d'une acétylation enzymatique de la QCT par la lipase B de *Candida antarctica* (CALB), la lipase la plus exploitée industriellement pour des estérifications régio- et énantiosélectives. Dans une perspective d'ingénierie rationnelle de l'enzyme, une démarche de modélisation moléculaire est mise en oeuvre pour mieux comprendre les interactions qui régissent le positionnement et l'orientation du substrat dans le site actif de la lipase. Dans une première partie expérimentale, l'absence d'activité d'acétylation de la CALB envers la QCT, en présence d'un excès d'acétate de vinyle, a été confirmée. Dans une seconde partie, cette inactivité de la CALB a été expliquée à l'aide de simulations de docking et de dynamique moléculaire. Elle résulte d'une orientation inappropriée du donneur d'acyle liée à la sérine catalytique et d'une proximité insuffisante des hydroxyles de la QCT vis-à-vis des résidus catalytiques. L'éloignement de la QCT de la triade catalytique est due à la rigidité de la molécule, l'étroitesse du site actif ainsi qu'à des interactions hydrophobes et électrostatiques entre le substrat et les résidus de la cavité. En revanche, cette approche de simulation moléculaire prédit un bon positionnement des deux substrats dans le site actif de la lipase de *Pseudomonas cepacia* (PCL), laquelle est capable d'acétyler la QCT. Dans une troisième partie, l'influence de mutations de deux résidus impliqués dans les liaisons de stabilisation hydrophobe de la QCT dans la CALB a été investiguée par simulation. La substitution d'isoleucines par des valines et des alanines conduit à une augmentation du volume de la poche catalytique et une mobilité accrue de la QCT. Mais ces mutations sont insuffisantes pour permettre un positionnement adéquat de l'acétate et de la QCT par rapport à la triade catalytique. La dernière partie focalise sur les interactions électrostatiques entre la QCT et le site actif de CALB. Les orientations du substrat dans la cavité suite à une méthylation ou une acétylation des groupements hydroxyles de la QCT sont précisées.

Mots-clés : Quercétine, Lipase B de *Candida antarctica*, Acétylation, Docking, Dynamique moléculaire

ABSTRACT

Quercetin (QCT) is a plant-produced polyphenolic compound well-known for its antioxidant activities and beneficial health effects. Its solubility, stability, bioavailability and biological activities may be improved by a selective acylation of its hydroxyl groups. This work aims at studying the possibility of QCT enzymatic acetylation by *Candida antarctica* lipase B (CALB), the most industrially exploited lipase for regio- and enantioselective esterifications. In prospect of the rational enzyme design, a molecular modeling approach was implemented to understand the interactions that govern the substrate positioning and orientation in the lipase's active site. In a first experimental part, the absence of CALB acetylation activity towards quercetin in excess of vinyl acetate was confirmed. In a second part, this inactivity of CALB was explained by means of docking and molecular dynamics simulations. This results from an inappropriate positioning of the acyl donor linked to the catalytic serine and from an insufficient proximity of QCT hydroxyls vis-à-vis catalytic residues. The distance of QCT from the catalytic triad is due to its rigidity and to the narrow active site as well as to hydrophobic and electrostatic interactions between the substrate and the cavity residues. On the contrary, this molecular simulation approach predicts an appropriate positioning of both substrates in the active site of *Pseudomonas cepacia* lipase (PCL), which can perform QCT acetylation. In a third part, the impact of mutations of two residues implicated in the stabilization of QCT by hydrophobic interactions in CALB was investigated through simulations. The substitution of isoleucines by alanines and valines led to an increase in the catalytic pocket volume which intensified the mobility of QCT. However, these mutations are insufficient to allow an appropriate positioning of acetate and QCT in relation to the catalytic triad. The last part of this work focuses on the electrostatic interactions between QCT and CALB's active site. The substrate orientation in the cavity following methylation or acetylation of QCT's hydroxyl groups was clarified.

Keywords: Quercetin, *Candida antarctica* Lipase B (CALB), Acetylation, Docking, Molecular Dynamics