



Université de Lille - Sciences et Technologies

École doctorale Sciences de la Matière, du Rayonnement et de l'Environnement

Thèse de Doctorat

Spécialité : Ingénierie des Fonctions Biologiques

Présentée pour l'obtention du grade de docteur de l'université de Lille

par

Maxime De Sousa Lopes Moreira

Valorisation du glycérol par catalyse enzymatique

Préparée à **Institut Charles Viollette – ICV – EA 7394**

Soutenue le 28 Novembre 2018, devant le jury composé de :

Rapporteur	Pr. Laurence HECQUET	Professeur des universités	Université Clermont Auvergne
Rapporteur	Dr. Bastien DOUMÈCHE	Maître de conférences-HDR	Université Lyon 1
Examinatrice	Dr. Christine HÉLAINE	Maître de conférences	Université Clermont Auvergne
Président	Pr. Pascal DHULSTER	Professeur des universités	Université de Lille
Invité	Dr. Jérôme LE NÔTRE	Responsable R&D	SAS PIVERT
Directeur de thèse	Pr. Renato FROIDEVAUX	Professeur des universités	Université de Lille

Remerciements

Mes premiers remerciements s'adressent aux Pr. Laurence Hecquet, Dr. Bastien Doumèche, Dr Christine Hélaine, Pr Pascal Dhulster et Dr Jérôme Le Nôtre, membres de mon jury de thèse qui ont eu l'amabilité de prendre sur leur temps pour juger mon travail.

Merci à la SAS PIVERT, le financeur de ma thèse et plus particulièrement à Madame Cécile Armielle responsable des ressources humaines qui m'a recruté et suivi tout au long de ma thèse et à Monsieur Jérôme Le Nôtre qui a participé activement au projet Glycybride et ma thèse. Merci aussi à tous les membres du projet Glycybride d'avoir été présents pendant ces trois années et pour ne citer qu'eux Mickael Capron et Pascal Fongarland.

Je voudrais aussi remercier l'Institut Charles Viollette dans sa globalité pour leur accueil, leur aide et leur bonne humeur pendant ces trois années. Plus particulièrement, un grand merci au directeur de l'Institut Monsieur Pascal Dhulster qui a toujours été à l'écoute et qui m'a recommandé pour mon futur post-doctorat.

Un immense merci à Corine, Gabrielle, Cathy et Julie pour tout ce qu'elles font au sein du laboratoire. Un merci tout particulier à Matthieu qui a eu la lourde tâche de me former à la biologie moléculaire, avec qui nous avons galéré mais aussi bien rigolé. Merci aussi à Christophe, Benoit et Rozy pour avoir toujours été présent pour des questions et des discussions, parfois même autour d'un verre. Merci à Frédéric AFFOUARD et Natalia CORREIRA pour leur aide sur la partie docking. Un immense merci à la plateforme Bilille et surtout à Hélène Touzet et Isabelle Guigon avec qui nous avons beaucoup travaillé sur la création du programme bio-informatique.

Un grand merci aux personnes qui m'ont donné goût à la recherche : Laurence Hecquet, Marielle Lemaire, Christine Hélène, Franck Charmantray, Virgile Hélène, Thierry Gefflaut, Federico Cisnetti, Arnaud Gautier, Wolf Dieter Fessner. J'espère n'avoir oublié personne.

Merci aussi à Marion, Ghina et Mathilte qui m'ont beaucoup appris durant mon stage de master 2. Petite pensée pour mes amis Kevin et Ombeline qui soutiennent eux aussi leur thèse et avec qui j'ai passé toutes mes années universitaires.

Un merci particulier à ceux qui sont devenus plus que des collègues au fil de ces trois ans et avec qui j'ai passé beaucoup de temps les week-end : Juju, Alexandra, Mon Chou, Rémi, Kalim, Antoine, BOB, Xavier, Marjo et Michel. Mais bien entendu merci aussi à tous les autres que j'aurais oublié et avec qui nous avons peut-être noué moins de liens mais passé d'excellents moments tout de même.

Merci à mon directeur de thèse Renato qui est devenu bien plus que cela et qui va encore devoir me supporter 2 ans en post-doctorat. Renato m'a toujours soutenu dans mes idées et aidé en cas de besoin. Je pense surtout à la correction de ce manuscrit et de toutes les communications effectuées... Les gens qui me connaissent savent que l'écriture n'est pas forcément mon fort. Je retiendrai particulièrement les moments passés hors travail, badminton, congrès et bien d'autres. J'espère que nous garderons contact pendant très longtemps.

Enfin mes dernières pensées vont à ma famille. À mes parents et les parents de ma compagne qui m'ont toujours soutenu financièrement et moralement tout le long de mes études. Un immense merci pour m'avoir poussé à faire ces 8 années d'études et un bon courage à mon frère qui a décidé de suivre le même chemin.

Pour finir, je voudrais remercier Angéline Aujon qui me supporte depuis plus de 6 ans maintenant. C'est grâce à elle si j'ai réalisé ma thèse à Lille. Elle a abandonné ses envies de thèse pour me permettre de réaliser la mienne et m'a suivi sans hésitation dans ce projet. Il n'y a pas de mot assez fort pour la remercier.

Ce travail a été réalisé, en partenariat avec la SAS PIVERT, dans le cadre de l'Institut pour la Transition Énergétique (ITE) P.I.V.E.R.T. (www.institut-pivert.com) retenu parmi les Investissements d'Avenir. Ce travail a bénéficié d'une aide de l'État au titre du Programme d'Investissements d'Avenir portant la référence ANR-001-01.

Sommaire

SOMMAIRE	1
TABLE DES FIGURES	4
TABLE DES TABLEAUX	9
INTRODUCTION	10
OBJECTIF DE LA THESE	16
CHAPITRE 1 : PARTIE BIBLIOGRAPHIQUE.....	19
<i>I. La création de liaisons C-C par les aldolases.....</i>	<i>20</i>
1. Introduction	20
2. Les aldolases	20
2.1. Définition et classification	20
2.2. Les pyruvate et phospho-énolpyruvate aldolases	23
2.3. Les Acétaldéhyde aldolases	24
2.4. Les Glycine aldolases	25
2.5. DHAP aldolases.....	26
3. La fructose-6-phosphate aldolase	28
3.1. Structure et mécanisme de la FSA _{wt}	28
3.2. Historique des travaux sur la FSA.....	31
3.2.1 La FSA _{wt}	31
3.2.2 Les mutants de la FSA	39
<i>II. Valorisation du glycérol.....</i>	<i>45</i>
1. Introduction	45
2. Valorisation par catalyse chimique	46
2.1. Généralités	46
2.2. Reformage du glycérol	47
2.3. Halogénéation du glycérol	48
2.4. Synthèse de carbonates de glycérol	48
2.5. Synthèse de polymères de glycérol	49
2.6. Déshydratation du glycérol	51
2.7. Éthérification du glycérol	52
2.8. Réduction du glycérol.....	52
2.9. Oxydation du glycérol par voie chimique	53
2.10. Utilisation des micro-ondes	56
3. Valorisation du glycérol par voies biotechnologiques	57
3.1 Valorisation catalysée par des microorganismes complets.....	57
3.2. Valorisation par enzymes isolées.....	60

3.2.1. Utilisation de lipases	60
3.2.2. Oxydation par voies enzymatiques.....	60
3.2.2.1 Les alcools oxydases	60
3.2.2.2. Transformation des alcools primaires du glycérol et du glycolaldéhyde en aldéhydes par les oxydases	62
CHAPITRE 2 : RESULTATS ET DISCUSSION.....	65
<i>I. Étude du Clivage du glycéraldéhyde</i>	<i>66</i>
1. Objectif	66
2. Production et criblage d'une banque de mutants de FSA pour cliver le glycéraldéhyde	67
3. Caractérisation du mutant FSA L107Y A129G	70
4. Production du mutant FSA L107Y A129G	72
5. Recherche d'une 3 ^{ème} mutation pour lever l'inhibition de la FSA L107Y A129G par le D-glycéraldéhyde	76
5.1. Stratégie	76
5.2. Création du fichier contenant les coordonnées spatiales du mutant L107Y A129G grâce au logiciel SCWRL4	77
5.3. Études de Docking	79
5.3.1. Avec le L-glycéraldéhyde.....	81
5.3.2. Avec le D-glycéraldéhyde	84
5.3.3. Comparaison des études de docking du L-glycéraldéhyde (cluster 1) et D-glycéraldéhyde (cluster 2)	87
5.4. Création des mutants définis par docking	90
<i>II. Recherche d'une glycérol oxydase dans la biodiversité</i>	<i>96</i>
1. Introduction	96
2. Création du programme bio-informatique.....	96
2.1. Stratégie	96
2.2. Création du set de références de séquences protéiques primaires	100
2.3. Création informatique de la banque d'enzymes	102
3. Création de la banque d'oxydases.....	106
3.1. Préparation des microorganismes et extraction de l'ADN	106
3.2. Amplification des gènes d'intérêt par PCR	109
3.3. Clonage des gènes amplifiés	114
3.4. Transformation et expression des protéines.....	117
3.4.1. Obtention de colonies transformées non productrices	117
3.4.2. Obtention de colonies transformées productrices de protéines recombinantes.....	118
3.4.3. Surexpression et production des protéines recombinantes	118
CONCLUSION GENERALE ET PERSPECTIVES	122
MATERIELS ET METHODES	128
1. Généralités.....	129
1.1. Réactifs	129
1.2. Appareils pour analyses	129
2. Protocoles standards pour les cultures bactériennes	129

2.1. Protocole pour la culture des souches bactériennes.....	129
2.2. Purification de la FSA par résine IMAC.....	130
2.3. Production et purification de la FSA par choc thermique.....	131
2.4. Dosage des protéines par le test de Bradford	131
2.5. Protocole standard SDS-PAGE pour la vérification de la production des protéines recombinantes	132
2.6. Composition des différents milieux de culture.....	133
3. Procédures standards de biologie moléculaire	139
3.1. Extraction de plasmides grâce au GeneJET Plasmid Miniprep Kit de chez ThermoScientific.....	139
3.2. Protocole pour gel d'électrophorèse ADN.....	139
3.3. Procédure générale pour la mutagénèse dirigée avec le kit Agilent Quick Change II Mutagenesis	140
3.4. Extraction d'ADN génomique à l'aide du kit DNA Purification de Promega.....	142
3.5. Procédure standard pour l'Amplification en Chaîne par Polymérase (PCR).....	143
3.6. Procédure standard pour le LIC clonage	143
3.7. Dosage de l'ADN (plasmides inclus) par NanoDrop	144
4. Protocoles d'enzymologie	145
4.1. Détermination des constantes cinétiques par la méthode de Hanes-Woolf	145
4.2. Test de criblage des fructose-6-phosphate aldolases.....	147
4.3. Synthèse et caractérisation RMN du L-xylose	147
4.4. Criblage des oxydases	148
ANNEXES.....	149
ANNEXE 1	150
ANNEXE 2	151
REFERENCES BIBLIOGRAPHIQUE.....	152

Table des figures

Figure 1 : la bio-économie en France	12
Figure 2 : utilisation du glycérol	13
Figure 3 : valorisation du glycérol par oxydation	14
Figure 4 : principe de la chimie verte par Anastas et Warner.....	15
Figure 5 : voie envisagée pour la valorisation du glycérol par catalyse enzymatique.....	17
Figure 6 : réaction classique d'aldolisation	21
Figure 7 : mécanismes réactionnels des aldolases de classe I et II.....	22
Figure 8 : réaction native de la DERA	25
Figure 9 : cross-aldéhyde addition de l'acétaldéhyde par la DERA.....	25
Figure 10 : mise en évidence des stéréoisomères obtenus par l'aldolisation de la glycine sur un aldéhyde accepteur catalysée par les différentes thréonine aldolases naturelles	26
Figure 11 : complémentarité des DHAP aldolases	27
Figure 12 : synthèse d'alcools tertiaires catalysée par la Rhamnulose-1-phosphate aldolase de <i>Bacteroides thetaiotamicron</i>	28
Figure 13 : réaction "native" de la FSA.....	28
Figure 14 : structure 3D de la FSA avec en a) une vue du dessus, en b) une vue de profil pour mettre en évidence les sites actifs et en c) la représentation d'un seul monomère	29
Figure 15 : mécanisme de rétro aldolisation catalysée par la FSA.....	30
Figure 16 : site actif de la FSA _{wt}	31
Figure 17 : réactions catalysées par la FSA _{wt} avec le DHA en tant que donneur et différents aldéhydes	32
Figure 18 : réactions catalysées par la FSA _{wt} avec HA en donneur et différents aldéhydes accepteurs	33

Figure 19 : synthèse one-pot d'iminocyclitols catalysée par la FSA _{wt} et palladium sur charbon	35
Figure 20 : synthèse one-pot d'iminocyclitols catalysée par la FSA _{WT} et palladium sur charbon	36
Figure 21 : catalyse de la dimérisation du GA par la FSA _{wt} suivie de l'aldolisation de HA pour former le 1-déoxy-D-ido-hept-2-ulose	37
Figure 22 : substrats accepteurs cycliques amino-aldéhydes protégés par groupement Cbz av	37
Figure 23 : synthèse de déoxynitrocyclitols catalysée par la FSA _{wt}	38
Figure 24 : synthèse de nitrocyclitols catalysée par la FSA A129S	38
Figure 25 : synthèse de précurseurs de thio-sucres catalysée par la FSA _{wt}	39
Figure 26 : possible mécanisme d'aldolisation catalysée par la FSA avec GA en donneur et un aldéhyde en accepteur	42
Figure 27 : double addition du glycolaldéhyde catalysée par deux variants de la FSA	44
Figure 28 : évolution de la production (histogramme bleu) et du nombre de publications (courbe orange) concernant la valorisation du glycérol	45
Figure 29 : synthèse par transestérification des huiles végétales des biocarburants	46
Figure 30 : les différentes voies de transformation du glycérol par catalyse chimique	46
Figure 31 : reformage photocatalytique du glycérol sous action d'un complexe Pt 0,2 %/TiO ₂ synthétisé en présence d'urée	47
Figure 32 : synthèse d'épichlorohydrines à partir du glycérol	48
Figure 33 : synthèse de carbonates de glycérol	49
Figure 34 : synthèse d'aryloxypropanediols avec formation in situ de carbonates de glycérol	49
Figure 35 : polymérisation de l'éther éthoxyéthyl glycidyl catalysée par l'hydroxyde de césium	50
Figure 36 : synthèse de dendrimère de poly(glycérol-co-acide lactique)	50

Figure 37 : déshydratation du glycérol en acroléine.....	51
Figure 38 : étherification du glycérol et du 1-phénylpropanol	52
Figure 39 : voies de formation du 1,2- et 1,3-propanediol	53
Figure 40 : produits obtenus par oxydation du glycérol	54
Figure 41 : oxydation sélective du glycérol en DHA.....	55
Figure 42 : Les quatre produits obtenus par action des micro-ondes sur le catalyseur à base de Sb et V sur le glycérol	57
Figure 43 : transestérification du glycérol et de l'acétate de méthyle pour la synthèse d'acétate de glycérol	60
Figure 44 : production du L-glycéraldéhyde à partir de glycérol	63
Figure 45 : synthèse du glyoxal par catalyse enzymatique	66
Figure 46 : méthode de criblage de la FSA.....	68
Figure 47 : thermostabilité de la FSA L107Y A129G	70
Figure 48 : Interprétation du spectre RMN du proton avec en noir le α -L-xylose et en rouge le β -L-xylose.....	71
Figure 49 : réaction secondaire entre le glycolaldéhyde produit et le L-glycéraldéhyde catalysée par la FSA L107Y A129G	72
Figure 50 : production de protéines recombinantes chez <i>E. coli</i>	73
Figure 51 : procédure pour réaliser la mutation à l'aide du KAQ	74
Figure 52 : séquences génétiques de la FSA _{wt}	75
Figure 53 : séquences peptidiques de la FSA _{wt} et de la FSA L107Y A129G	75
Figure 54 : différentes étapes de l'utilisation du logiciel SCWRL4 pour la création du fichier .PDB de la FSA L107Y A129G	78
Figure 55 : position de la boîte d'amarrage du glycéraldéhyde (ligand) dans le site actif de la FSA L107Y A129G.....	80
Figure 56 : position du L-glycéraldéhyde des clusters 1 et 3 après docking	81

Figure 57. a et b : interactions glycéraldéhyde/FSA	82
Figure 58 : acides aminés du site actif de la FSA mutant impliqués dans les interactions avec le L-glycéraldéhyde des clusters 1 et 3.....	83
Figure 59 : position du D-glycéraldéhyde des clusters 1, 2 et 3 dans le site actif après docking	84
Figure 60 : interactions D-glycéraldéhyde/FSA	85
Figure 61 : interactions entre le D-glycéraldéhyde et la poche de liaisons hydrogènes de la FSA.....	86
Figure 62 : acides aminés du site actif de la FSA impliqués dans les interactions avec le D-glycéraldéhyde des clusters 1,2 et 3.	87
Figure 63 : superposition des deux énantiomères sélectionnés dans le site actif de la FSA L107Y A129G	88
Figure 64 : comparaison des acides aminés du site actif de la FSA impliqués dans les interactions avec le D-glycéraldéhyde et le L-glycéraldéhyde.....	89
Figure 65 : séquence génétique de la FSA L107Y A129G en noir et des amorces contenant les mutations positions 26G et 131G	90
Figure 66 : séquence peptidique des FSA L107Y A129G, T26G L107Y A129G et L107Y A129G Y131G	91
Figure 67 : criblage des mutants produits sur la réaction de clivage du L-glycéraldéhyde	91
Figure 68 : criblage des mutants sur la réaction de clivage du L/D-glycéraldéhyde	93
Figure 69 : interaction hydrogène entre le L-glycéraldéhyde et le site actif du mutant de la FSA L107Y A129G.....	94
Figure 70 : oxydation du glycérol en L-glycéraldéhyde et du glycolaldéhyde en glyoxal	97
Figure 71 : stratégie de recherche de nouvelles oxydases dans la biodiversité	98
Figure 72 : arbre d'identité du set de références	101
Figure 73 : arbre phylogénique des 85 représentants de clusters	102

Figure 74 : procédure suivie depuis la réception des bactéries lyophilisées jusqu'à l'obtention des ADNs bactériens purs	106
Figure 75 : principe du clonage d'un ORF dans le vecteur d'expression pLATE11 selon la méthode LIC	116
Figure 76 : étapes du processus de transformation	118
Figure 77 : gel SDS-page pour la vérification de la production des protéines	120
Figure 78 : criblage des oxydases	121
Figure 79 : Voie envisagée pour la valorisation du glycérol par catalyse enzymatique	123
Figure 80 : criblage des oxydases	126
Figure 81 : purification d'enzymes sur résine IMAC, a : étapes de purification, b : interaction entre l'étiquette 6-His et la résine Ni-NTA.....	130
Figure 82 : représentation graphique des vitesses initiales en fonction des concentrations en substrat : $V=f(\text{concentration})$	146
Figure 83 : représentation de Hanes-Woolf : $S/V = f(S)$	146
Figure 84 : méthode de criblage et de détermination des constantes cinétiques de la FSA L107Y A129G	147
Figure 85 : criblage des oxydases	148
Figure 86 : carte du plasmide contenant les ARNt incluse dans les cellules compétentes Rosetta 2 (DE3).....	150
Figure 87 : carte du plasmide pLATE11	151

Table des tableaux

Tableau 1 : nouvelle classification des aldolases	23
Tableau 2: principales PEP aldolases.....	24
Tableau 3 : principales Pyr aldolases.....	24
Tableau 4 : comparaison des paramètres cinétiques de la FSA _{wt} et de la FSA A129S pour différents substrats	40
Tableau 5 : oxydation du glycérol catalysée par l'or sur différents supports	56
Tableau 6 : récapitulatif des microorganismes métabolisant le glycérol	59
Tableau 7 : classification par le Comité International de l'Union de la Biochimie et de la Biologie Moléculaire.....	61
Tableau 8 : résultat du criblage des microorganismes pour la recherche d'activité glycérol oxydase.....	62
Tableau 9 : résultat du criblage de la banque de FSA	69
Tableau 10 : détermination des constantes cinétiques de la FSA L107Y A129G	70
Tableau 11 : comparaison des valeurs d'énergies et de constantes d'inhibition pour les deux représentants sélectionnés	88
Tableau 12 : enzymes du set de références	100
Tableau 13 : matrice d'identité des enzymes du set de références	101
Tableau 14 : référencement des bactéries dont les gènes vont être clonés.....	105
Tableau 15 : quantité d'ADN extraite à partir des bactéries cultivées et corrélation avec la PCR 16S.....	109
Tableau 16 : récapitulatif des gènes amplifiés et caractéristiques d'amplification.....	113
Tableau 17 : gènes amplifiés avec succès	115

Introduction

Depuis les années 90, la réglementation française sur le traitement des déchets ne cesse d'évoluer au travers du Code de l'Environnement qui a défini, à l'époque, quatre priorités :

- Principe de réduction à la source de la production et de la nocivité des déchets
- Principe de proximité pour organiser et limiter le transport en distance et en volume
- Principe de valorisation des déchets par réemploi, recyclage ou toute autre action visant à obtenir à partir des déchets des matériaux réutilisables ou de l'énergie
- Principe d'information du public sur les effets pour l'environnement et la santé publique des opérations de production et d'élimination des déchets

En 2007, le Grenelle de l'Environnement ainsi que les réglementations européennes REACH (enregistrement, évaluation et autorisation des produits chimiques) sont venus compléter ce code. Ces règles s'appliquent aussi bien à la population qu'aux industriels. Pour les faire appliquer, l'état a mis en place une grande campagne publicitaire basée sur le recyclage des déchets. Du côté industriel, la loi a connu de nombreuses modifications notamment en 2015 avec la mise en place de la Loi de Transition Énergétique et le plan de réduction et de valorisation des déchets avec des slogans comme « bien trier c'est mieux respirer », « tout le monde a le droit à une seconde chance » (<https://www.legifrance.gouv.fr>). De nombreux spots publicitaires ont aussi été créés et plus récemment une application smartphone a vu le jour, un guide du tri qui permet aux utilisateurs une réalisation parfaite du tri sélectif. Des taxes sont instaurées pour les entreprises les plus polluantes (taxe carbone et taxe générale sur les activités polluantes) ce qui pousse les industriels à changer leur mode de fonctionnement et à tendre vers la bio-économie.

Ce terme de bio-économie est considéré par le gouvernement français seulement depuis 2017 alors que T.I. Baranoff aurait « créé le concept » en 1925. L'état français par l'intermédiaire du ministère de l'agriculture, de l'agroalimentaire et de la forêt a défini ce terme comme « l'ensemble des activités économiques fondées sur les bio-ressources, c'est à dire les matières organiques terrestres ou marines, végétales ou animales. Cette économie est basée sur du carbone renouvelable, de la production à la transformation, jusqu'à la valorisation des coproduits et des bio-déchets, tout en visant la sécurité alimentaire ». Elle

crée les conditions du passage d'une économie fondée sur les ressources fossiles à une économie fondée sur la biomasse » (Figure 1).



Figure 1 : la bio-économie en France © Ministère de l'agriculture

Cette bio-économie a bien entendu pour but l'utilisation « totale » des matières premières mais aussi un objectif de création d'emplois puisqu'en 2016 on comptait environ 1,6 millions d'emplois autour de la bio-économie en France.

Dans ce contexte, la valorisation des déchets devient un enjeu aussi bien environnemental qu'économique. Le « zéro déchet » est devenu l'un des objectifs majeurs aussi bien pour les multinationales que pour les gouvernements. Le bois est par exemple valorisé en énergie par combustion, les matières organiques en compost... La biomasse agricole, comprenant la biomasse oléagineuse et lignocellulosique, peut elle aussi être entièrement valorisée. Le colza, le tournesol et beaucoup d'autres plantes oléagineuses, regroupées sous le nom de biomasse oléagineuse, sont très présentes en Hauts de France puisqu'elles représentent plus de 160000 hectares (environ 10 % de la surface exploitée

dans ce domaine en France) et rentrent dans la composition de produits cosmétiques, alimentaires... Les utilisations les plus connues de cette biomasse concernent la production d'énergies renouvelables (biocarburants) et de bio-polymères (plastique « végétal »)^[1]. Depuis maintenant plusieurs années, une nouvelle tendance vise, à partir de la biomasse, à produire des « molécules plateformes », petites molécules élémentaires à partir desquelles il est possible de synthétiser des molécules chimiques complexes^[2].

Prenons l'un des exemples les plus populaires de cette transition énergétique, l'utilisation des biocarburants pour limiter la consommation des énergies fossiles. Lors de la production du biodiesel par transestérification des huiles végétales avec le méthanol, un coproduit non intéressant pour le biocarburant est produit à hauteur de 10 % (en masse), le glycérol. Actuellement le glycérol est purifié et utilisé principalement pour des applications cosmétiques et pharmaceutiques^[3] (Figure 2).

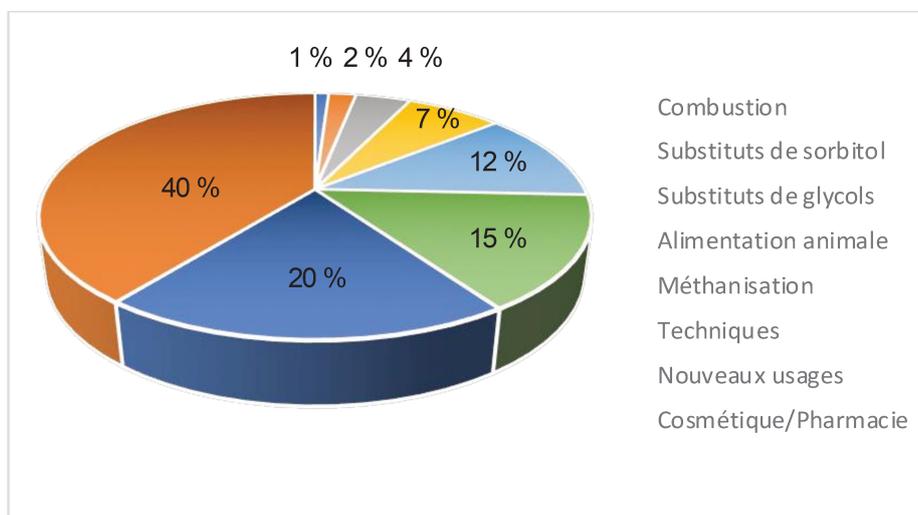


Figure 2 : utilisation du glycérol^[4]

Les biocarburants ne cessent de se développer, la production mondiale de biodiesel est passée de 15 000 à 430 000 barils par jour^[4] entre 2000 et 2015 ce qui a entraîné la multiplication par neuf de la production de glycérol. Néanmoins, il faut nuancer ce propos puisque depuis fin 2017, l'utilisation du diesel ainsi que des biodiesels stagne pour ne pas dire diminue. L'utilisation du glycérol reste tout de même à l'heure actuelle une énorme manne financière et un enjeu pour le développement durable. L'oxydation du glycérol est beaucoup étudiée ces 10 dernières années, un exemple récent concerne les travaux de Katryniok *et al.*^[5] sur la synthèse de dérivés carbonylés du glycérol (Figure 3).

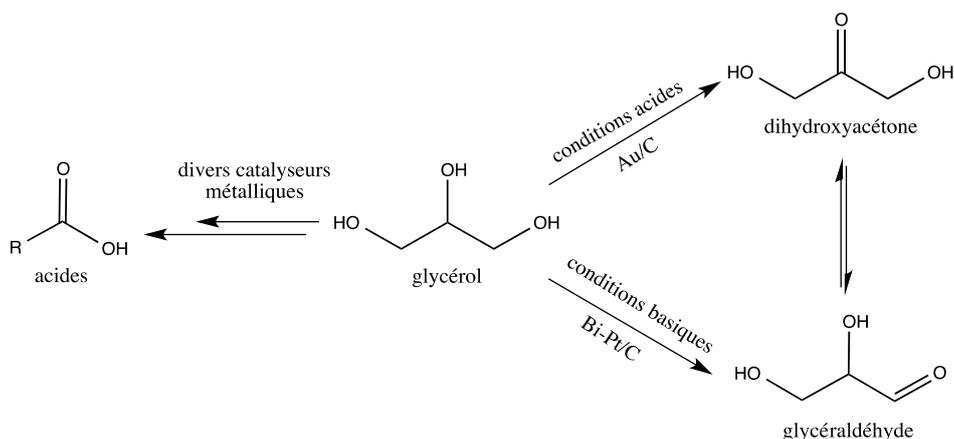


Figure 3 : valorisation du glycérol par oxydation. L'utilisation de catalyseur métallique mène à différents acides, en condition basique avec un catalyseur AU/C (argent supporté sur charbon) à la dihydroxyacétone et en condition acide avec un catalyseur Bi-Pt/C (bismuth et platine supportés sur charbon) au glycéraldéhyde^[5]

Les conditions de température pour l'oxydation du glycérol en glycéraldéhyde et dihydroxyacétone (DHA) sont situées entre 50 et 100°C sous pression atmosphérique. Le DHA est une cible très intéressante car il est utilisé dans l'industrie cosmétique. De plus, son prix de vente est de 150 \$ le kilo tandis que le glycérol ne coûte que 0,3 \$ le kilo. La production de glycéraldéhyde est encore plus rentable financièrement, la société Sigma Aldrich le commercialise à 102 000 € le kilo. Concernant la formation d'acides, les températures sont du même ordre de grandeur, cependant, la pression est généralement entre 3 et 5 bars.

Les produits ciblés sont donc très intéressants financièrement mais les conditions réactionnelles doivent être étudiées pour tendre vers une chimie plus éco-responsable :

- températures élevées entraînant un coût plus élevé de la production ainsi qu'une « pollution » (réchauffement climatique)
- utilisation de base, dangereuse pour le manipulateur et difficile à recycler
- sélectivité variable conduisant très souvent à des mélanges racémiques lors de la synthèse de dérivés carbonyles notamment

Pour pallier à ces problèmes, une solution est de créer des catalyseurs qui seraient capables de travailler en conditions douces, c'est à dire à pH neutre, à température ambiante et être plus sélectif vis-à-vis du produit souhaité. C'est dans ce contexte de valorisation de la biomasse oléagineuse issue de la synthèse de biodiesel que le projet GLYCYBRIDE financé par la SAS PIVERT, a vu le jour. L'un des axes de recherche consiste dans un premier temps à transformer le glycérol par catalyse chimique. Dans un second temps,

cette catalyse chimique est combinée à une catalyse enzymatique, concept de catalyse hybride, tout en respectant les principes de chimie verte décrits par Paul Anastas et John C. Warner en 1998 (Figure 4).



Figure 4 : principe de la chimie verte par Anastas et Warner^[6]

Cette combinaison des deux « types » de catalyse est très novatrice puisque l'utilisation d'un catalyseur organométallique supporté et d'un autre enzymatique en one-pot one-step (utilisation d'un unique réacteur sans étape de purification intermédiaire) n'a que très rarement été étudiée. Un des verrous à la mise en œuvre de ce concept concerne les conditions réactionnelles. En effet, il faut rechercher des conditions adéquates aux deux catalyses avec comme cahier des charges le respect des principes évoqués ci-dessus, ce qui est un véritable challenge.

Ce défi technologique intéresse la SAS Pivert, Institut de Transition Énergétique (ITE) à l'interface entre le monde industriel et la recherche académique, pour l'obtention de molécules plateformes à partir de la biomasse oléagineuse. Leur objectif est de concevoir et développer la bioraffinerie oléagineuse du futur. La SAS Pivert s'est donnée trois missions :

- développer la chimie du végétale dans une logique de bioraffinerie et de valorisation de la plante entière

- offrir de nouveaux débouchés au monde agricole et aux acteurs des technologies de l'équipement
- participer à la réindustrialisation du territoire national

Pour mener à bien ces missions, quatre axes stratégiques ont été définis :

- un programme de recherche nommé GENESYS, mené avec consortium académique international
- une plateforme technologique évolutive, le BIOGIS Center, pour accélérer le transfert de la recherche à l'industrie
- des projets de maturation menés par les industries partenaires
- des programmes de formation continue et diplômante

Le programme de recherche GENESYS est divisé en 7 work-packages (groupe de travail) allant de la production de la biomasse à la production de biomolécules. Depuis 2012, plus de 60 projets ont été financés. La thèse qui vous est présentée aujourd'hui est incluse dans le programme de recherche GENESYS et plus précisément dans le work-package 3 intitulé « catalyse et biocatalyse » au travers de l'appel à projet Glycybride faisant partie du programme de recherche GENESYS.

Objectif de la thèse

L'objectif du projet Glycybride est de transformer le glycérol pour obtenir le glyoxal, molécule plateforme. Cette molécule est principalement utilisée comme agent de réticulation dans l'industrie papetière, textile et cuir. Ses propriétés sont intéressantes puisque le glyoxal en plus d'être biodégradable est aussi un substitut du formaldéhyde et du glutaraldéhyde. Pour cela deux thèses et un post-doctorat en collaboration ont été financés :

- Une thèse en catalyse chimique du solide au sein de l'Unité de Catalyse et Chimie du Solide (UCCS) de l'Université de Lille pour étudier la réaction d'oxydation du glycérol en glycéraldéhyde ainsi que l'oxydation du glycolaldéhyde en glyoxal.
- Une thèse, au sein de l'Institut Charles Viollette (ICV-Université de Lille), celle qui va être décrite dans ce manuscrit et qui vise à utiliser la biocatalyse pour cliver le glycéraldéhyde en glycolaldéhyde et formaldéhyde.

- Un post-doctorat sur la mise au point d'un réacteur pour la combinaison de ces deux catalyses au Laboratoire de Génie des Procédés Catalytiques (LGPC) de l'Université de Lyon.

Pour atteindre l'objectif final qui est de transformer le glycérol en glyoxal en couplant les deux types de catalyse, les deux projets de thèses ont donc dû prendre en compte les contraintes de conditions réactionnelles de chaque catalyse pour permettre ensuite leur combinaison tout en maintenant une bonne efficacité et stabilité.

Le choix du glycérol comme substrat de départ, de la cible, le glyoxal ainsi que le schéma réactionnel a été décidé par le comité scientifique de la SAS PIVERT. Durant cette thèse nous avons dû nous plier à ce schéma sans pouvoir en sortir. La cascade réactionnelle comprend une première étape d'oxydation du glycérol en glycéraldéhyde, une deuxième étape de clivage du glycéraldéhyde en glycolaldéhyde et enfin une étape d'oxydation du glycolaldéhyde en glyoxal (Figure 5).

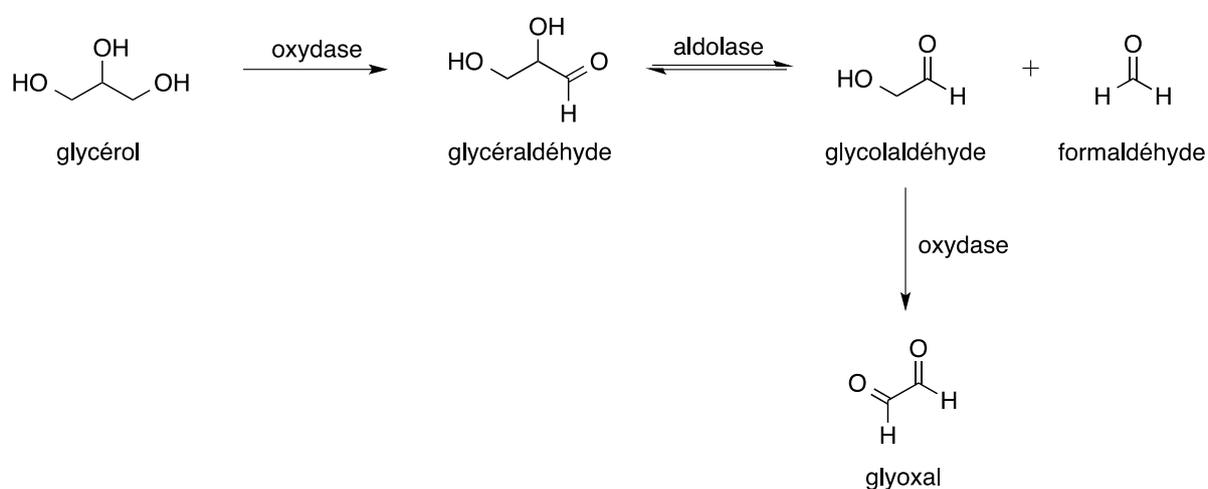


Figure 5 : voie envisagée pour la valorisation du glycérol par catalyse enzymatique

Le schéma réactionnel est discutable pour différentes raisons mais les deux principales sont la production de formaldéhyde produit CMR (cancérogène, mutagène et reprotoxique) et l'utilisation d'une aldolase qui va catalyser préférentiellement les réactions d'aldolisation et non de rétro-aldolisation.

La thèse portant sur la catalyse chimique s'est orientée sur l'oxydation du glycérol dans des conditions douces, à savoir un pH non contrôlé neutre, une température entre 60 et 100°C et une pression de 3 à 5 bars. La température optimale déterminée pour la

production du mélange racémique de glycéraldéhyde (sélectivité de 83 % pour une conversion de 10 %) est de 80°C. Le pH n'est pas contrôlé et varie entre 6 au départ de la réaction et 3 en fin de réaction.

Mon travail s'est dans un premier temps focalisé sur la réaction de clivage du glycéraldéhyde, avec la recherche d'une aldolase capable de catalyser cette réaction puis, la production et la caractérisation de cette enzyme. La fructose-6-phosphate aldolase (FSA) a été sélectionnée comme point de départ, pour la création de mutants par mutagenèse dirigée, après une étude bibliographique.

Dans un second temps, je me suis intéressé à la réaction d'oxydation du glycérol en glycéraldéhyde en utilisant une approche bio-informatique pour la recherche d'oxydase dans la biodiversité en raison de la non compatibilité des conditions réactionnelles de la partie oxydation chimique et rétro-aldolisation.

Chapitre 1 : Partie bibliographique

Ce premier chapitre s'articulera en deux parties. Une première concernant les aldolases avec un état de l'art des différentes classes d'aldolases et un historique sur la fructose-6-phosphate aldolase que j'ai utilisée pour la réaction de clivage du glycéraldéhyde. Une seconde partie portera sur les différentes méthodes chimiques et biotechnologiques de valorisation du glycérol.

I. La création de liaisons C-C par les aldolases

1. Introduction

La création de liaisons C-C (carbone-carbone) de façon stéréosélective est cruciale pour la synthèse asymétrique de squelettes moléculaires d'intérêt. Tout l'enjeu est de réussir à créer, à partir de synthons simples (achiraux), une grande diversité de molécules complexes dont la stéréochimie est parfaitement contrôlée. Pour arriver à ce résultat, les méthodes les plus conventionnelles sont l'utilisation de bases de Lewis ou encore la catalyse organométallique. Cependant, ces stratégies s'avèrent pour la plupart en désaccord avec les principes de la « chimie verte » évoqués dans l'introduction du manuscrit, notamment les conditions réactionnelles : pH neutre, faible température... C'est pourquoi, depuis plusieurs années maintenant, l'utilisation d'enzymes se développe et arrive à lever les contraintes environnementales et de sélectivités dans ce type de réaction. Ici nous allons nous focaliser sur les aldolases et plus particulièrement la fructose-6-phosphate aldolase (FSA). Il existe bien entendu d'autres classes d'enzymes capables de réaliser ce type réactions^[7] telles que les transférases^[8] et les hydroxynitrile lyases^[9].

2. Les aldolases

2.1. Définition et classification

La réaction d'aldolisation est l'une des réactions les plus importantes en synthèse organique^[10]. Elle est également indispensable d'un point de vue métabolique puisqu'elle intervient dans la synthèse des hydrates de carbones et des acides aminés^[11]. Cette réaction permet la formation de liaisons C-C tout en créant 1 ou 2 nouveaux centres asymétriques. Le

contrôle de la stéréochimie se fait généralement par l'utilisation d'acides de Lewis^[12], de catalyseurs organiques^[13] ou organométalliques^[14]. Plus récemment, une alternative visant à utiliser des catalyseurs enzymatiques, les aldolases, s'est développée en raison de ses avantages liés aux principes de la « chimie verte », tels que l'économie d'atomes, la forte stéréosélectivité (contrôlée généralement par l'aldolase et non par la structure du substrat) et spécificité. Il existe un grand nombre d'aldolases^[15], ce qui permet de synthétiser une grande diversité de produits.

Les aldolases sont des enzymes qui appartiennent à la classe des lyases (EC 4.1.2.X). Elles catalysent l'addition stéréosélective d'un composé donneur (nucléophile) sur un composé accepteur (électrophile) avec création de 1 ou 2 centres asymétriques (Figure 6).

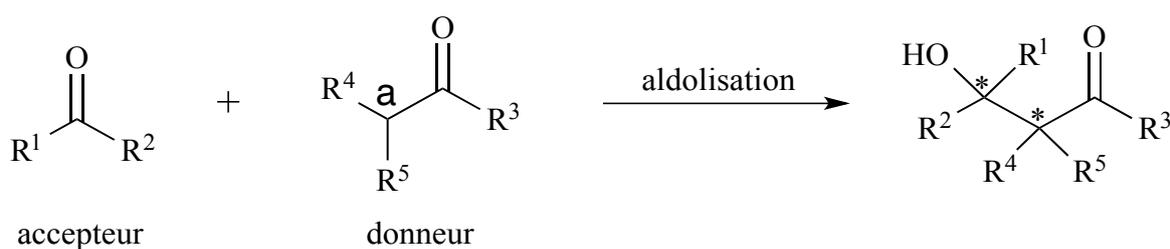


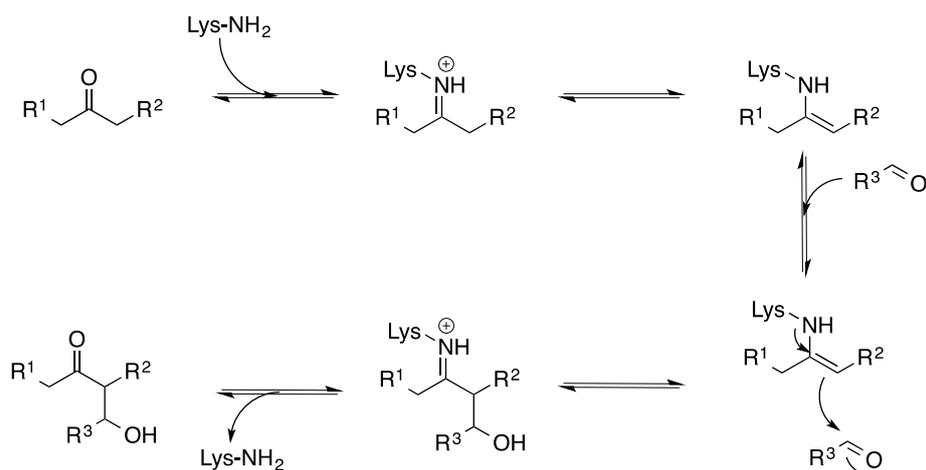
Figure 6 : réaction classique d'aldolisation

Cette réaction consiste en l'attaque du carbone alpha (nucléophile) d'un aldéhyde ou d'une cétone donneur, sur le carbonyle électrophile d'un autre aldéhyde ou cétone accepteur.

Les aldolases sont classées historiquement selon leur mécanisme^[11] (Figure 7).

- Les aldolases de classe I : l'intermédiaire réactionnel est une énamine formée entre le substrat donneur et un résidu lysine de l'enzyme.
- Les aldolases de classe II : le mécanisme réactionnel fait intervenir un cofacteur métallique divalent (Zn^{2+} , Mg^{2+} ...), l'intermédiaire est un énolate.

Mécanisme des aldolases de classe I



Mécanisme des aldolases de classe II

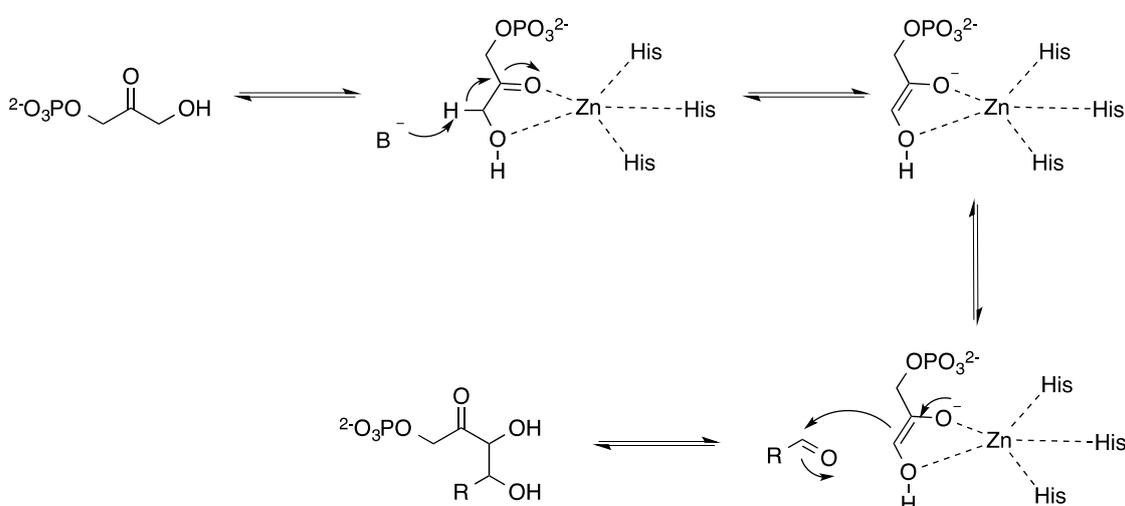


Figure 7 : mécanismes réactionnels des aldolases de classe I et II

Dans le cas d'une aldolase de classe I, le résidu lysine forme avec le donneur une énamine plus réactive que la cétone. L'azote étant moins électronégatif que l'oxygène, il peut plus facilement donner des électrons ce qui entraîne l'aldolisation. Dans le cas d'une aldolase de type II, l'enzyme utilise un métal comme cofacteur (coordonné par des résidus histidine) pour faciliter la déprotonation du substrat donneur et ainsi former l'énolate réactif.

La plupart des aldolases ont une très forte spécificité vis-à-vis du substrat donneur, ce qui a entraîné une refonte de la classification cette fois-ci en fonction des substrats et non du mécanisme d'action^[10] (Tableau 1). Cependant, ces enzymes sont capables d'accepter également différentes structures de substrat accepteur.

Donneur	Accepteur	Produit	Enzyme
$\text{HO}-\text{CH}_2-\overset{\text{O}}{\parallel}{\text{C}}-\text{CH}_2-\text{OH}$ DHA et dérivés	$\text{HO}-\text{CH}_2-\overset{\text{O}}{\parallel}{\text{C}}-\text{OPO}_3^{2-}$ DHAP	$\text{R}-\overset{\text{OH}}{\underset{\text{OH}}{\text{C}}}-\overset{\text{O}}{\parallel}{\text{C}}-\text{CH}_2-\text{OH}$	DHA(P) aldolases
$\text{O}=\text{C}-\text{CO}_2\text{H}$ ou $\text{OPO}_3^{2-}-\text{C}=\text{C}-\text{CO}_2\text{H}$ Pyruvate ou phosphoenolpyruvate	$\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{H}$	$\text{R}-\overset{\text{OH}}{\text{C}}-\text{CH}_2-\overset{\text{O}}{\parallel}{\text{C}}-\text{CO}_2\text{H}$	Pyruvate et phosphopyruvate aldolases
$\text{O}=\text{C}-\text{H}$ Acétaldéhyde	$\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{H}$	$\text{R}-\overset{\text{OH}}{\text{C}}-\text{CH}_2-\overset{\text{O}}{\parallel}{\text{C}}-\text{H}$	Acétaldéhyde aldolases
$\text{H}_2\text{N}-\text{CH}_2-\text{COOH}$ Glycine	$\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{H}$	$\text{R}-\overset{\text{OH}}{\underset{\text{NH}_2}{\text{C}}}-\overset{\text{O}}{\parallel}{\text{C}}-\text{OH}$	Glycine aldolases

Tableau 1 : nouvelle classification des aldolases

2.2. Les pyruvate et phospho-énolpyruvate aldolases

Cette famille d'aldolases fût découverte en 1972^[16]. *In vivo*, ces enzymes ont une fonction de dégradation d'acide 3-deoxy-2-ulosonique. Elles interviennent aussi dans la biosynthèse des acides carboxyliques. Ces enzymes catalysent l'aldolisation réversible du pyruvate (Pyr) ou dérivés^[17], tels que le phosphoenolpyruvate (PEP), sur différents substrats accepteurs. Elles sont pour la majorité de classe I, cependant il en existe un petit nombre de classe II dont le potentiel est moins important^[18,19] en raison de la spécificité de substrat. Les PEP (Tableau 2) et Pyr (Tableau 3) aldolases peuvent être utilisées en synthèse organique avec un fort déplacement de l'équilibre vers le produit d'addition. Cependant, pour les PEP aldolases, le déplacement de l'équilibre peut être irréversible suite au relargage des groupements phosphates.

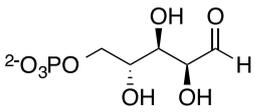
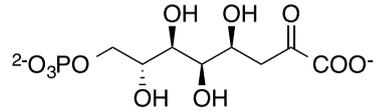
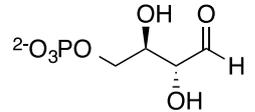
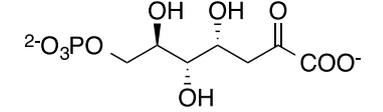
Accepteurs naturels	Enzymes	Produits
 D-arabinose-5-phosphate	3-deoxy-D-manno-oculosonate-8-phosphate (KDOP) synthétase (EC 2.5.1.55)	 3-deoxy-D-manno-oculosonate-8-phosphate (KDOP)
 D-érythrose-4-phosphate	3-deoxy-D-arabino-2-heptulosonate-7-phosphate (DAHP) synthétase (EC 2.5.1.54)	 3-deoxy-D-arabino-2-heptulosonate-7-phosphate (DAHP)

Tableau 2: principales PEP aldolases

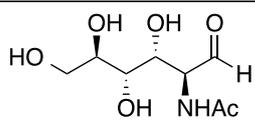
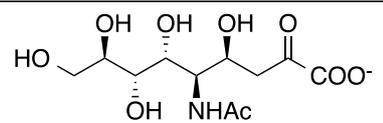
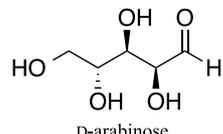
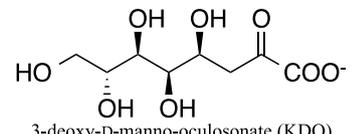
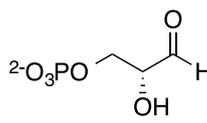
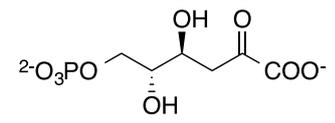
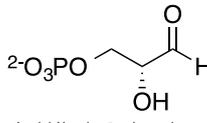
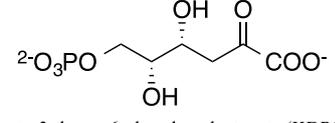
Accepteurs naturels	Enzymes	Produits
 N-acétyl mannosamine	N-acétyl neuraminic (NeuAc) aldolase (NeuA, EC 4.1.3.3)	 Acide N-acétyl neuraminique (NeuAc)
 D-arabinose	3-deoxy-D-manno-oculosonate (KDO) aldolase (EC 4.1.2.23)	 3-deoxy-D-manno-oculosonate (KDO)
 D-glycéraldéhyde-3-phosphate (DG3P)	2-ceto-3-deoxy-6-phosphogluconate (KDPG) aldolase (EC 4.1.2.14)	 2-ceto-3-deoxy-6-phosphogluconate (KDPG)
 D-glycéraldéhyde-3-phosphate (DG3P)	2-ceto-3-deoxy-6-phosphogalactonate (KDPGal) aldolase (EC 4.1.2.21)	 2-ceto-3-deoxy-6-phosphogalactonate (KDPGal)

Tableau 3 : principales Pyr aldolases

Ces enzymes ont été étudiées pour le développement de synthèses stéréocontrôlées d'acides non naturels^[20].

2.3. Les Acétaldéhyde aldolases

Dans ce groupe d'aldolases, une seule enzyme est connue pour le moment la 2-deoxyribose-5-phosphate aldolase^[21,22] (DERA, EC 4.1.2.4). C'est une enzyme de classe I qui a la particularité de réaliser des réactions d'aldolisation entre deux aldéhydes. La réaction native consiste en l'aldolisation réversible de l'acétaldéhyde et du D-glycéraldéhyde-3-

phosphate(DG3P) pour donner le 2-deoxyribose-5-phosphate avec un équilibre encore une fois déplacé vers la réaction de l'addition.

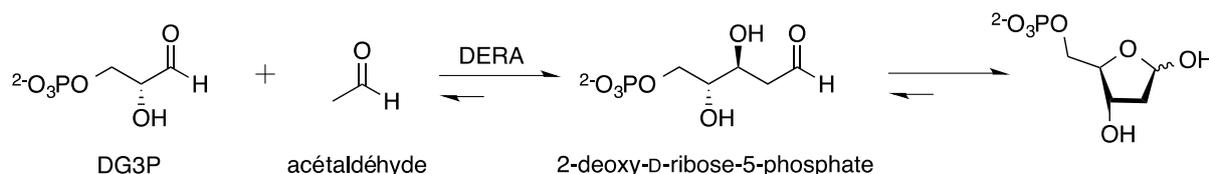


Figure 8 : réaction native de la DERA

Cependant, la DERA possède l'avantage de tolérer de nombreux substrats non naturels accepteurs^[22] (aldéhyde aliphatique, 2-hydroxyaldéhyde...) ainsi que différents substrats donneurs ce qui est assez rare pour les aldolases^[11,23-25] (Figure 9). Il est à noter que les énantiomères D sont en général de meilleurs substrats que les énantiomères L.

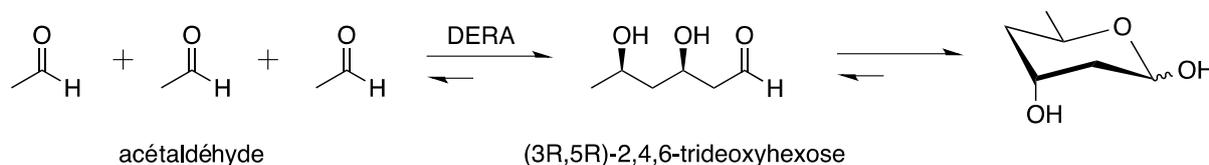


Figure 9 : cross-aldéhyde addition de l'acétaldéhyde par la DERA

La DERA sert aussi pour la synthèse de déoxy sucres car elle accepte notamment des substrats chlorés ou comportant des azotures en bout de chaîne. La DERA est donc un outil intéressant pour la synthèse organique.

2.4. Les Glycine aldolases

Cette classe regroupe deux types d'enzymes, les thréonine aldolases^[26] (ThrA, EC 4.1.2.5) et les sérine hydroxyméthyl-transférases^[27] (SHMT, EC 2.1.2.1). Ce sont des enzymes qui utilisent de manière indispensable le phosphate de pyridoxal (PLP) comme cofacteur de la même façon que les transaminases. Elles catalysent de façon réversible l'aldolisation de la glycine sur différents aldéhydes pour former deux nouveaux carbones asymétriques. Cependant, il n'est pas possible à l'heure actuelle d'obtenir avec des ThrA naturelles les 4 stéréoisomères de façon sélective^[28,29]. En effet, alors que la L-ThrA et la L-allo-ThrA catalysent la formation de composés de configuration L-thréo et L-érythro respectivement, les D-low ThrA^[26,30] (faible spécificité) ne sont pas sélectives sur le carbone β (Figure 10).

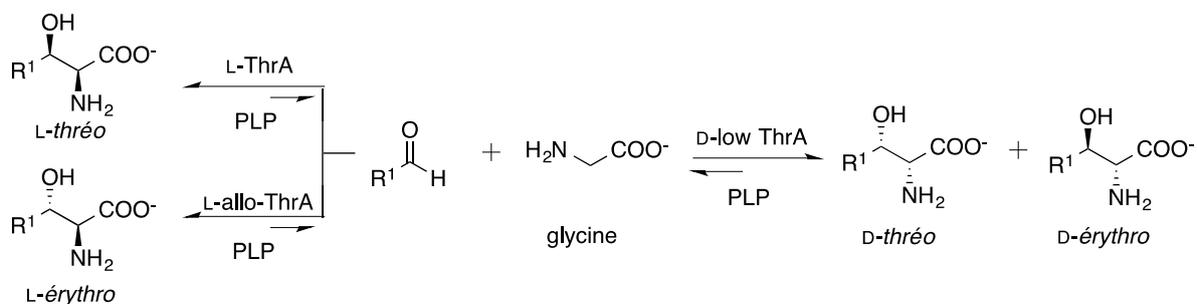


Figure 10 : mise en évidence des stéréoisomères obtenus par l'aldolisation de la glycine sur un aldéhyde accepteur catalysée par les différentes thréonine aldolases naturelles

Cependant, l'obtention d'un mélange de deux stéréoisomères^[31] peut dans certains cas être intéressant. Les enzymes de ce groupe tolèrent de nombreux aldéhydes accepteurs comme des aldéhydes aliphatiques, aromatiques mais pas d'aldéhydes α,β insaturés^[28,32].

2.5. DHAP aldolases

Ce groupe contient de nombreuses enzymes capables d'accepter une grande variété de substrats. Les DHAP aldolases sont étudiées depuis longtemps car elles présentent un grand intérêt pour la synthèse de polyols chiraux. L'utilisation de ces enzymes conduit à la création d'une liaison C-C et à la formation de deux nouveaux centres chiraux. Il est possible d'obtenir les 4 stéréoisomères. Le contrôle de la stéréochimie du C3 (carbone numéro 3 en suivant les principes de nomenclature) est très élevé, cependant la stéréochimie du C4 reste plus « aléatoire » notamment pour la tagatose-1,6-*bis*phosphate aldolase^[33,34]. On peut distinguer 4 principales enzymes dans les DHAP aldolases (Figure 11), la fructose-1,6-*bis*phosphate aldolase^[35] (FruA, EC 4.1.2.13), la fucose-1-phosphate aldolase^[36] (FucA, EC 4.1.2.17), la tagatose-1,6-*bis*phosphate aldolase^[33] (TagA, EC 4.1.2.40) et la rhamnULOse-1-phosphate aldolase^[37] (RhuA, EC 4.1.2.19). La FruA et la TagA catalysent nativement l'aldolisation du DHAP sur le DG3P pour former respectivement le D-fructose-1,6-*bis*phosphate (3S, 4R) et le D-tagatose-1,6-*bis*phosphate (3S, 4S) alors que la FucA et la RhuA catalysent nativement l'aldolisation du DHAP sur le L-lactaldéhyde pour former respectivement le L-fucose-1-phosphate (3R, 4R) et le L-rhamnULOse-1-phosphate (3R, 4S). Ces enzymes sont utilisées en synthèse organique car elles possèdent un large panel de substrats accepteurs^[38]. Cependant, leur très forte dépendance au DHAP entraîne quelques difficultés : le DHAP est très coûteux et relativement instable.

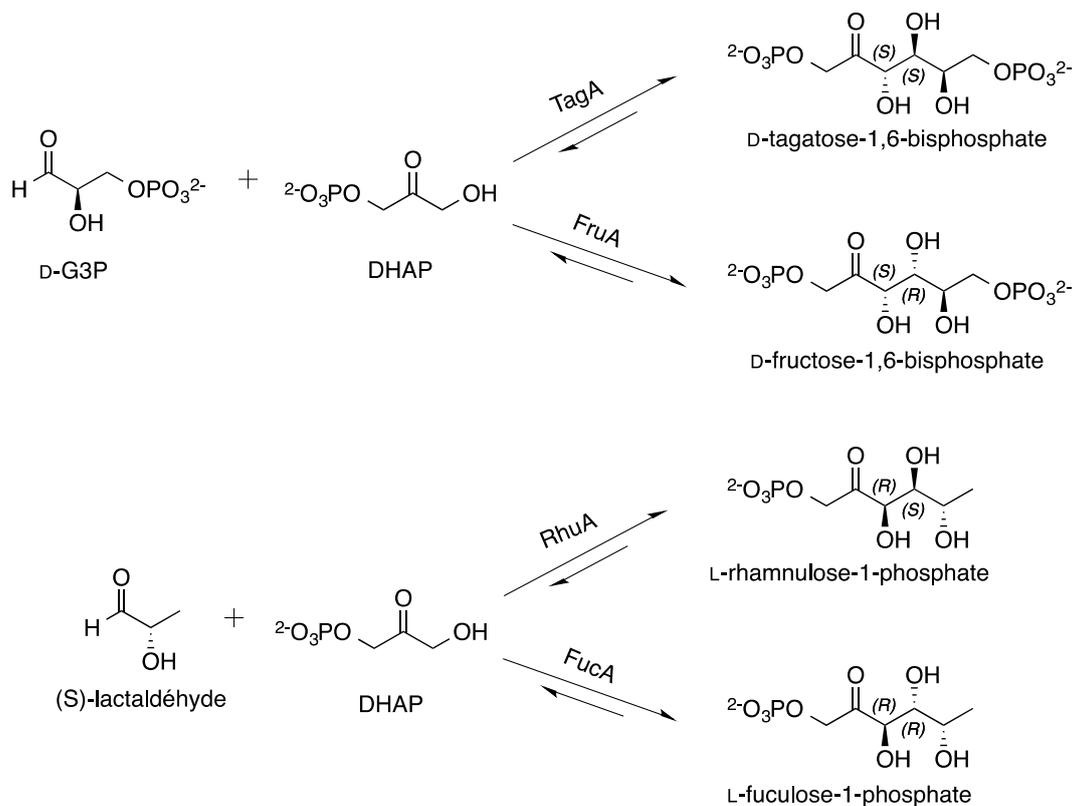


Figure 11 : complémentarité des DHAP aldolases

Les réactions d'aldolisation enzymatiques sont toujours réversibles, cependant l'équilibre est souvent en faveur de la synthèse. Certains paramètres comme la concentration en substrat, peuvent jouer sur cet équilibre.

Le DHAP est une molécule relativement chère (1240 euros le gramme chez Sigma Aldrich) ; des méthodes de génération *in situ* du DHAP ont donc été étudiées notamment grâce à l'action d'acide phosphatase sur des substrats phosphorylés bon marché^[39] mais aussi de façon chimique à partir de dérivés du DHA ou du glycérol^[40].

Récemment, la synthèse d'alcools tertiaires a été réalisée en utilisant une rhamnulose-1-phosphate aldolase de *Bacteroides thetaiotamicron* (RhuA_{BTHET}). Cette découverte est très importante et inattendue puisque c'est la première fois qu'une « réaction d'aldolisation » entre deux cétones a été décrite (Figure 12)^[41].

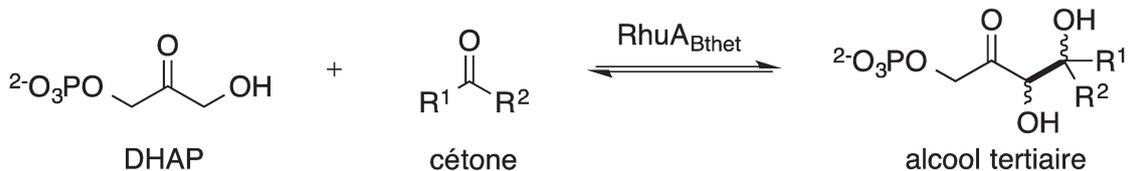


Figure 12 : synthèse d'alcools tertiaires catalysée par la Rhamnulose-1-phosphate aldolase de *Bacteroides thetaiotamicron*

3. La fructose-6-phosphate aldolase

3.1. Structure et mécanisme de la FSA_{wt}

La fructose-6-phosphate aldolase (FSA) est une aldolase de classe I, qui fait partie des DHA aldolases découvertes chez *Escherichia coli* (*E.coli*) par Schürmann et Sprenger^[42] en 2001. Son rôle biologique reste encore inconnu mais la réaction qui l'a mise en évidence est la rétro-aldolisation du D-fructose-6-phosphate (DF6P) en dihydroxyacétone (DHA) et D-glycéraldéhyde-3-phosphate (DG3P) (Figure 13).

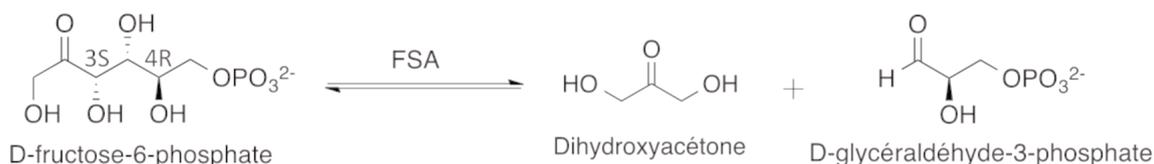


Figure 13 : réaction "native" de la FSA

Depuis la découverte de la FSA, d'autres DHA aldolases ont été découvertes dans différents organismes^[43] mais aussi une deuxième dans *E.coli*, la FSAB^[44]. Ce travail de thèse s'est focalisé sur la FSAA de *E.coli* découverte par Schürmann et Sprenger qui sera décrite plus précisément par la suite.

Si l'on regarde la structure de la FSA, elle se présente sous la forme d'un homodécamère^[45] (Figure 14a) d'environ 257 KDa.

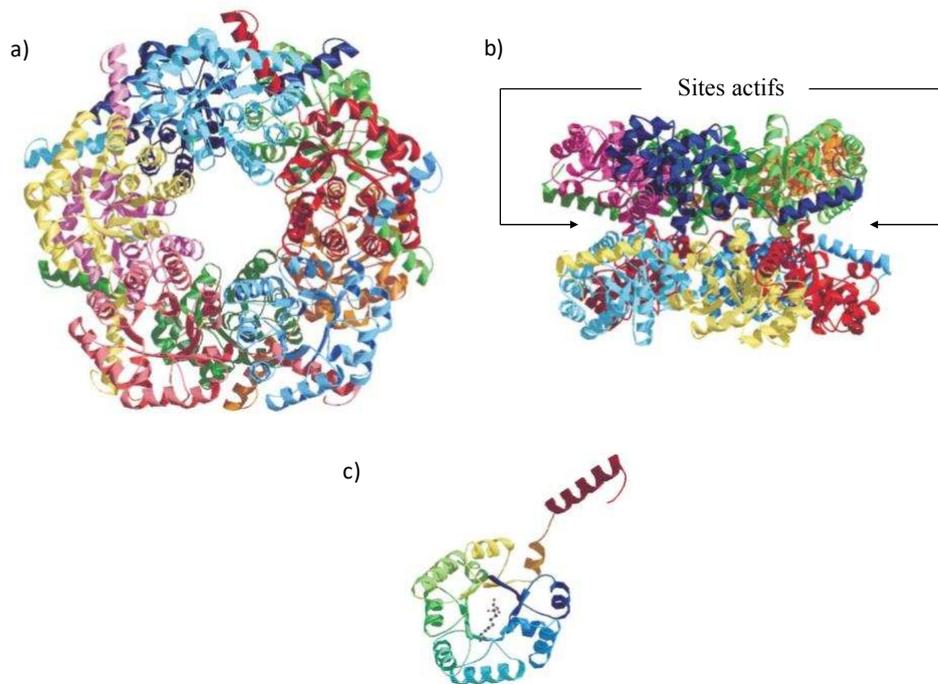


Figure 14 : structure 3D de la FSA avec en a) une vue du dessus, en b) une vue de profil pour mettre en évidence les sites actifs et en c) la représentation d'un seul monomère^[45]. Pour les figures a et b chaque couleur représente 1 monomère (code couleur de la base de données de protéines (PDB, protein data bank))

Elle est constituée de deux pentamères empilés de telle sorte que les sites actifs se situent entre un monomère (Figure 14c) de chaque pentamère (Figure 14b).

Le mécanisme de rétro aldolisation fait intervenir le groupement amine d'un résidu lysine (position 85) de la FSA qui va former avec le carbonyle du DF6P un iminium. Sous l'action d'une base (B⁻), il va y avoir clivage du DF6P en DG3P et formation d'un carbanion qui va se reprotonner grâce à un acide (HB). Une étape finale d'hydrolyse va enfin libérer le DHA et régénérer la lysine (Figure 15). L'enzyme est stéréosélective et conduit uniquement à l'isomère 3S, 4R lors de l'aldolisation.

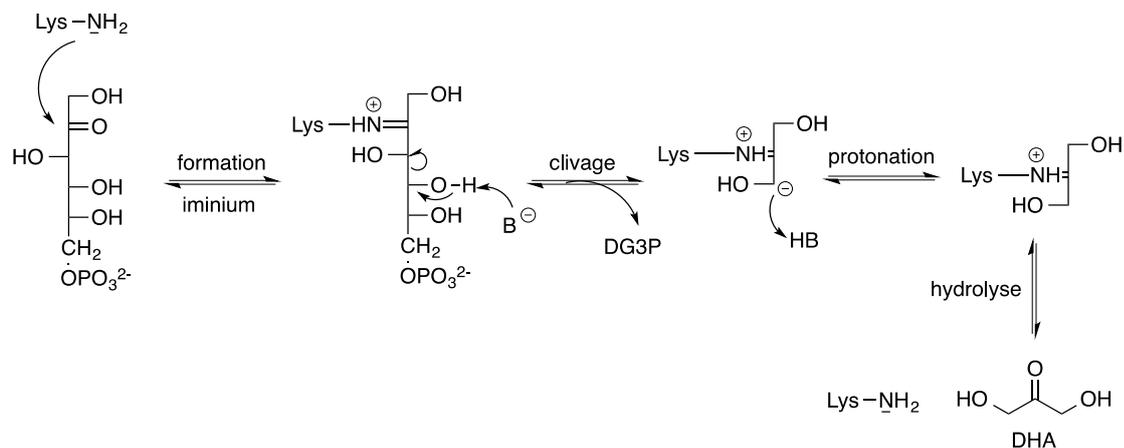


Figure 15 : mécanisme de rétro aldolisation catalysée par la FSA

L'activité de tous les résidus du site actif (Figure 16) n'est pas encore déterminée. Cependant certains résidus ont tout de même un rôle défini dans le mécanisme réactionnel. La lysine en position 85 va former l'iminium (Figure 15) qui va favoriser le clivage dans la rétro-aldolisation par stabilisation du carbanion grâce à son effet mésomère -M. De plus la présence de l'alcool qui possède un effet inductif -I stabilise aussi le carbanion. Des travaux ont mis en évidence le résidu tyrosine en position 131 qui est indispensable à l'activité de rétro-aldolisation^[46].

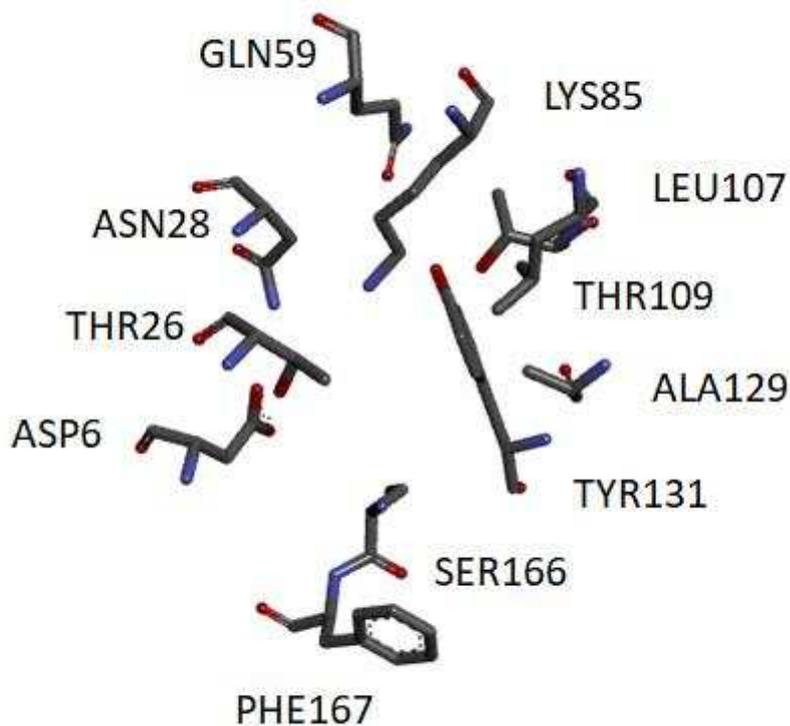


Figure 16 : site actif de la FSA_{wt} avec en rouge les atomes d'oxygène, en gris les atomes de carbone et en bleu les atomes d'azote

En effet, le groupement carbonyle du résidu glutamine (GLN) en position 59 active la tyrosine (TYR) en position 131 par l'intermédiaire d'une molécule d'eau. Cette interaction va rendre le proton du groupement alcool de la tyrosine 131 labile et permettre ainsi la reprotonation du carbanion (Figure 15) mais aussi d'autres échanges de protons. Les résidus Asparagine (ASN) en position 28 et acide aspartique (ASP) en position 26 sont décrit comme favorisant le positionnement du substrat dans le site actif par interactions hydrogènes. La thréonine (THR) en position 109 intervient dans le processus de formation de l'iminium par interactions hydrogènes sur le carbonyle du substrat lors de la déshydratation^[47].

3.2. Historique des travaux sur la FSA

3.2.1 La FSA_{wt}

L'un des grands avantages de la FSA « wild type » (FSA_{wt}) est qu'elle tolère différents substrats donneurs et accepteurs sous sa forme native. Cette particularité en fait

un outil très intéressant pour la synthèse organique. De plus, la FSA_{wt} est thermostable jusqu'à 60°C avec un temps de demi-vie de 16 h à 75°C^[44,48] ce qui ajoute encore plus d'intérêt à l'utilisation de cette enzyme notamment pour combiner son pouvoir catalytique avec de la catalyse chimique. Une première série d'étude menée par Schürmann et Sprenger^[48] a permis de mettre en évidence plusieurs aldéhydes accepteurs de la FSA_{wt} (Figure 17) ainsi qu'un nouveau donneur, l'hydroxyacétone (HA)(Figure 18).

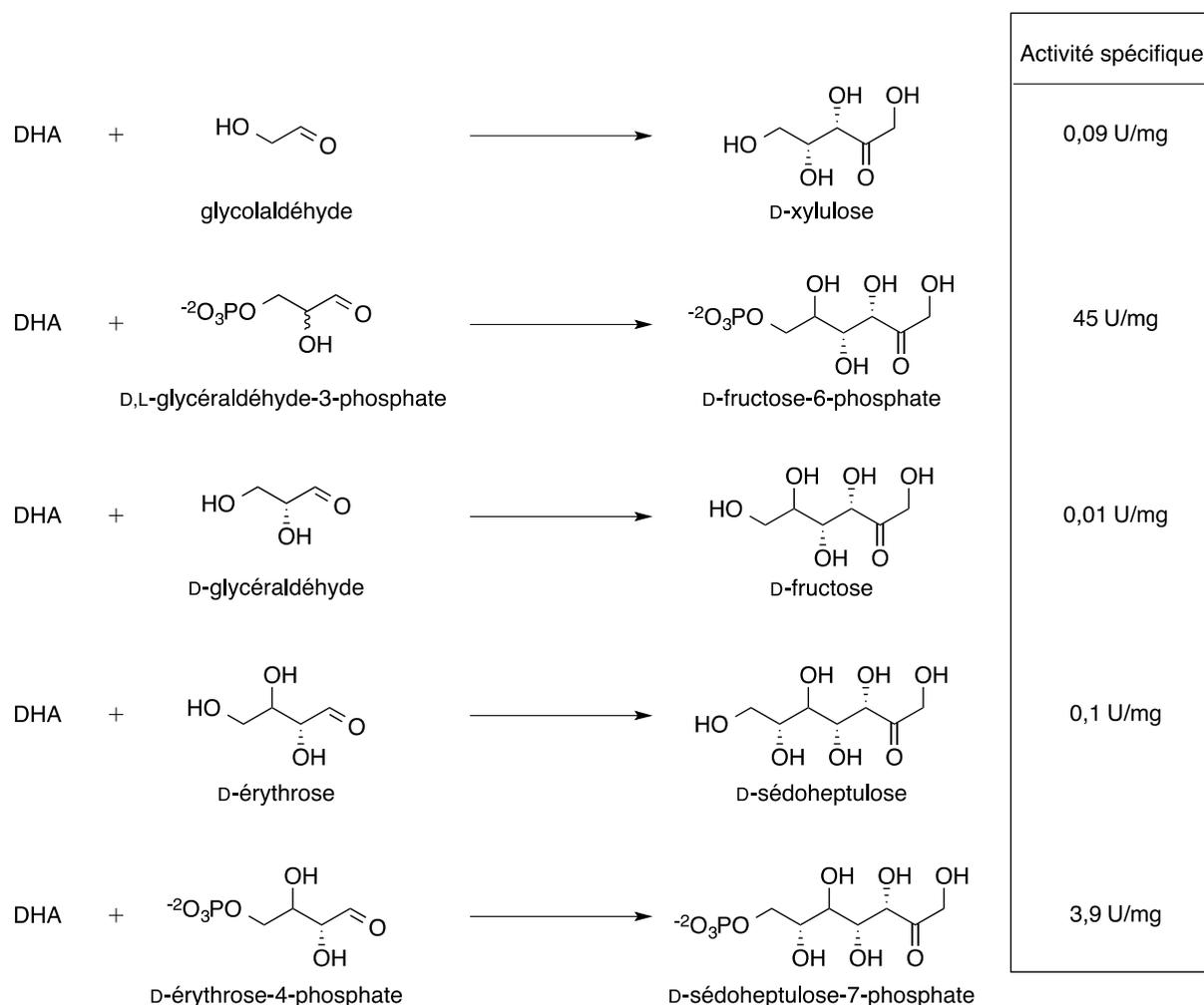


Figure 17 : réactions catalysées par la FSA_{wt} avec le DHA en tant que donneur et différents aldéhydes

On remarque que la FSA_{wt} possède de meilleures activités spécifiques lorsque le substrat accepteur est phosphorylé et que le DHA est utilisé comme donneur. De plus, le passage d'un accepteur à 3 carbones (G3P) à un accepteur à 4 carbones (D-érythrose-4-phosphate) divise l'activité de la FSA_{wt} par 11. L'activité diminue fortement lorsque le produit de la réaction dépasse 6 carbones puisqu'elle passe de 45 U.mg⁻¹ (synthèse F6P) à 3,9 U.mg⁻¹

(synthèse du D-sédoheptulose-7-phosphate) même si cela reste largement acceptable pour réaliser des synthèses.

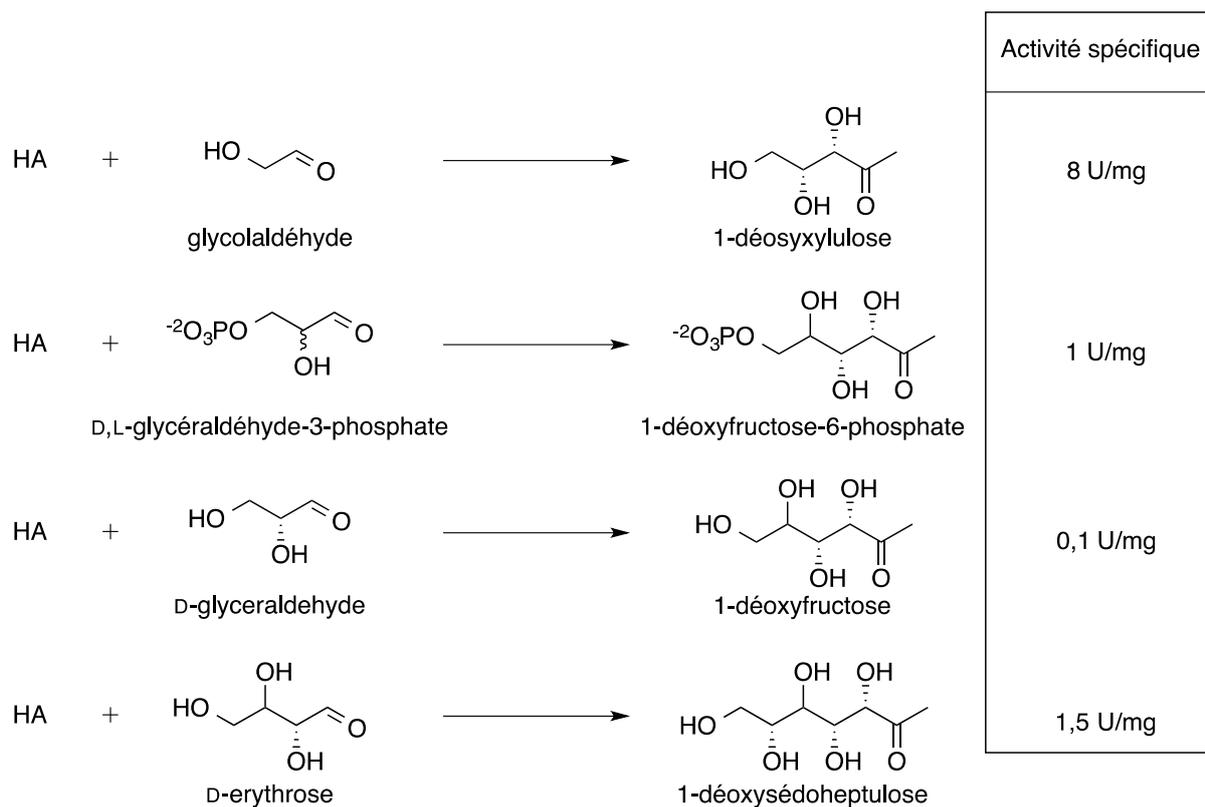


Figure 18 : réactions catalysées par la FSA_{wt} avec HA en donneur et différents aldéhydes accepteurs

La FSA_{wt} est donc capable d'utiliser un autre donneur, le HA, mais cette fois-ci le meilleur accepteur devient le glycolaldéhyde pour former le 1-déoxyxylulose avec une activité spécifique de 8 U.mg⁻¹. L'utilisation de HA entraîne une augmentation de l'activité spécifique pour les mêmes accepteurs non phosphorylés étudiés avec le DHA comme donneur. Que ce soit avec HA ou DHA comme donneur, l'équilibre des réactions est fortement en faveur de la synthèse des produits. Cette première étude de la FSA_{wt} a permis de mettre en évidence sa capacité à synthétiser des sucres 1-déoxy qui sont précurseurs de différents isoprénoides, vitamine B1 et B6^[49].

Cependant, ces études ont été par la suite en partie « invalidées »^[44,51]. En effet, le meilleur substrat donneur est le HA avec un k_{cat}/k_M de 145 s⁻¹mM⁻¹ lorsque le D,L-glycéraldéhyde-3-phosphate est utilisé comme donneur. Enfin le meilleur accepteur est le D,L-glycéraldéhyde-3-phosphate avec un k_{cat}/k_M de 216 s⁻¹mM⁻¹ avec la dihydroxyacétone comme donneur.

En 2007, Sugiyama *et al.*^[51] ont mis en évidence la possibilité pour la FSA_{wt} d'utiliser des substrats accepteurs de type azido-aldéhyde et Cbz-amino aldéhyde mais aussi un nouveau substrat donneur le 1-hydroxy-2-butanone (HB). Cette capacité est très intéressante puisqu'elle permet de réaliser la synthèse d'iminocyclitols, inhibiteurs de glycosidases, par catalyse hybride one-pot avec un catalyseur métallique (palladium sur charbon). Trois azido-aldéhydes ont été testés avec les 3 substrats donneurs de la FSA (DHA, HA, HB). Le rendement ne dépend pas du donneur puisqu'ils sont du même ordre de grandeur. Cependant, lorsque l'on augmente la complexité de l'azido-aldéhyde accepteur, le rendement chute de moitié (Figure 19).

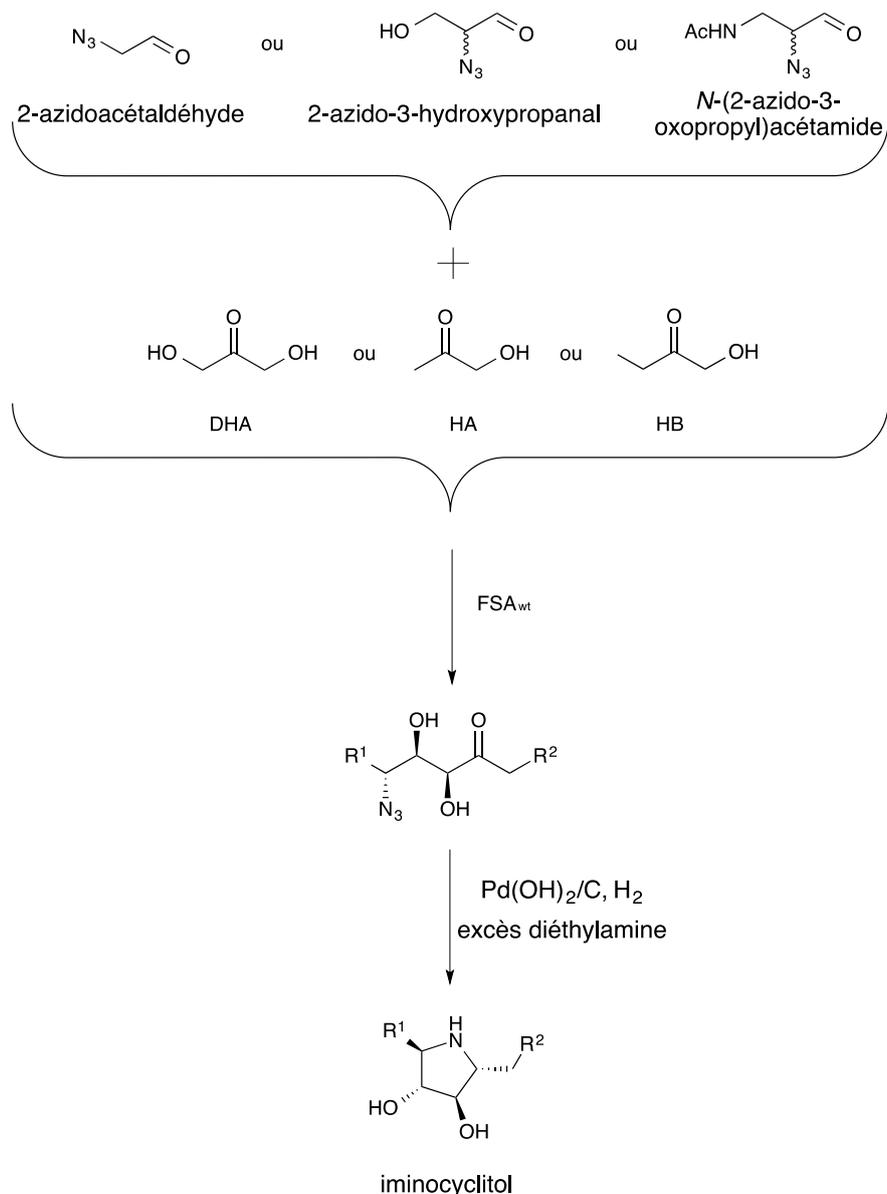


Figure 19 : synthèse one-pot d'imino-cyclitols catalysée par la FSA_{wt} et palladium sur charbon avec R¹ = H, CH₂OH, CH₂NHAc et R² = OH, H, CH₃ à partir de substrats accepteurs azido-aldehydes

La synthèse d'imino-cyclitols avec des aminoaldéhydes protégés par des groupements Cbz (carboxybenzyl) a également été étudiée. Trois aldéhydes accepteurs énantiomériquement purs ont été testés en utilisant toujours les mêmes donneurs « classiques » de la FSA. De même que pour les accepteurs de type azido-aldéhyde, le donneur ne joue pas un rôle fondamental dans le rendement mais c'est l'encombrement de l'accepteur qui entraîne les variations. Le rendement passe de 70-80 % pour l'accepteur non encombré à 50-60 % pour les plus encombrés. La stéréochimie de l'aldéhyde accepteur n'a également pas d'influence sur le rendement (Figure 20).

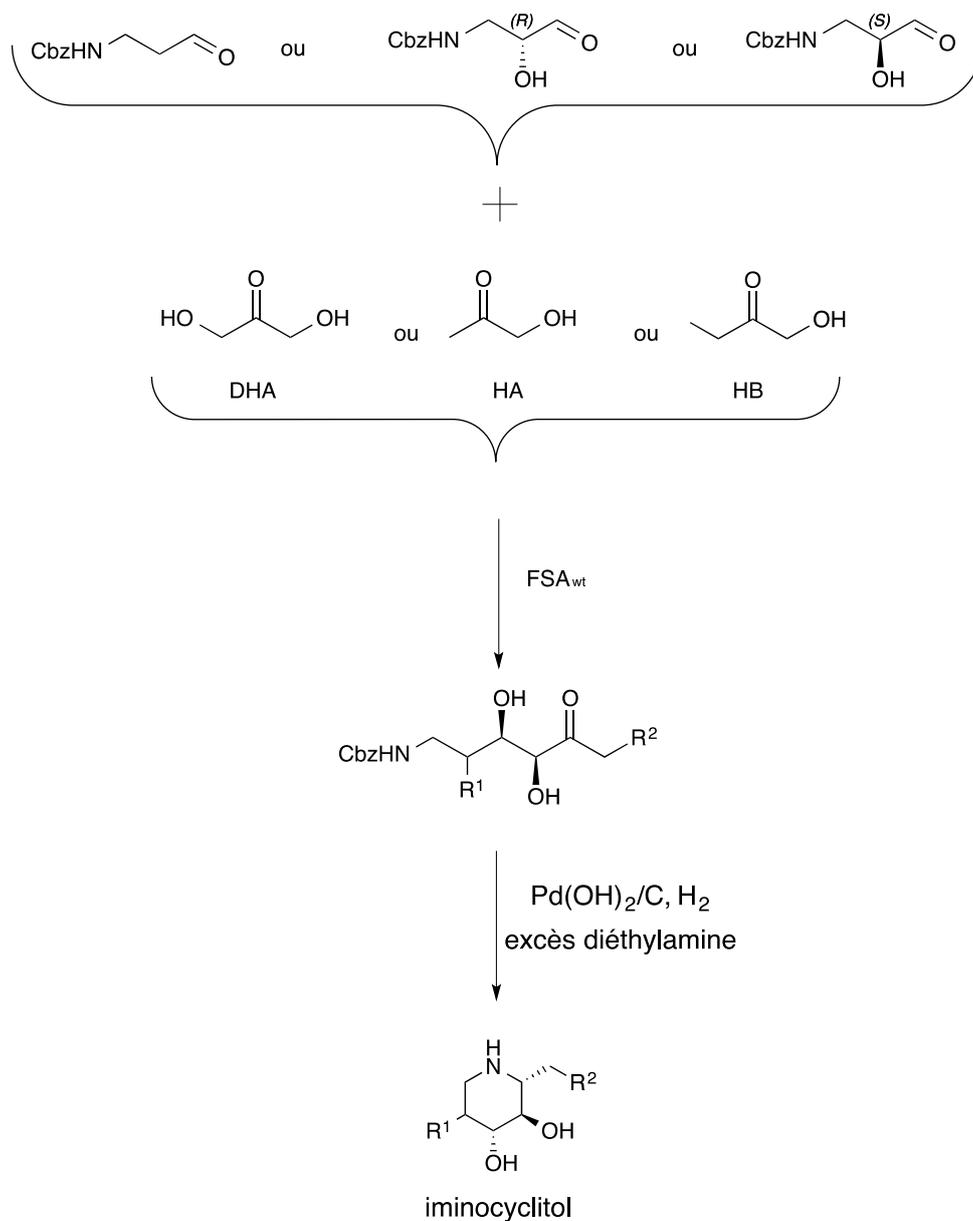


Figure 20 : synthèse one-pot d'imino-cyclitols catalysée par la FSA_{WT} et palladium sur charbon avec R¹ = H, OH et R² = CH₂OH, CH₃, CH₂CH₃

En 2009, Garrabou *et al.*^[50] ont mis en évidence la capacité de la FSA_{wt} à utiliser le glycolaldéhyde (GA) comme donneur et comme accepteur. De plus, l'enzyme est également capable de réaliser des réactions d'aldolisation entre deux aldéhydes ce qui n'est pas commun. La synthèse de 1-déoxy-D-ido-hept-2-ulose a notamment été réalisée en deux étapes par la FSA_{wt} en réacteur one-pot : une première réaction d'aldolisation du glycolaldéhyde sur lui-même pour donner du D-(-)-thréose puis une deuxième réaction d'aldolisation de cette molécule avec HA (Figure 21).

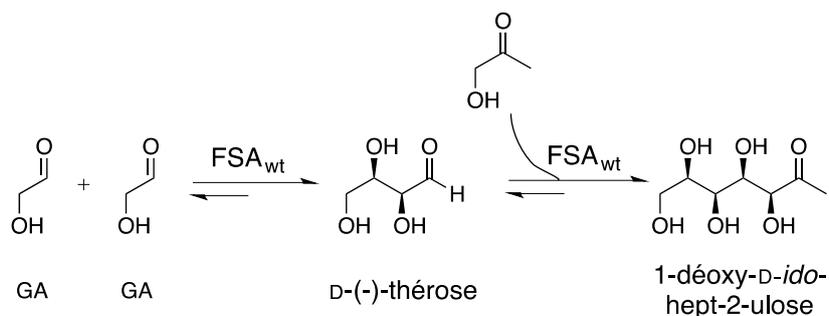


Figure 21 : catalyse de la dimérisation du GA par la FSA_{wt} suivie de l'aldolisation de HA pour former le 1-déoxy-D-idohept-2-ulose

En 2015, Soler *et al.*^[52] ont réalisé différentes mutations pour découvrir de nouveaux substrats accepteurs plus imposants, notamment en augmentant l'espace dans le site actif. Cette fois-ci, les substrats amino-aldéhydes protégés par des groupements Cbz sont cycliques et comportent 5 ou 6 carbones. Les amines protégées sont incluses dans le cycle et les groupements aldéhydes sont branchés au cycle avec des chaînes plus ou moins longues (Figure 22).



Figure 22 : substrats accepteurs cycliques amino-aldéhydes protégés par groupement Cbz avec R¹ = H, (R ou S) OH et R² = H, (R ou S) OH

Les taux de conversion obtenus sont au minimum de 80 % dans la majorité des cas, cependant des mélanges épimériques sont souvent obtenus. Les deux principaux mutants retenus pour leur activité sont A129S/R134X/A165G/S166G et L107Y/A129G/R134X/A165G/S166G avec X équivalent à R, V P ou S. Une différence est aussi à noter au niveau des substrats donneurs puisque en plus du DHA et HA, le GA a aussi été testé avec succès.

La FSA_{wt} a aussi été utilisée pour catalyser la formation de déoxynitrocyclitols. En effet trois nouveaux substrats de type nitro-aldéhyde se sont avérés être d'excellents accepteurs^[53]. Les rendements sont importants (80 %) cependant, la cyclisation n'étant pas stéréo-contrôlée des diastéréoisomères sont obtenus (Figure 23). Le mutant A129S utilisé avec les mêmes accepteurs et le DHA comme donneur conduit à la formation de

nitrocyclitols avec des rendements un peu plus faibles de l'ordre de 70 %. Ce mutant fût le premier réalisé et ses différentes activités seront détaillées par la suite (Figure 24).

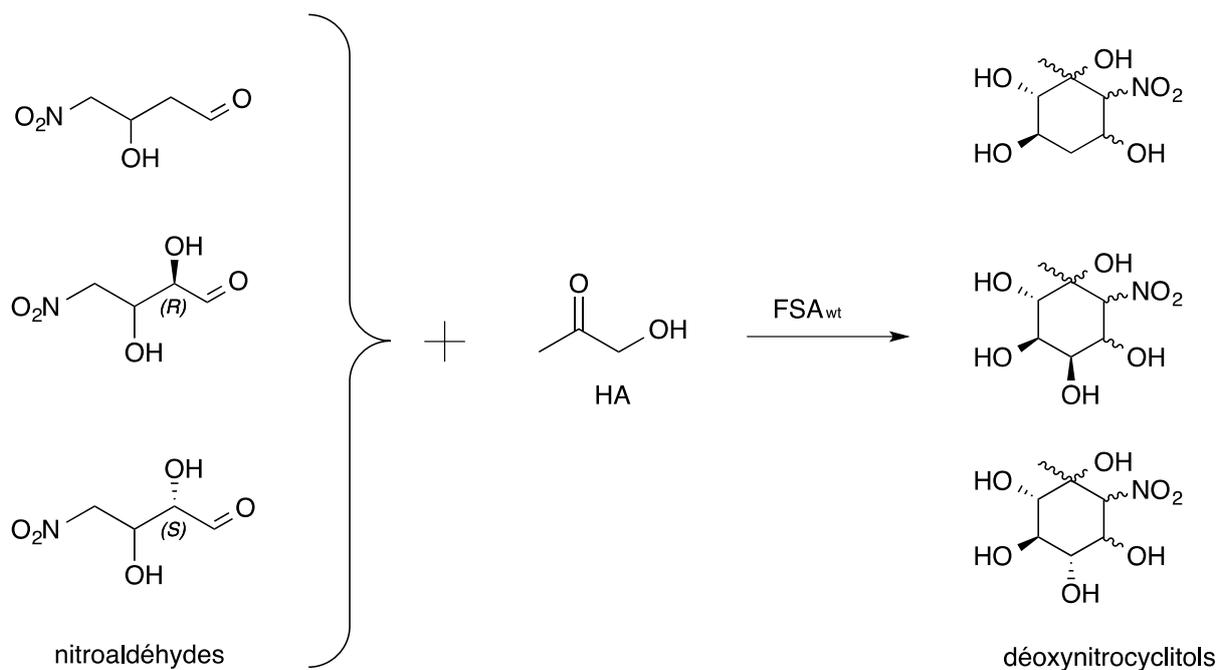


Figure 23 : synthèse de déoxynitrocyclitols catalysée par la FSA_{wt}

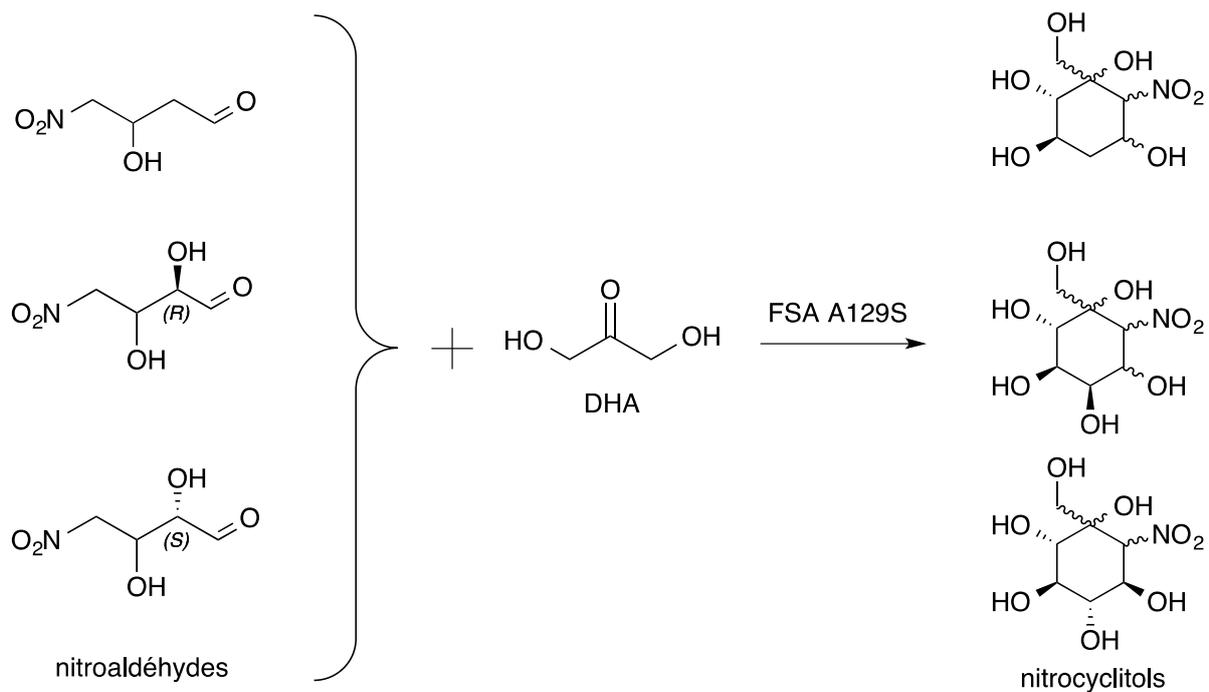


Figure 24 : synthèse de nitrocyclitols catalysée par la FSA A129S

Enfin, il a été démontré que la FSA_{wt} pouvait aussi synthétiser des précurseurs de thio-sucres^[54]. En effet Fanton *et al.* ont réalisé la synthèse du 5-chloro-5-déoxy-D-xylulose et du 5-bromo-5-déoxy-D-xylulose avec des rendements de 84 % en 36 h en utilisant 20 U (quantité d'enzyme qui catalyse la formation d'1 µmole de substrat par minute) de FSA_{wt} et le DHA comme substrat donneur. Une substitution nucléophile de type 2 (SN2) avec du sulfate de sodium est réalisée pour obtenir le thio-sucre (Figure 25).

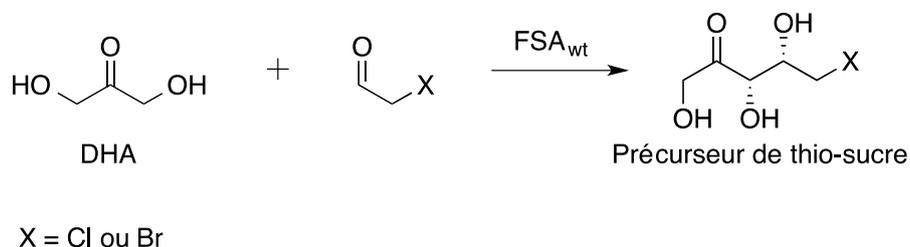


Figure 25 : synthèse de précurseurs de thio-sucres catalysée par la FSA_{wt}

La FSA, sous sa forme native, est un catalyseur très intéressant puisqu'elle tolère à la fois une grande diversité d'accepteurs (aldéhydes phosphorylés, nitrés, azidés...) et plusieurs donneurs différents (DHA, HA, HB, GA). C'est pourquoi de nombreuses équipes ont utilisé cette enzyme pour augmenter son panel de substrats donneurs déjà bien fourni. La partie du site actif qui contient le substrat donneur est relativement hydrophobe (alanine en position 129, leucine en position 107). La stratégie a donc été de réaliser des mutations pour la rendre plus hydrophile (Figure 16).

3.2.2 Les mutants de la FSA

Comme annoncé précédemment, le mutant de la FSA A129S fut le premier synthétisé. Cette mutation de l'alanine en sérine en position 129 de la FSA a eu pour but d'augmenter l'activité de la FSA vis-à-vis du DHA en créant, par l'intermédiaire de la sérine, de potentielles interactions hydrogènes^[55]. Le mutant obtenu s'est révélé être presque 20 fois plus actif que la FSA_{wt} vis-à-vis du DHA (Tableau 4).

Substrat	k_{cat}/K_M FSA _{wt} (s ⁻¹ .mM ⁻¹)	k_{cat}/K_M FSA A129S (s ⁻¹ .mM ⁻¹)
DHA	4	69
HA	145	41
GA donneur	83	25
GA accepteur	0,3	0,4
4-nitrobutanal	0,5	1
Formaldéhyde	Non déterminé	0,4
G3P	216	4,22

Tableau 4 : comparaison des paramètres cinétiques de la FSA_{wt} et de la FSA A129S pour différents substrats

Ce mutant a également été utilisé pour réaliser la synthèse de nitrocyclitols et de divers polyols en utilisant comme donneur HA, GA ou DHA avec des rendements isolés de l'ordre 80 %. Ce premier travail de mutation sur la FSA et les premiers résultats montrent aussi que le GA peut être utilisé à la fois comme donneur et comme accepteur.

Ces études ont lancé une série de travaux de la part de Clapés *et al.* sur l'utilisation du GA par la FSA^[56]. En effet, en s'appuyant sur les travaux réalisés sur le mutant A129S, cette équipe a créé différentes combinaisons de mutations sur 3 positions 107, 129, 165 pour augmenter l'activité de la FSA à réaliser la synthèse d'aldoses. La stratégie utilisée est la mutagenèse par saturation de sites sur la position 129. Trois mutants sont ressortis comme augmentant le rendement et la vitesse de la synthèse du thréose, A129G, A129T et A129V. Le A129G est le plus efficace des trois puisqu'il est 20 fois plus rapide que la FSA_{wt} pour catalyser l'aldolisation entre deux molécules de GA.

Fort de ce constat et des travaux précédents réalisés sur la FSA A129S/A165G^[56], le mutant A129G/A165G a été créé et testé. Cependant, comme attendu en raison de la proximité spatiale de deux glycines, les interactions avec le donneur et l'accepteur diminuent ce qui entraîne une baisse de l'activité de la FSA. L'idée pour pallier à cette diminution d'activité a été de rajouter une troisième mutation proche du site actif pour rétablir les interactions avec le substrat donneur. Pour cela, le résidu 107 a été désigné comme potentiel candidat. Différentes combinaisons de mutations ont été testées et deux ont été retenues comme les plus actives : L107Y/A129G et L107Y/A129G/A165G. Ce gain

d'activité résulte d'une synergie entre les mutations L107Y et A129G. Le fait de retirer la chaîne de l'alanine 129 va permettre à la tyrosine en position 107 de mieux interagir avec les substrats donneurs. Grâce à ce travail, les auteurs ont proposé un mécanisme pour l'aldolisation de deux aldéhydes (dont le GA comme donneur) catalysée par les variants FSA en position 107 et 129 (Figure 26).

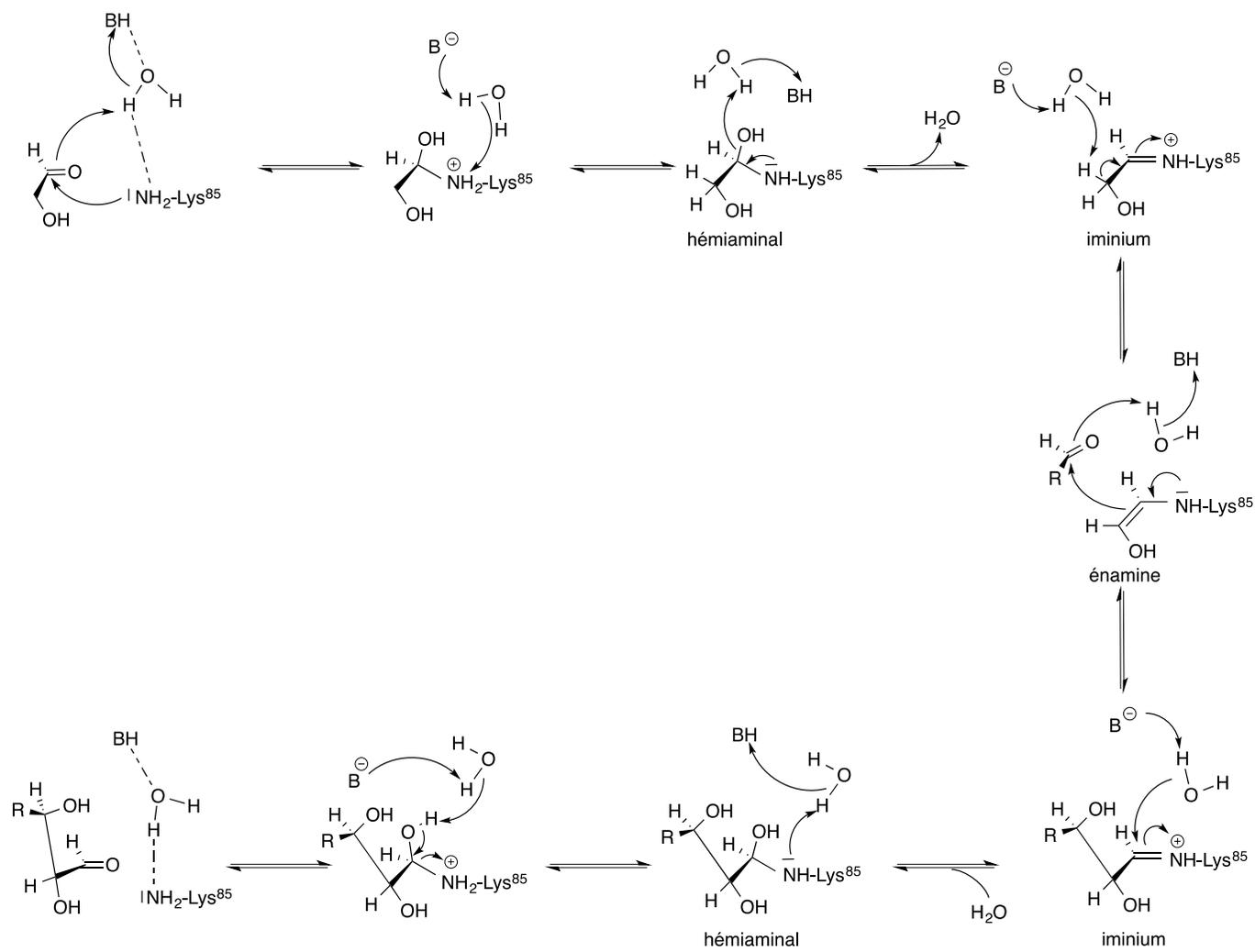


Figure 26 : possible mécanisme d'aldolisation catalysée par la FSA avec GA en donneur et un aldéhyde en accepteur^[54-56]

Dans ce mécanisme réactionnel les molécules d'eau ont une importance capitale (comme dans toutes réactions enzymatiques) puisqu'elles sont médiatrices de tous les transferts de protons durant la réaction. Le résidu basique, proposé comme intervenant dans le mécanisme par les auteurs, est l'acide aminé Y131. Cependant, la mutation L107Y provoque une interaction avec le groupement hydroxyl de l'hémiaminal formé par l'intermédiaire du motif phénol présent. Cette étude a permis de mettre en évidence l'importance des résidus 107 et 129 dans l'activité et la sélectivité vis-à-vis du donneur.

En 2015, la même équipe continue ses travaux sur l'utilisation de la FSA mutée pour la synthèse d'aldoses^[57] en incluant dans l'étude le formaldéhyde (FA) et en réalisant des réactions biocatalytiques en tandem. La synthèse de prébiotiques (pentoses et hexoses) stéréo-définis reste à l'heure actuelle un challenge qui peut être relevé par la biocatalyse. La majeure partie des méthodes pour la synthèse de polyols chiraux consiste à utiliser des substrats chiraux pour réaliser des réactions de biocatalyses ou de catalyses organométalliques. L'objectif ici est de préparer des molécules chirales fonctionnelles à partir de molécules achirales grâce à la biocatalyse et plus précisément la FSA dont nous avons précédemment détaillé le potentiel.

Grace à l'utilisation d'un programme de modélisation de structure, les auteurs ont créé une banque de FSA pour répondre à leur objectif. En plus des positions 107, 129 et 165 déjà étudiées, la mutation de la position 166 qui est une sérine dans la forme native de la FSA a été réalisée. Cette position 166 est très proche de l'accepteur dans le site actif ce qui peut laisser à penser que son remplacement par une glycine peut diminuer l'encombrement stérique. Six mutants ont été sélectionnés par la suite pour réaliser les synthèses. Douze substrats accepteurs ont été testés avec comme unique donneur le GA. Les résultats obtenus sont très intéressants puisque à partir de substrats achiraux et un seul mutant FSA des réactions en tandem sont réalisées avec des taux de conversion d'au moins 50 % (Figure 27).

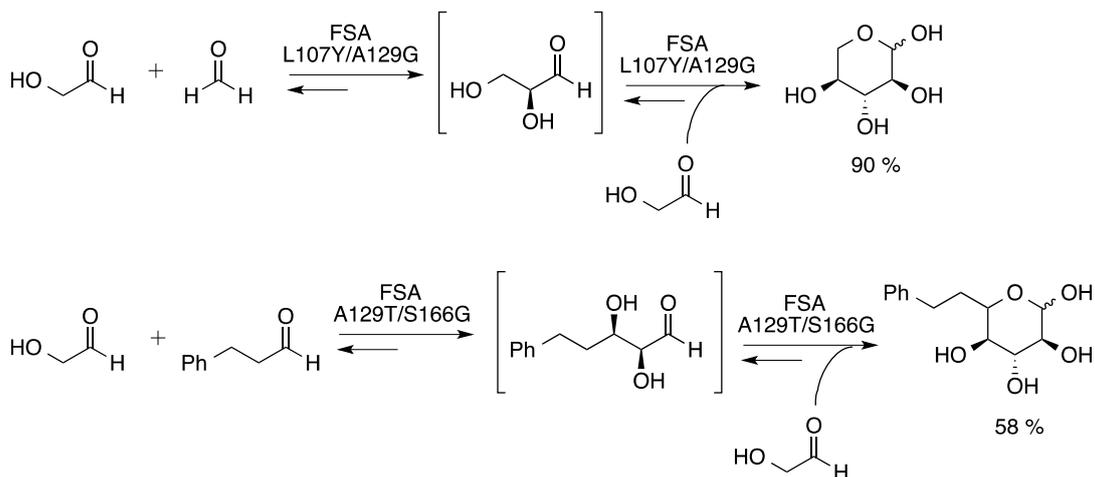


Figure 27 : double addition du glycolaldéhyde catalysée par deux variants de la FSA

Deux exemples ont été présentés parmi tous ceux figurant dans la publication de Szekrenyi *et al.*^[57] pour montrer qu'à partir d'un seul mutant FSA il est possible de réaliser en one-pot une réaction en tandem. Bien entendu, d'autres substrats accepteurs ont été testés pour observer l'influence de l'encombrement et du type de chaînes hydrocarbonées des accepteurs sur les taux de conversion en fonction des différentes mutations. Les auteurs ont réussi à créer à partir de molécules achirales, des dérivées du L-glucose. Cette publication montre l'importance qu'il faut accorder à l'ingénierie enzymatique pour la synthèse de molécules complexes à partir de molécules simples et ainsi relever les challenges évoqués dans l'introduction du manuscrit.

II. Valorisation du glycérol

1. Introduction

L'essor des biocarburants depuis plusieurs années entraîne une forte hausse de production du glycérol. En 2012, la France produisait plus de deux millions de tonnes de biodiesel pour environ 200 000 tonnes de glycérol^[3]. A l'échelle mondiale, la production de biodiesel est passée de 15 000 à 430 000 barils par jour, ce qui a multiplié par 9 la production de glycérol^[4]. Cette forte croissance a fait apparaître une manne financière importante liée à la valorisation du glycérol, ce qui a entraîné une augmentation des recherches sur ce sujet (Figure 28).

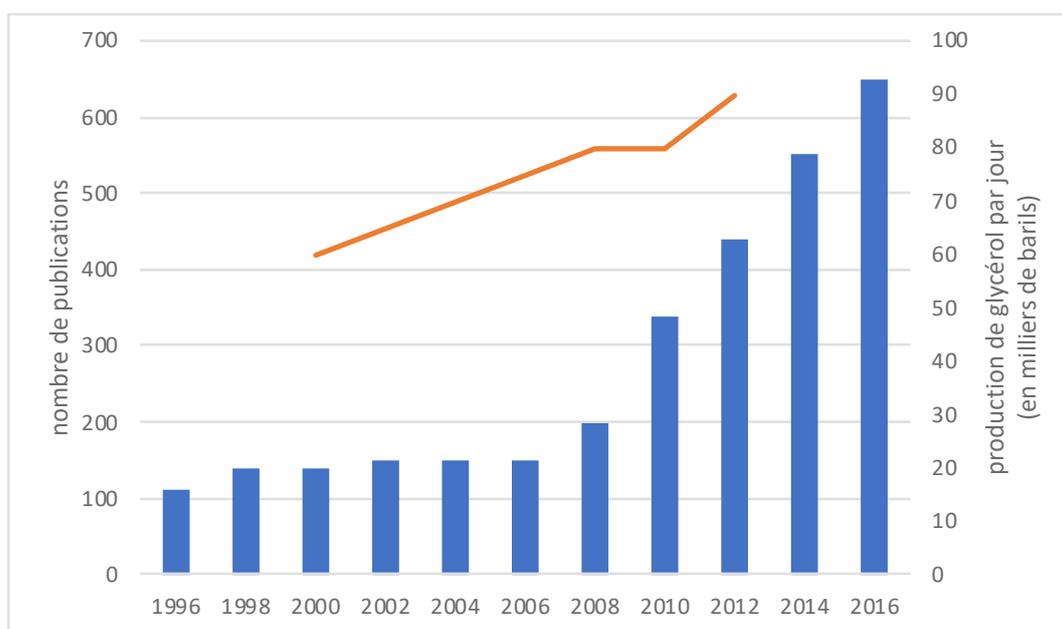


Figure 28 : évolution de la production (histogramme bleu) et du nombre de publications (courbe orange) concernant la valorisation du glycérol^[4,58]

De nombreuses approches de valorisation voient le jour : la catalyse organique, organométallique, enzymatique et l'utilisation de micro-ondes. Cette partie bibliographique a pour but de faire un état de l'art sur la valorisation du glycérol en distinguant les approches « chimiques traditionnelles » et les méthodes plus « innovantes » utilisant les microorganismes ou les enzymes isolées.

2. Valorisation par catalyse chimique

2.1. Généralités

Le glycérol est une molécule plateforme très simple puisqu'il ne possède que 3 fonctions alcools réactives. Une grande quantité de glycérol est produite chaque année puisque 10 kg de cette molécule sont produit pour 100 kg de biodiesel lors de la synthèse des biocarburants par transestérification des huiles végétales (Figure 29).

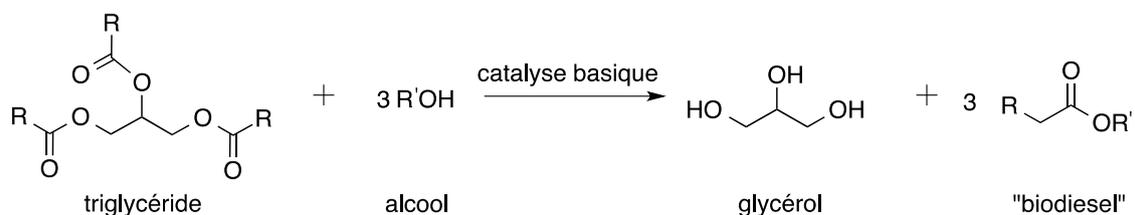


Figure 29 : synthèse par transestérification des huiles végétales des biocarburants

Les différentes molécules pouvant être synthétisées à partir du glycérol par catalyse chimique sont résumées Figure 30. Il est possible de mettre en œuvre un grand nombre de réactions fondamentales pour transformer le glycérol : oxydation, déshydratation, estérification...

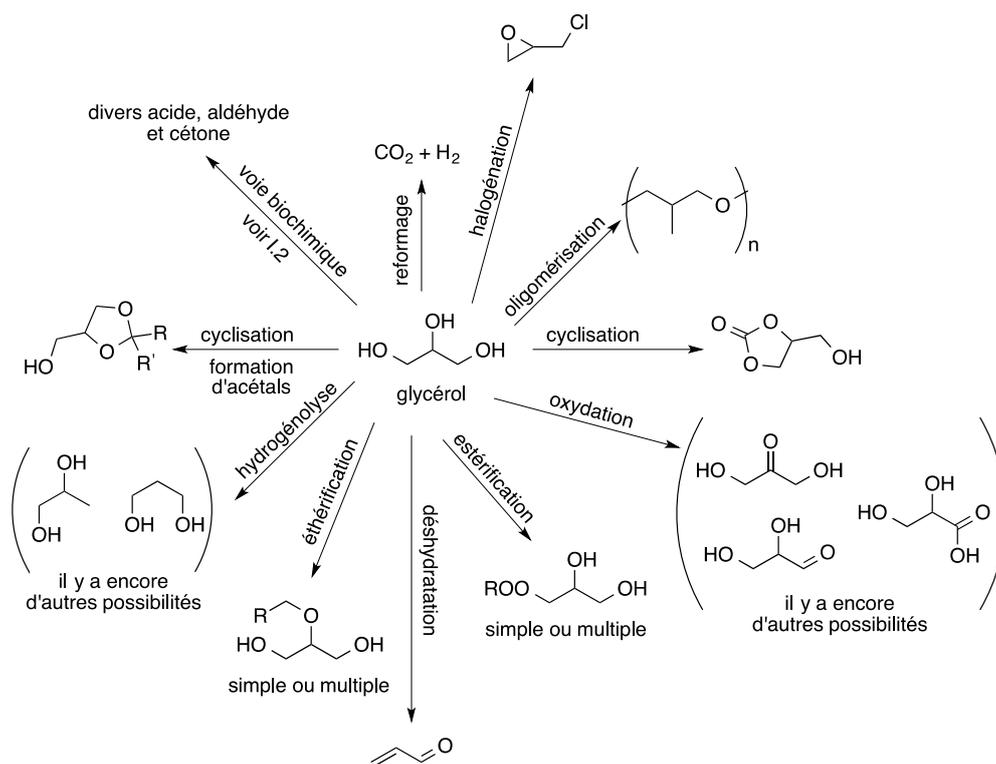


Figure 30 : les différentes voies de transformation du glycérol par catalyse chimique^[58]

Le glycérol peut mener à de nombreux produits très intéressants tels que des gaz, acides, aldéhydes...^[1,59]. La majorité des oxydations du glycérol s'effectuent à des pH basiques supérieurs à 9 et à des températures entre 50 à 100°C. Les autres réactions, telles que la déshydratation, la réduction et la formation de gaz, sont réalisées à des températures supérieures à 180°C. La plupart des conditions réactionnelles sont donc loin des principes de la « chimie verte » puisque soit les températures sont élevées, soit une concentration importante de base est nécessaire. A cela s'ajoute la présence de catalyseurs métalliques notamment pour les oxydations qui augmente encore le coût de production.

2.2. Reformage du glycérol

Beltram *et al.*^[60] ont mis au point une technique photocatalytique de valorisation du glycérol sous l'action de platine supporté par de l'oxyde de titane polymorphe préparé en présence d'urée pour la production d'hydrogène (Figure 31).

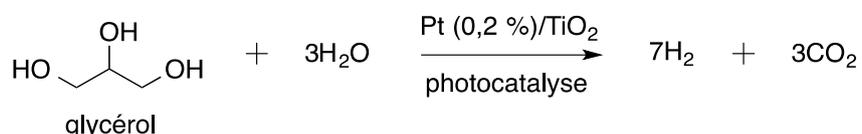


Figure 31 : reformage photocatalytique du glycérol sous action d'un complexe Pt 0,2 %/TiO₂ synthétisé en présence d'urée

L'étude a montré que la production de H₂ et CO₂ est dépendante du ratio molaire urée/Ti. C'est la conformation du support urée/Ti qui explique l'activité du catalyseur puisque à faible pourcentage en urée, le Ti n'est observé que sous sa forme rutile (0 à 3%), de 3 à 20 % le Ti est sous sa forme anatase et au-dessus de 20 % on observe la forme brookite. Ces trois formes diffèrent par l'organisation spatiale du Ti (mailles différentes^[61]). Il est important de ne pas dépasser 20 % car la forme brookite diminue la surface du catalyseur et donc son efficacité, la forme anatase est donc à privilégier.

L'équipe de Remón *et al.*^[62] a également travaillé sur le reformage du glycérol mais en utilisant des « catalyseurs classiques » (Pt, Ni, Cu...) sur des supports oxydés de type alumine. L'objectif était de déterminer les conditions de synthèse industrielle, notamment de pression, de température, le ratio masse de catalyseur/masse de glycérol et leur influence sur la conversion du glycérol en CO, CH₄, CO₂ et H₂.

2.3. Halogénéation du glycérol

La chloration du glycérol permet la production d'épichlorohydrines qui sont utilisées dans les résines plastiques ou encore dans les mousses polyuréthanes et les détergents. La synthèse consiste en une hydrochloration du glycérol catalysée par un acide dans un premier temps, puis en milieu basique lors d'une seconde étape pour la formation de l'époxyde (Figure 32)^[63].

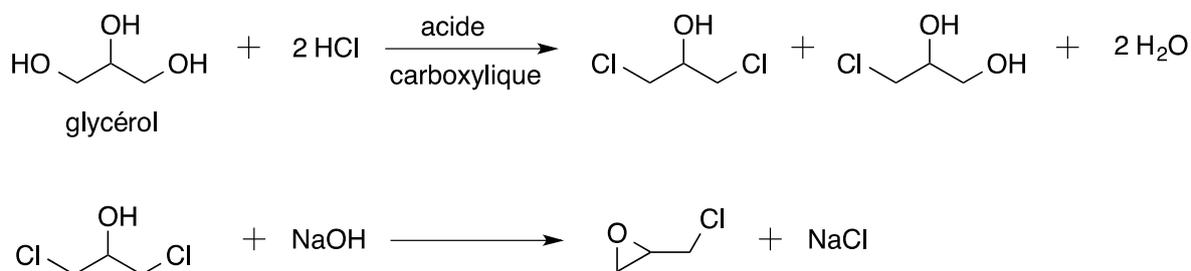


Figure 32 : synthèse d'épichlorohydrines à partir du glycérol

Un des groupements hydroxyles du glycérol est estérifié par l'acide carboxylique en milieu acide. La proximité du groupement alcool secondaire du glycérol va provoquer la cyclisation/déshydratation grâce à une catalyse acide. Enfin l'ouverture du cycle est provoquée par l'attaque nucléophile du chlore. Les conditions basiques de la deuxième étape entraînent la formation de l'époxyde grâce à un mécanisme de substitution nucléophile de type 2 (SN2) intramoléculaire.

2.4. Synthèse de carbonates de glycérol

Les carbonates de glycérol sont aussi des molécules très intéressantes puisqu'ils sont utilisés dans de nombreux domaines allant de la cosmétique aux batteries lithiums. Il existe beaucoup de méthodes pour obtenir ces molécules^[64]. L'une d'entre elles consiste en une réaction de transestérification du glycérol avec le diméthyl carbonate ou l'éthylène carbonate à des températures entre 50 et 70°C en utilisant pour catalyseur le méthoxide de potassium^[65] (Figure 33).

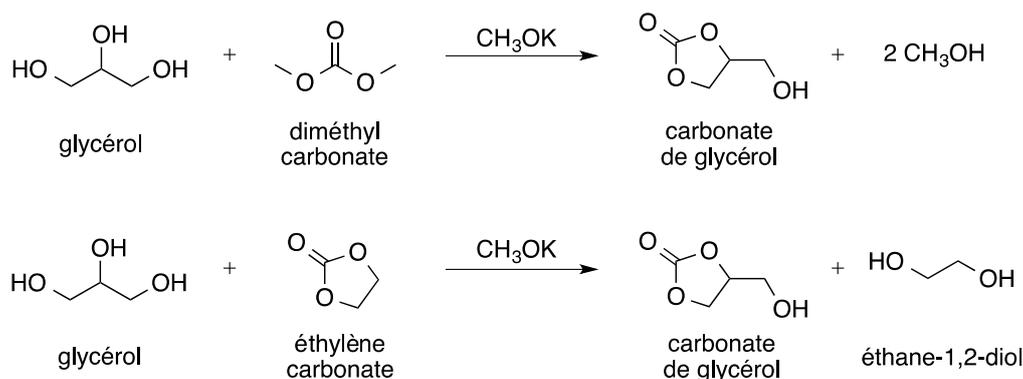


Figure 33 : synthèse de carbonates de glycérol

Truscello *et al.*^[66] ont travaillé sur la fonctionnalisation du glycérol en one-pot via la formation de carbonate de glycérol *in situ*, pour l'obtention d'aryloxypropanediols à activité biologique (Figure 34) comme la guaiphénésine, la méphénésine et la chlorophénésine.

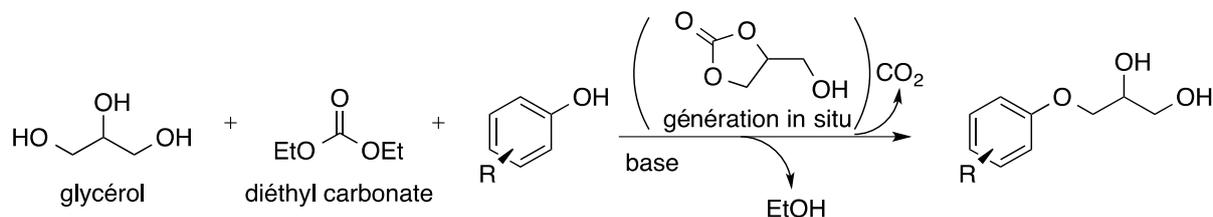


Figure 34 : synthèse d'aryloxypropanediols avec formation *in situ* de carbonates de glycérol, la molécule entre parenthèses correspond à l'espèce réactive mais n'a cependant pas été isolée dans cette publication

Cette méthode a permis d'obtenir des rendements de plus de 90 % tout en passant par des intermédiaires réactionnels et réactifs non toxiques.

2.5. Synthèse de polymères de glycérol

Le glycérol est une molécule intéressante à polymériser. En effet, comme le montre la Figure 30, il est possible d'obtenir de nombreux monomères différents pour arriver à la synthèse de polymères à activités variées^[67]. La première étude menée en 1994 par Tanton *et al.*^[68] consiste en la polymérisation d'un dérivé du glycérol, l'éther d'étoxyéthyl glycidyl pour la formation d'éther de poly-1,2-glycérol catalysée par de l'oxyde de césium (Figure 35).

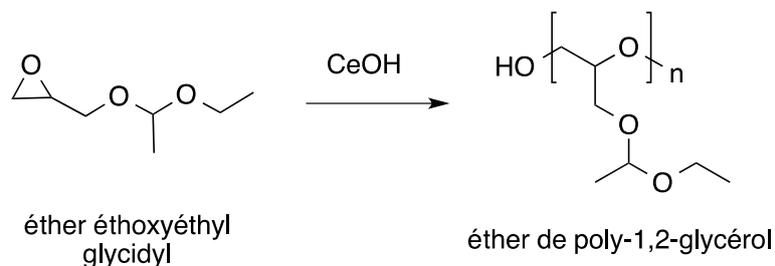


Figure 35 : polymérisation de l'éther éthoxyéthyl glycidyl catalysée par l'hydroxyde de césium

La synthèse de copolymère comportant des acides lactiques ou glycoliques est aussi très utile dans le domaine médical puisque ces polymères biodégradables sont présents dans les pansements ainsi que les filtres pour le sang^[67].

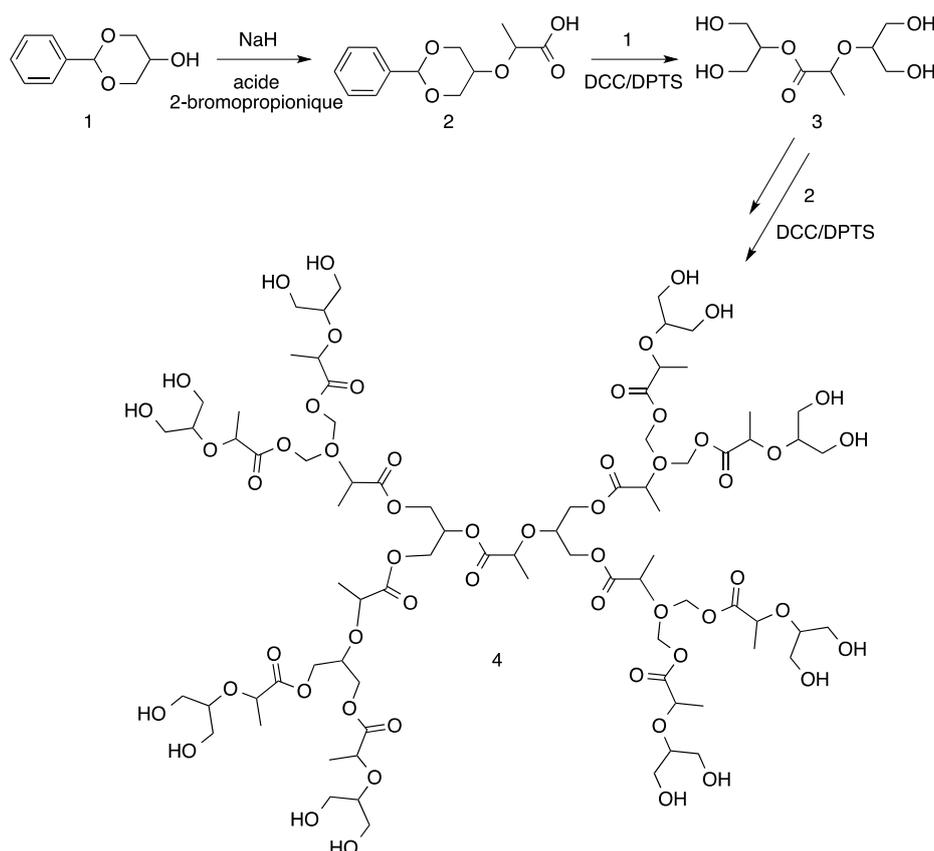


Figure 36 : synthèse de dendrimère de poly(glycérol-co-acide lactique)

La synthèse d'un polymère d'acide lactique a notamment été décrite par Carnahan *et al.*^[69] avec des masses moléculaires élevées entre 1200 et 4000. Le *cis*-1,3-*O*-benzylidèneglycérol (1) est traité par NaH puis par de l'acide 2-bromopropionique pour former l'acide 2-[(*cis*-1,3-*O*-benzylidèneglycérol)-2-propionique] (2). Les composés 1 et 2 sont ensuite couplés en présence de *N,N*-dicyclohexylcarbodiimide (DCC) et de 4-(diméthylamino)pyridinium 4-

toluènesulfonate (DPTS) pour créer un monomère (3). Ce monomère est ensuite couplé avec le composé 2 pour la synthèse du polymère final 4 (Figure 36).

2.6. Déshydratation du glycérol

La déshydratation du glycérol pour donner l'acroléine a beaucoup été étudiée ces dernières années. L'acroléine est une molécule très réactive de par son groupement aldéhyde et la présence d'une double liaison conjuguée ce qui en fait l'une des molécules plateformes les plus importantes pour l'industrie chimique. On retrouve cette molécule dans la synthèse de parfums, de polymères plastiques... Comme évoqué par Katryniok *et al.*^[5] le problème majeur pour réaliser cette déshydratation est le phénomène de coking sur le catalyseur (dépôt de résidus organiques désactivant) dû principalement aux conditions de températures élevées. Zhang *et al.*^[71] ont travaillé sur un catalyseur d'étain hydroxyphosphaté supporté par la zéolite HZSM-5 avec des rendements de 68 % au bout de 30 h en phase liquide. Le catalyseur possède de bonnes performances et une bonne résistance au phénomène de coking. Un traitement alcalin est effectué sur le support pour augmenter le volume des pores de la zéolite ce qui entraîne des transferts de molécules plus simples et explique la bonne résistance au coking. De nombreux oxydes de métaux ont aussi été étudiés pour la déshydratation du glycérol^[72]. Le meilleur oxyde recensé d'un point de vue conversion du glycérol (80 %) et sélectivité envers l'acroléine (75 %) est réalisé à partir d'oxyde de tungstène supporté sur une matrice poreuse de phosphate de zirconium en phase gaz^[73].

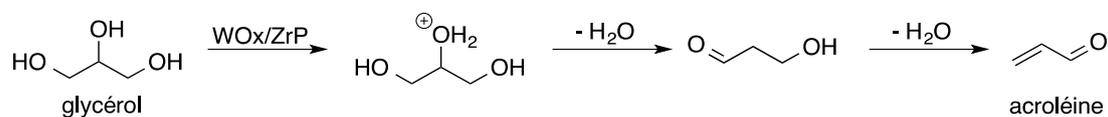


Figure 37 : déshydratation du glycérol en acroléine

Le passage du glycérol à l'acroléine se réalise en 3 étapes, une première étape de protonation catalysée par un acide de Brønsted (catalyseur décrit précédemment) et deux étapes de déshydratation (Figure 37).

2.7. Éthérification du glycérol

Les molécules résultantes de l'éthérification et de l'estérification du glycérol sont très utilisées dans l'industrie cosmétique en tant que surfactant et lubrifiant. De plus certaines molécules possèdent aussi des activités anti-inflammatoires, antibactériennes^[74]...

Gu *et al.*^[75] ont développé une approche par catalyse hétérogène en testant différents types de catalyseurs à base d'oxyde de silicium ou de résines (amberlite, zéolite...) pour catalyser l'éthérification du 1-phénylpropanol et du glycérol (Figure 38).

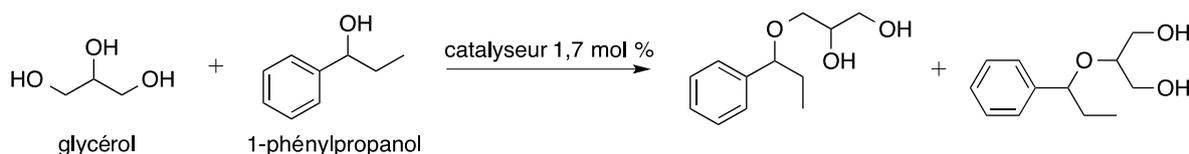


Figure 38 : éthérification du glycérol et du 1-phénylpropanol

La synthèse est sélective à plus de 99 %, seulement 1 % de double éthérification sur les deux alcools primaires du glycérol est observé. La double éthérification est augmentée lorsque le temps de réaction est allongé. Cependant, la sélectivité envers le composé voulu (éthérification de l'alcool primaire du glycérol) est fortement augmentée avec un large excès de glycérol. Le catalyseur SiO₂-SO₃H a donné les meilleurs résultats et a par la suite été utilisé pour la synthèse d'autres éthers en faisant varier la nature de l'alcool. Une chute du rendement est néanmoins observée pour les alcools allyliques à longue chaîne ne comportant pas d'insaturations.

2.8. Réduction du glycérol

La réduction du glycérol a été très étudiée ces dernières années pour donner notamment les 1,2- et 1,3-propandiol qui sont utilisés pour la synthèse de polyesters, polypropylène (Figure 39)^[76] .

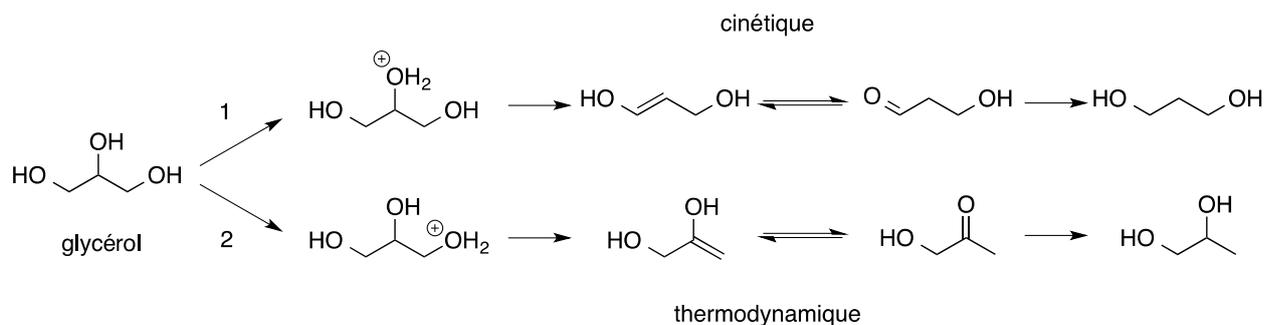


Figure 39 : voies de formation du 1,2- et 1,3-propanediol

La première déshydratation du glycérol pour former l'hydroxyacétone (2) est régie par la thermodynamique. Les auteurs ont donc extrapolé que la formation du 3-hydroxypropanal était contrôlée par les lois cinétiques. La formation d'un intermédiaire carbocation relativement stable grâce à une élimination de l'alcool en catalyse acide est le fondement de cette hypothèse. Ten Dam *et al.* ont montré que la voie 1 était contrôlée par la cinétique alors que la voie 2 était contrôlée par la thermodynamique. Ils ont montré que l'utilisation de Pt/Al₂O₃ avec de l'acide silicotungstique donne les meilleurs résultats de sélectivité du 1,3-propanediol. En effet la température ainsi que la quantité d'acide sont les principaux paramètres à prendre en compte pour favoriser la synthèse du 1,3-propanediol.

A contrario, Akiyama *et al.*^[77] se sont penchés sur la sélectivité envers le 1,2-propanediol. Ils ont mis au point une méthode à pression atmosphérique avec un gradient de température en utilisant un catalyseur au cuivre. En effet, à 190°C le glycérol est converti en 1,2-propanediol à 80 % « seulement ». L'étude de l'hydrogénation et de la déshydratation, réalisées séparément, ont montré que l'hydrogénation se fait à des températures moins élevées. A 200°C la déshydratation est complète c'est pourquoi un gradient de température est appliqué. Celui-ci commence à 200°C pour convertir le maximum de glycérol en HA puis la température est baissée jusqu'à 120°C (suffisant pour l'hydrogénation) pour limiter l'apparition de produits de dégradation et favoriser la synthèse du produit souhaité. Cette étude montre que la synthèse du 1,2-propanediol est thermodynamique.

2.9. Oxydation du glycérol par voie chimique

L'oxydation du glycérol a beaucoup été étudiée et notamment pour synthétiser le DHA très utilisée en cosmétique, l'acide tartronique pour une utilisation pharmaceutique

(obésité, ostéoporose...) mais aussi de nombreux autres acides à courte chaîne (synthons pour la synthèse organique) comme l'acide acétique, l'acide glycolique^[59] ... Une revue de 2011 fait état de toutes les molécules possiblement atteintes par oxydation chimique à partir du glycérol (Figure 40).

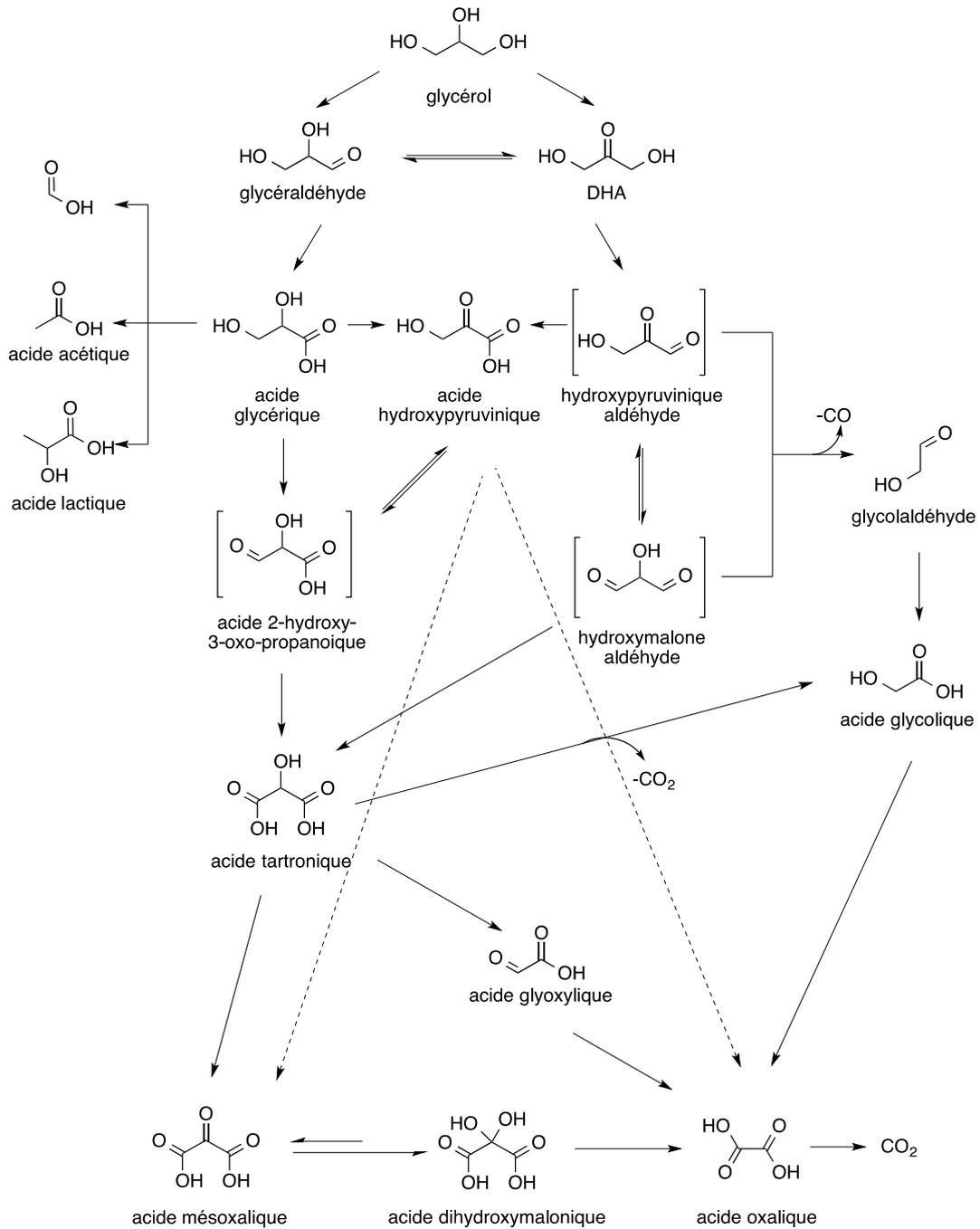


Figure 40 : produits obtenus par oxydation du glycérol^[59]

Le verrou majeur dans l'oxydation du glycérol est la sélectivité envers un seul produit et notamment lorsque le produit recherché est un aldéhyde.

Intéressons-nous plus particulièrement au DHA et donc à l'oxydation de l'alcool secondaire du glycérol. Painter *et al.*^[78] ont mis au point une méthode sélective d'oxydation pour obtenir le DHA avec des rendements de 90 % (Figure 41). Pour cela, ils utilisent un catalyseur métallique à base de palladium ainsi que la benzoquinone (agent oxydant). Le catalyseur au palladium couplé à la benzoquinone, dans des conditions de synthèse douce (température et pression atmosphérique) entraîne une très forte sélectivité pour le DHA (> 96 %).

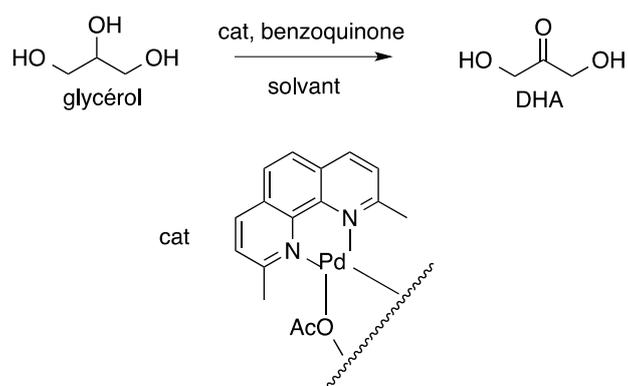


Figure 41 : oxydation sélective du glycérol en DHA

L'étude a aussi mis en évidence l'importance du solvant dans le taux de conversion puisque pour un mélange CH₃CN/H₂O (7/1) la totalité du glycérol est converti en 3 h alors que dans le DMSO seulement 15 min suffisent. Les propriétés aprotiques polaires du DMSO sont certainement la raison de cette vitesse de réaction.

De nombreux travaux ont aussi été réalisés sur l'oxydation sélective de l'un des alcools primaires du glycérol pour donner entre autre le glycéraldéhyde. Chornaja *et al.*^[79] ont réussi à oxyder plus ou moins sélectivement le glycérol en glycéraldéhyde en utilisant du platine supporté par différents oxydes (titane, alumine...). Cependant la plupart du temps la sélectivité n'est élevée que pour des taux de conversion très faibles en glycérol. Lorsque la concentration en glycéraldéhyde est trop élevée, il se dégrade ou s'oxyde d'avantage.

La synthèse d'acides a aussi beaucoup été étudiée. Cui *et al.* en 2014^[80] ont utilisé une catalyse à l'or sur différents supports pour oxyder le glycérol. Plus de 98 % de conversion sont obtenus pour chaque essai, cependant la nature du support a une importance capitale puisque, à mêmes conditions réactionnelles, les sélectivités sont totalement différentes.

Catalyseurs	Conversion en %	Rendements en %			
		Acide glycérique	Acide tartronique	Acide oxalique	Autres
Au/Ac	90	80	6	2	2
Au/CeO ₂	98	2	24	55	17
Au/NaY	98	26	43	4	25
Au/REY	99	5	69	6	19
Au/HY	98	3	80	12	3

Tableau 5 : oxydation du glycérol catalysée par l'or sur différents supports

La sélectivité n'est pas optimale. Cependant pour l'acide glycérique et pour l'acide tartronique elle atteint 80 % dans des conditions douces. Ces sélectivités sont attribuées à la taille des particules du catalyseur, pour l'acide tartronique les particules sont de très petites tailles (nm) alors que pour l'acide glycérique les particules sont quinze fois plus grandes.

2.10. Utilisation des micro-ondes

L'acrylonitrile est un produit très utilisé dans l'industrie chimique et plastique. A l'heure actuelle, il est synthétisé à partir de propylène qui est une ressource fossile. En 2009, Calvino-Casilda *et al.*^[81] ont réussi à synthétiser cette molécule à partir du glycérol en utilisant une méthode simple et rapide sans solvant, les micro-ondes. En effet l'utilisation des micro-ondes active le catalyseur organométallique à base d'antimoine et de vanadium supporté par de l'alumine (Figure 42).

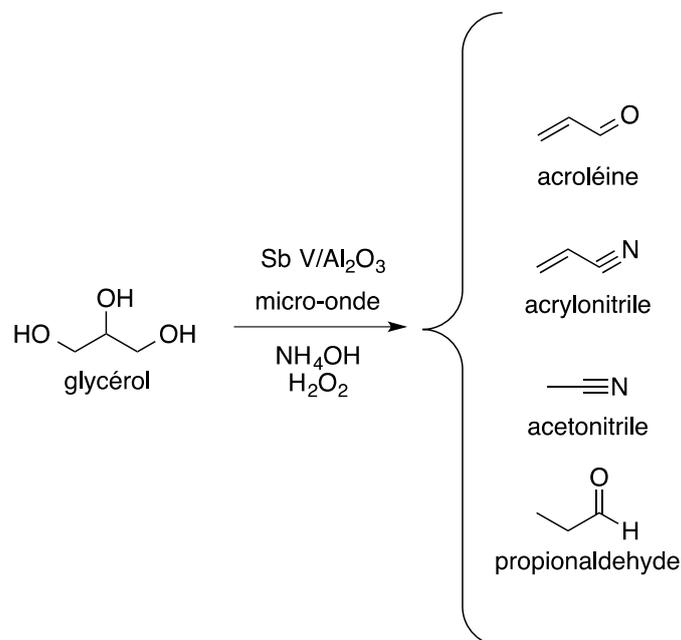


Figure 42 : Les quatre produits obtenus par action des micro-ondes sur le catalyseur à base de Sb et V sur le glycérol

Les différents « blancs » ont été réalisés et ont montré que c'est la combinaison des deux métaux et des micro-ondes qui donne 50 % de conversion du glycérol avec plus de 85 % de sélectivité en 1 h seulement. De plus, l'origine du Sb utilisé pour créer le catalyseur a son importance puisque le Sb-tartrate donne une meilleure sélectivité en acrylonitrile et une meilleure conversion que le Sb₂O₃ (conversion de 20% et sélectivité de 45 %). Lorsque le Sb-tartrate est utilisé, la formation de rutilé VSbO₄ est observée, ce qui favorise la formation de liaisons carbone-azote (transformation de l'acroléine en acrylonitrile). L'acroléine est donc un intermédiaire clé obtenu facilement grâce aux micro-ondes et transformé en acrylonitrile grâce au catalyseur sous structure rutilé. L'utilisation des micro-ondes est une méthode efficace, sans solvant et relativement peu onéreuse pour réaliser la synthèse de l'acrylonitrile à partir du glycérol. Cependant l'utilisation des micro-ondes à l'échelle industrielle paraît difficilement réalisable.

3. Valorisation du glycérol par voies biotechnologiques

3.1 Valorisation catalysée par des microorganismes complets

Les voies de synthèse biochimiques ont de nombreux avantages par rapport aux méthodes chimiques traditionnelles évoquées précédemment. Les conditions douces de réaction ainsi que la forte sélectivité en font un atout pour les industriels. En effet, entre

2004 et 2015, l'utilisation des microorganismes et enzymes a été multiplié par 5 dans les industries puisque les produits issues d'approches biotechnologiques sont passés de 3 % à plus de 15 %^[58]. Les voies biotechnologiques sont donc des approches d'avenir qui peuvent non pas remplacer la chimie mais compléter ou améliorer le panel de réactions possibles.

Le glycérol peut être métabolisé par de nombreuses bactéries pour synthétiser principalement différents acides, cétones, alcools et aldéhydes. Une fois le glycérol « absorbé » par la bactérie, il existe différentes voies métaboliques indépendantes qui utilisent des enzymes pour réaliser les synthèses. J'ai choisi de ne détailler qu'un exemple parmi la multitude de possibilités, celui de la synthèse de bioéthanol. En effet, le glycérol est principalement produit lors de la synthèse des biocarburants par transestérification d'huile végétale. Cependant, le glycérol peut aussi être transformé en biocarburant (bioéthanol) par l'intermédiaire de bactéries, telles que des mutants d'*E. coli*, *O. polymorpha*.

Cofré *et al.*^[82] ont mis au point les conditions de synthèse du bioéthanol et de l'hydrogène à partir du glycérol catalysée par une souche mutée de *E. coli* : MG1655. Ils ont déterminé les conditions réactionnelles optimales pour la fermentation et ont obtenu des conversions intéressantes puisque dans 200 litres, 100 % du glycérol étaient consommés et environ 9 g.L⁻¹ de bioéthanol étaient obtenus au bout de 250 h.

Le Tableau 6 présente un résumé des différentes molécules obtenues par catalyse microbienne à partir du glycérol.

Produits	Microorganismes	Applications	Références
1,3-Propanediol	<i>Klebsilla pneumoniae</i> , <i>Clostridia butyricum</i> , <i>Clostridium pasteurianum</i> , <i>Enterobacter agglomerans</i> , <i>Enterobacter gerogenes</i> , <i>Citrobacter freudij</i> , <i>Lactobacillus reuteri</i> ,	Synthèse de polyéthers, de polyuréthanes et de polyesters pour la cosmétique, l'alimentaire, la pharmacie et le textile. Production d'adhésifs, de solvants, de détergents et de résines.	[83–87]
Éthanol	<i>Citrobacter sp.</i> , <i>Klebsilla sp.</i> , <i>Enterobacter sp.</i> , <i>Eschérichia sp.</i> ,	Production de biodiesel, d'additifs pour diesel, industries pharmaceutiques, cosmétiques et chimiques	[88–92]
Acide lactique	Bactéries, champignons, levures,	Industries agroalimentaires, pharmaceutiques et cosmétiques. Synthons pour la synthèse organique.	[93–96]
Acide citrique	<i>Candida guilliermondii</i> , <i>Candidat lipolytica</i> , <i>Yarrowia lipolytica</i> , <i>Aspergillus niger</i> ,	Industries pharmaceutiques et chimiques, utilisation dans l'alimentation et les boissons.	[97,98]
Acide succinique	<i>Angerobiospirillum</i> , <i>succiniciproducens</i> , <i>Actinobacillus succinogenes</i>	Intermédiaire pour la synthèse de résines, polymères biodégradables, additifs et industries chimiques.	[99–101]
Acide propionique	<i>Propionibactérium sp.</i>	Herbicides, additifs, industries pharmaceutiques, cosmétiques, alimentaires et synthèse d'agents antifongiques	[102–105]
DHA	<i>Gluconobacter oxydans</i>	Cosmétique, médecine, industries pharmaceutiques et alimentaires et la chimie fine.	[106–111]

Tableau 6 : récapitulatif des microorganismes métabolisant le glycérol

3.2. Valorisation par enzymes isolées

3.2.1. Utilisation de lipases

Les premières transformations du glycérol étaient des réactions d'estérification catalysées par des lipases. La lipase Novozym 435 a été utilisée pour la synthèse d'acétates de glycérol par transestérification (Figure 43).

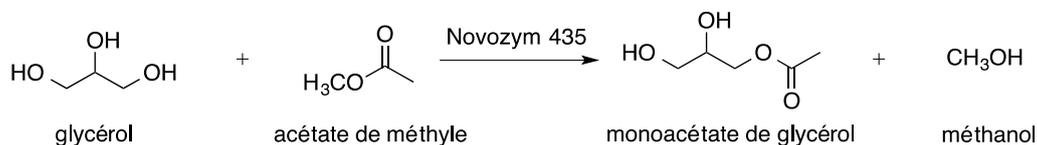


Figure 43 : transestérification du glycérol et de l'acétate de méthyle pour la synthèse d'acétate de glycérol

La double et la triple transestérification sont également observées. 95 % de conversion sont obtenus en 12 h dans les conditions optimales à 40°C dans le tert-butanol. Un ratio de 1/40 entre le glycérol et le méthyl acétate^[112] favorise la formation du monoacétate de glycérol.

Le glycérol 4-methoxy cinnamoyl utilisé dans la formulation cosmétique a aussi été synthétisé par la Novozym 435. Cette enzyme catalyse l'estérification de l'acide 4-méthoxy cinnamique par le glycérol avec un rendement de 34 %^[113]. Ces travaux mettent aussi en avant la « facilité » d'utilisation de cette lipase puisqu'elle tolère des températures de 70°C ainsi que différents solvants organiques. Des réactions d'oxydations du glycérol sont aussi étudiées.

3.2.2. Oxydation par voies enzymatiques

3.2.2.1 Les alcools oxydases

Les oxydoréductases actives sur les groupements CH-OH (EC 1.1.X) font partie de la famille des oxydases. La classification de ces oxydoréductases est divisée en sous-familles en fonction du groupement oxydé puis en fonction de l'accepteur (agent oxydant). Les C-OH oxydases sont classées selon le même principe (Tableau 7).

Code enzymatique	Appellation
EC 1.1.1	Avec NAD ⁺ ou NADP ⁺ comme accepteur
EC 1.1.2	Avec un cytochrome pour accepteur
EC 1.1.3	Avec l'oxygène comme accepteur
EC 1.1.4	Avec un disulfide comme accepteur
EC 1.1.5	Avec une quinone ou composé similaire comme accepteur
EC 1.1.9	Avec le cuivre comme accepteur
EC 1.1.98	Avec d'autres accepteurs physiologiques connus
EC 1.1.99	Avec des accepteurs dont le rôle physiologique est inconnu

Tableau 7 : classification par le Comité International de l'Union de la Biochimie et de la Biologie Moléculaire

Plus simplement, on peut aussi classer les alcools oxydases en quatre groupes en fonction uniquement de la nature du substrat oxydé^[114] : les alcools oxydases à courte chaîne (SCAO), les alcools oxydases à longue chaîne (LCAO), les alcools oxydases aromatiques (AAO) et les alcools oxydases d'alcool secondaire (SAO). Ces enzymes sont principalement retrouvées dans les bactéries, levures, champignons, plantes, insectes et mollusques. Chacun de ces organismes contient majoritairement l'une ou l'autre catégorie. Les SCAO et les LCAO sont de nature intracellulaire alors que les AAO et SAO sont principalement sécrétées dans le milieu. Les SCAO et LCAO sont toujours décrites comme des protéines multimériques avec des masses moléculaires très élevées, cependant les caractéristiques moléculaires de ces enzymes restent peu connues. On sait cependant que ces enzymes nécessitent quasiment toutes la présence d'un cofacteur lors de la réaction catalytique. L'oxydation directe des alcools en aldéhydes ou cétones correspondantes est catalysée principalement par deux enzymes, les alcools déshydrogénases (ADH) et les alcools oxydases (AOX). Les ADH catalysent la réaction d'oxydation mais celle-ci est réversible alors que les AOX ne peuvent pas catalyser la réduction. Ceci s'explique par la différence du cofacteur utilisé, les ADH utilisent le NAD⁺/NADH ou le NADP⁺/NADPH alors que les AOX utilisent le FAD et l'oxygène comme second substrat. D'autres alcools oxydases utilisent des cofacteurs métalliques comme le cuivre, des acides aminés ou bien aucun cofacteur.

3.2.2.2. Transformation des alcools primaires du glycérol et du glycolaldéhyde en aldéhydes par les oxydases

Durant cette thèse je me suis intéressé à la valorisation du glycérol par oxydation de l'alcool primaire. Cette réaction est très peu décrite et principalement dans des champignons. On dénombre plusieurs classes d'enzymes qui peuvent réaliser ces réactions :

- Glycérol déshydrogénase et oxydase
- Alcool oxydase
- Galactose oxydase

L'un des champignons les plus étudiés pour catalyser la réaction d'oxydation du glycérol en glycéraldéhyde est *Aspergillus japonicus*. En 1980, l'équipe de Uwajima^[115] met en évidence la présence de glycérol oxydase au sein d'une souche particulière de ce champignon (AT 008). 11 microorganismes ont révélé une activité glycérol oxydase (Tableau 8).

Souche	Activité en U par masse de cellule par mL de culture
<i>Aspergillus japonicus</i> AT 001	0,52
<i>Aspergillus japonicus</i> AT 008	1,12
<i>Aspergillus oryzae</i> AT 105	0,0028
<i>Aspergillus parasticus</i> AT 462	0,001
<i>Aspergillus flavus</i> AT 853	0,015
<i>Aspergillus tamarii</i> AT 857	0,024
<i>Aspergillus itaconicus</i> AT 923	0,0027
<i>Neurospora crassa</i> AT 003	0,025
<i>Neurospora sitophila</i> AT 045	0,0035
<i>Neurospora tetrasperma</i> AT 053	0,0040
<i>Penicillium</i> sp UT 1750	0,023

Tableau 8 : résultat du criblage des microorganismes pour la recherche d'activité glycérol oxydase

Comme nous pouvons le constater sur tous les microorganismes criblés, seuls les champignons présentent une activité glycérol oxydase. L'enzyme de *Aspergillus japonicus* AT 008 a été purifiée et testée avec différents cofacteurs. Fait très intéressant, seul l'oxygène est nécessaire à la réaction, aucun autre cofacteur ne donne d'activité. Cependant, aucune information concernant le couplage alcool/cofacteur n'est fournie. De plus, le FAD et le NADP n'ont pas été testés. Le produit d'oxydation a été déterminé comme un mélange des deux énantiomères du glycéraldéhyde par chromatographie sur couche mince, cependant aucune de leurs méthodes de détection ne permet d'affirmer la stéréochimie du produit obtenu.

Une alcool oxydase de *Phanerochaete chrysosporium* a été identifiée par Linke *et al.*^[116] comme ayant une activité glycérol oxydase. 36 souches de *Phanerochaete chrysosporium* ont été criblées par dosage du H₂O₂ et une alcool oxydase possédant une activité glycérol oxydase a été mise en évidence. Cette enzyme possède une activité nettement meilleure pour le méthanol (6 fois plus). Les constantes cinétiques déterminées avec le glycérol comme substrat montrent une forte affinité au travers du K_M de 1,9 mM et un V_{max} de 16,9 μmol.min⁻¹.L⁻¹ (k_{cat} non fourni). Le produit de la réaction a été déterminé comme étant du D-glycéraldéhyde grâce à une enzyme auxiliaire l'alcool déshydrogénase de levure de boulanger. Cette alcool oxydase est la première à avoir été décrite comme pouvant oxyder des alcools monovalents comme le méthanol et le glycérol.

Une glycérol déshydrogénase (GlyDH) provenant de *Gluconobacter oxydans* a révélé une activité d'oxydation du glycérol^[117]. Elle est capable d'oxyder le glycérol en un mélange racémique de glycéraldéhyde. Grâce à une seconde enzyme isolée du même microorganisme, la glycérol déshydrogénase 1615 (GlyDH 1615), le D-glycéraldéhyde produit est réduit en glycérol jusqu'à atteindre des excès énantiomériques >99% (Figure 44).

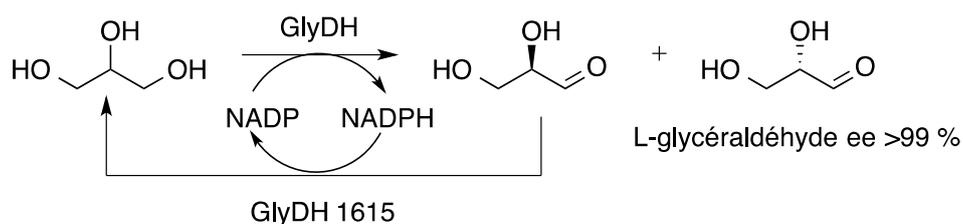


Figure 44 : production du L-glycéraldéhyde à partir de glycérol

Une galactose oxydase commerciale est aussi décrite comme pouvant oxyder le glycérol en L-glycéraldéhyde lors d'une cascade enzymatique^[118]. Cependant, le rendement de production n'est que de 5 % en 2 jours pour 3750 U (nombre de U correspondant à la conversion du D-galactose en D-galacto-hexodialdose). L'activité existe mais elle est insuffisante pour envisager l'utiliser en synthèse organique. Une solution serait peut-être de muter cette enzyme pour augmenter son activité.

Concernant l'oxydation du glycolaldéhyde en glyoxal, celle-ci est beaucoup moins décrite. La plupart du temps ce sont des microorganismes entiers (champignons ou levures) qui ont cette activité. Les rares fois où des enzymes isolées sont étudiées, les auteurs parlent d'oxydation de l'éthylène glycol en glyoxal en passant par le GA^[119] .

ISOBE *et al.* ont montré qu'une enzyme isolée de *Paenibacillus* sp AUI 311 est capable d'oxyder le glycolaldéhyde en glyoxal^[120]. Il est intéressant de constater que cette enzyme oxyde préférentiellement les alcools primaires des molécules possédant un aldéhyde puisque le glycéraldéhyde est aussi un substrat.

Chapitre 2 :

Résultats et

discussion

L'objectif de mon travail de thèse est de réaliser à partir du glycérol la synthèse du glyoxal par cascade enzymatique (Figure 45). Cependant, il ne faut pas perdre de vue l'aspect technologique du projet GLYCYBRIDE qui vise à combiner la catalyse chimique et enzymatique dans un même réacteur.

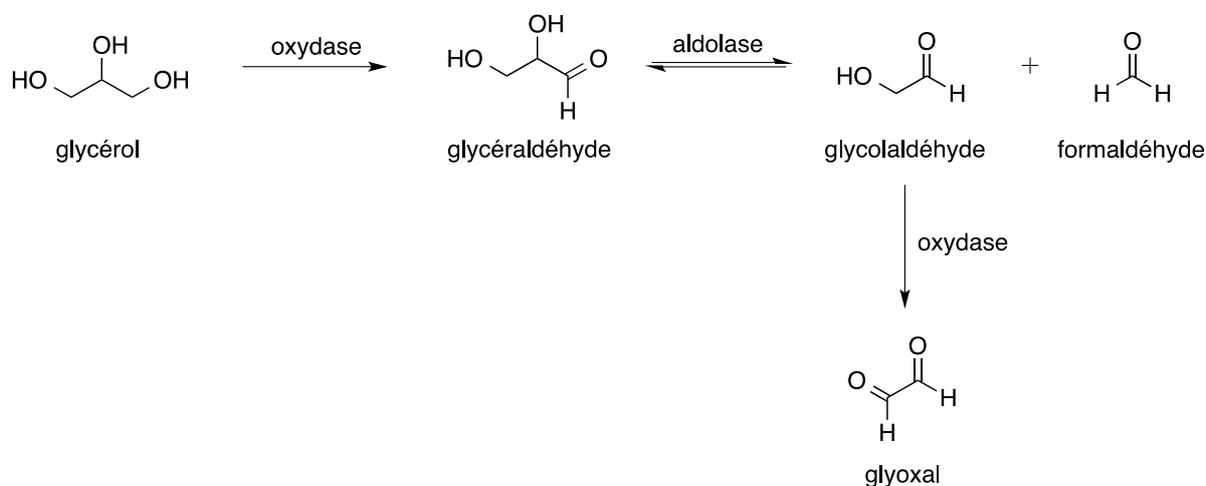


Figure 45 : synthèse du glyoxal par catalyse enzymatique

Le glyoxal est une cible intéressante, c'est une molécule plateforme très utilisée en synthèse chimique en raison de ses propriétés:

- Biodégradable
- Agent de solubilisation et de réticulation dans les industries textiles, papetières et du cuir
- Substitut du formaldéhyde et du glutaraldéhyde
- Substance accréditée pour les emballages alimentaires

I. Étude du Clivage du glycéraldéhyde

1. Objectif

La première année de ma thèse s'est focalisée sur l'étape de rétro-aldolisation du glycéraldéhyde en glycolaldéhyde et formaldéhyde. Nous avons sélectionné la fructose-6-phosphate aldolase (FSA) décrite dans la partie bibliographique pour réaliser cette étape en raison de sa capacité à effectuer des réactions d'aldolisation entre deux aldéhydes. En partenariat avec l'Institut de Chimie de Clermont-Ferrand (ICCF), nous avons mis au point un test de criblage et criblé leur collection de FSA. Le clivage du L-glycéraldéhyde a été mis en

évidence sous l'action d'un double mutant, la FSA L107Y A129G. Cependant, nous avons constaté une perte d'activité de cette enzyme en présence du D-glycéraldéhyde. Cette inhibition est un problème pour deux raisons. La première est bien entendue pour des questions de rendement et de production. La seconde concerne le couplage avec la catalyse chimique puisque les études menées à l'UCCS ont montré qu'il n'est pas possible dans leurs conditions de produire spécifiquement le L-glycéraldéhyde. En effet, le catalyseur mis au point ne permet l'obtention que du mélange racémique du glycéraldéhyde. Nous avons donc choisi d'explorer deux approches pour pallier cette inhibition :

- Utiliser les méthodes de docking pour étudier une troisième mutation sur la FSA L107Y A129G
- Réaliser la synthèse du L-glycéraldéhyde par voie enzymatique à partir du glycérol

2. Production et criblage d'une banque de mutants de FSA pour cliver le glycéraldéhyde

La FSA_{wt} est l'une des deux seules enzymes, avec la déoxyribose-5-phosphate aldolase^[121], capable de réaliser des réactions d'aldolisation entre deux aldéhydes. Cette enzyme a la particularité d'être thermostable ce qui est un avantage pour un éventuel couplage avec la catalyse chimique. Notre laboratoire possède une souche d'*Escherichia coli* (*E.coli*), fournie par l'ICCF et productrice de la FSA_{wt}. C'est donc naturellement que nous avons défini cette enzyme comme un bon candidat pour cliver le glycéraldéhyde. Au sein de l'ICCF, l'équipe du Professeure Lemaire est spécialisée dans l'utilisation de la FSA comme le montrent les nombreuses références citées dans la partie bibliographique et possède des mutants de la FSA. Nous avons donc engagé une collaboration de deux mois pour produire et cribler ces mutants dans leurs locaux à Clermont-Ferrand.

Nous avons mis au point un test de criblage utilisant une enzyme auxiliaire, l'alcool déshydrogénase (ADH) de *Saccharomyces cerevisiae*. Cette enzyme est capable de réduire le glycolaldéhyde produit par la rétro-aldolisation, potentiellement catalysée par la FSA, en utilisant un cofacteur, le NADH, dont la disparition est suivie à 340 nm (Figure 46).

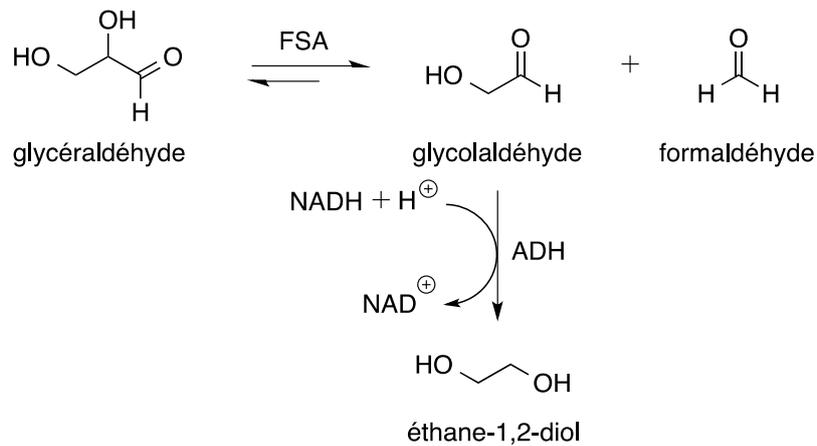


Figure 46 : méthode de criblage de la FSA en plaque 96 puits de 200 μ L avec une concentration de glycéraldéhyde de 50 mM, 5 U d'ADH, 1,2 mg de FSA non purifiée, 0,4 mM de NADH et pour solvant du tampon Glygly 50 mM pH 8

Les blancs ont été effectués en présence de chacun des substrats pour s'assurer de la spécificité de substrat de l'enzyme auxiliaire. Une dérive de 0,006 DO/min en présence de glycéraldéhyde est à noter mais négligeable par rapport à l'activité de réduction du glycolaldéhyde de l'ADH. La glycérol déshydrogénase (GDH) a été également testée comme enzyme auxiliaire mais elle est trois fois plus chère et moins sensible que l'ADH.

Dix-sept mutants de la FSA ainsi que quelques FSA issues de la biodiversité ont été produits et purifiés par résine IMAC en utilisant le protocole standard détaillé dans la partie matériel et méthode. Les enzymes ont été criblées grâce au test décrit précédemment en utilisant une quantité d'enzyme équivalente en masse (1,2 mg d'enzyme purifiée par puits de 200 μ L) ainsi que des concentrations en glycéraldéhyde de 25 et 50 mM. Aucune activité n'a été détectée en présence de D-glycéraldéhyde ou du mélange racémique ce qui laisse à penser la présence d'une éventuelle inhibition par le D-glycéraldéhyde ou un autre produit formé. Cependant des activités ont été révélées en présence de L-glycéraldéhyde (Tableau 9).

FSA	Résultat du criblage (DO/min)
Blanc	0,006 (+/- 0,002)
FSA _{wt}	0,007 (+/- 0,002)
A8AZ46	0,021 (+/- 0,004)
Q8DVJ4	0,024 (+/- 0,004)
Q99XT4	0,020 (+/- 0,003)
Q4HEJ1	0,012 (+/- 0,002)
A129G	0,017 (+/- 0,002)
A129S	0,013 (+/- 0,004)
A165S	0,024 (+/- 0,002)
A165G	0,012 (+/- 0,001)
L107Q	0,015 (+/- 0,002)
L107Y	0,010 (+/- 0,004)
L107Y A129G	0,078 (+/- 0,004)
A129S L107Y	0,013 (+/- 0,005)
A129G A165G L107Y	0,026 (+/- 0,002)
A129S A165S L107Q	0,012 (+/- 0,008)
A129S L107Q	0,016 (+/- 0,003)
R134A	0,024 (+/- 0,004)
D6H	0,010 (+/- 0,001)

Tableau 9 : résultat du criblage de la banque de FSA réalisé en plaque 96 puits dans 200 µL de tampon Gly-gly 50 mM pH 8 en présence de 50 mM de L-glycéraldéhyde et de 1,2 mg d'enzyme.

Les résultats ne diffèrent pas entre les deux concentrations de L-glycéraldéhyde testées pour vérifier que la concentration en glycéraldéhyde n'inhibe pas les FSA criblées. Cinq FSA testées ont répondu positivement au crible, seule la FSA contenant les mutations L107Y

A129G a été retenue pour les étapes suivantes car les autres mutants ne montraient pas une activité suffisamment élevée et nous suspicions des faux.

3. Caractérisation du mutant FSA L107Y A129G

Les constantes cinétiques de la FSA L107Y A129G (Tableau 10) ont été déterminées en utilisant une enzyme auxiliaire l'alcool déshydrogénase de *Saccharomyces cerevisiae* (Figure 46) et la représentation graphique de Hanes-woolf, plus précise que celle de Lineweaver-Burk pour des valeurs très basses concentrations.

Substrat	K_M	k_{cat}/K_M	V_M
L-glycéraldéhyde	2,7 mM (+/- 0,11)	0,104 min ⁻¹ .mM ⁻¹ (+/- 0,004)	0,013 μM.min ⁻¹ (+/- 0,001)

Tableau 10 : détermination des constantes cinétiques de la FSA L107Y A129G, réalisée en triplicata

La valeur de k_{cat}/K_M est très faible si l'on compare avec le k_{cat}/K_M de la réaction de clivage du fructose-6-phosphate de la FSA_{wt} qui est de 180 min⁻¹.mM⁻¹.

La thermostabilité de ce mutant a été étudiée en incubant une solution de FSA à 60°C pendant plusieurs jours. Aucune perte d'activité n'a été remarquée sur une durée de 50 h (Figure 47)^[42]. Par conséquent, les deux mutations en positions 107 et 129 n'ont pas d'effets sur la thermostabilité.

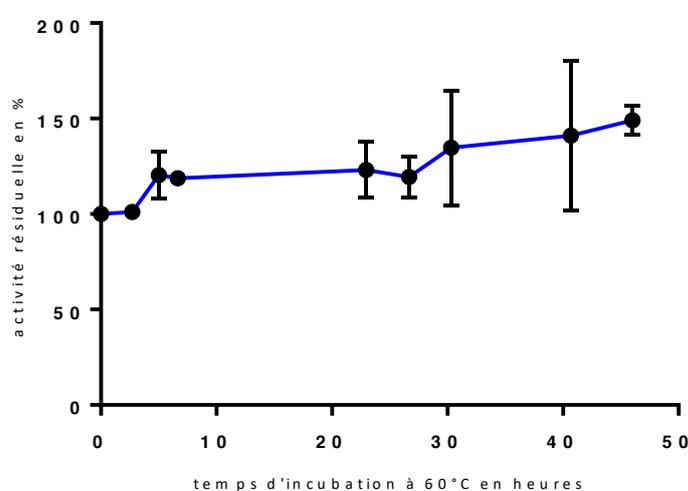


Figure 47 : thermostabilité de la FSA L107Y A129G incubée à 60°C pendant 50 heures, réalisée en plaque 96 puits de 200 μL avec 1,2 mg de FSA par puits dans du tampon Glygly 50mM pH 8 et une concentration en L-glycéraldéhyde de 25 mM

L'augmentation d'activité que l'on observe veut dire que en plus d'être thermostable, la FSA voit son activité augmentée avec la température ce a déjà été décrit notamment sur la Transcétolase de *Geobacillus stearothermophilus*^[8].

Le produit de clivage du L-glycéraldéhyde par la FSA L107Y A129G a ensuite été caractérisé par RMN ¹H (Figure 48).

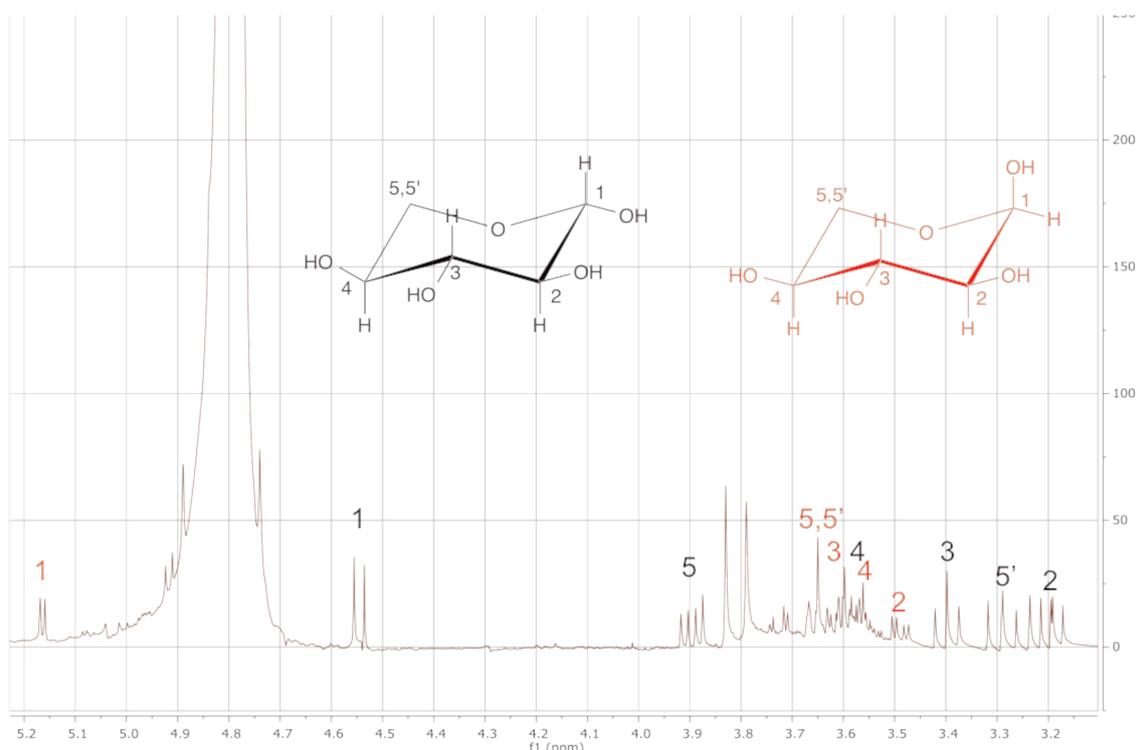


Figure 48 : Interprétation du spectre RMN du proton avec en noir le α -L-xylose et en rouge le β -L-xylose

Les signaux RMN du glycolaldéhyde ne sont pas visibles. Ce résultat était prévisible puisque la FSA catalyse préférentiellement l'aldolisation. Le L-glycéraldéhyde est bien clivé puisque les signaux correspondant ont disparus, une réaction secondaire s'est produite. Le L-glycéraldéhyde est bien clivé, cependant le glycolaldéhyde produit est ensuite substrat de la FSA L107Y A129G qui va catalyser la réaction d'aldolisation sur le L-glycéraldéhyde pour donner le L-xylose (Figure 49). La réaction d'aldolisation catalysée par la FSA L107Y A129G du glycolaldéhyde sur L-glycéraldéhyde pour former le L-xylose est décrite par Clapes *et al.*^[56]. La caractérisation RMN du proton du L-xylose a été effectuée dans cet article et nous a permis de comparer nos résultats. Aucune trace de glycolaldéhyde n'a été observée par RMN, il se retrouve dans l'eau sous sa forme hydratée. Son signal RMN est masqué par les signaux RMN des protons de l'eau (4,79 ppm). Au bout de 24 h, plus de 95 % de glycéraldéhyde sont consommés (estimation RMN du proton avec TSPD4 en standard).

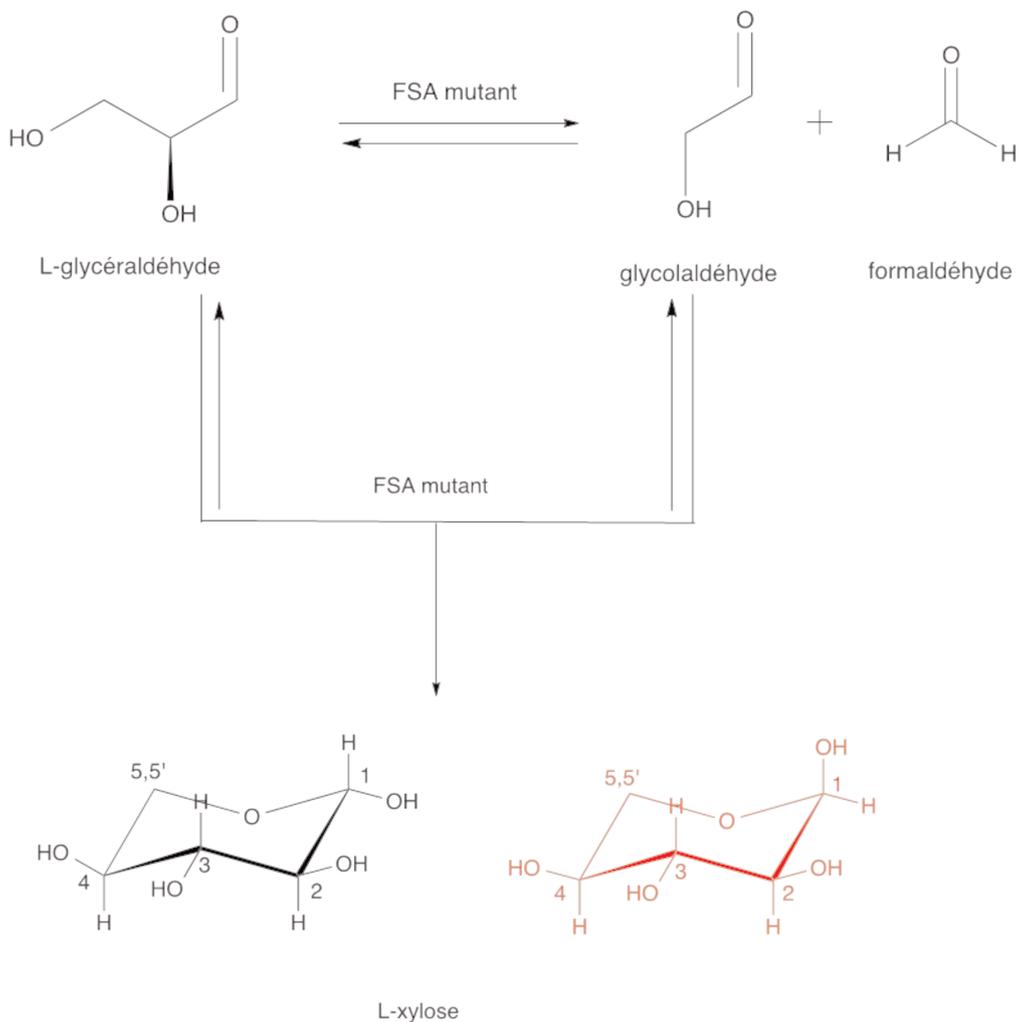


Figure 49 : réaction secondaire entre le glycolaldéhyde produit et le L-glycéraldéhyde catalysée par la FSA L107Y A129G

La présence de cette réaction « parasite » va devoir être maîtrisée par la suite car l'objectif est d'oxyder le glycolaldéhyde obtenu en glyoxal et non de produire du xylose. Cependant, les oxydases catalysent des réactions très souvent irréversibles. Si l'oxydase catalyse sélectivement l'oxydation du glycolaldéhyde en glyoxal, la réaction sera donc naturellement déplacée dans le sens de la synthèse du glyoxal.

4. Production du mutant FSA L107Y A129G

Une fois l'enzyme caractérisée, nous avons créé le mutant par biologie moléculaire puisque l'ICCF n'était pas en mesure de nous fournir la souche productrice de ce mutant. Une stratégie de mutagenèse dirigée a donc été adoptée, à partir du gène codant pour la FSA_{wt} que nous possédons, dans un plasmide avec un gène de résistance à l'ampicilline et un promoteur d'induction de production à l'IPTG.

Notre stratégie repose sur l'utilisation de protéines recombinantes, qui est de plus en plus présente, pour réaliser des synthèses enzymatiques. Cette méthode permet de produire de grandes quantités d'enzymes grâce à une induction par un promoteur fort (Figure 50) et aussi d'utiliser une plus grande diversité d'enzymes car le nombre d'enzymes commerciales est assez limité.

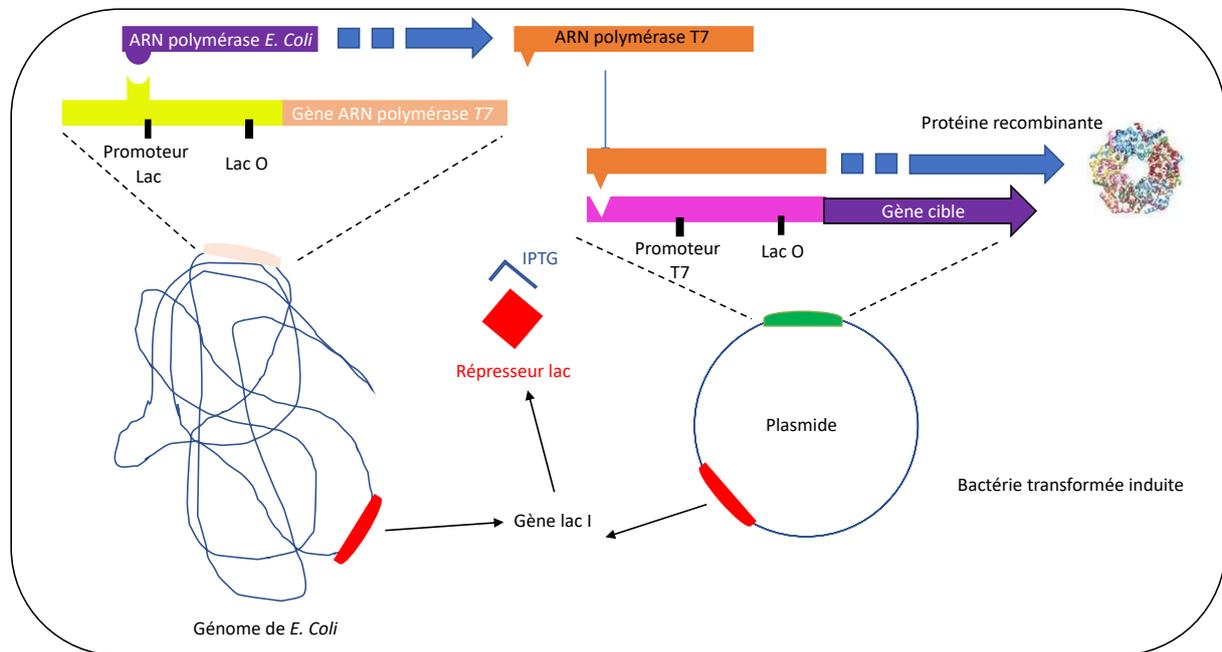


Figure 50 : production de protéines recombinantes chez *E. coli*

La Figure 50 illustre le mode de production des protéines recombinantes dans *E. coli* (BL21 DE3 pLys, Rosetta 2 (DE3) et XL1 BLUE utilisées lors de ce travail). En l'absence de l'inducteur (IPTG), le gène Lac I va initier la synthèse du répresseur de l'opéron lactose qui va se fixer sur l'opérateur Lac O. Ceci va empêcher la transcription et la traduction du gène qui code l'ARN polymérase T7. La présence d'IPTG va inhiber le répresseur par complexation ce qui va supprimer sa fixation au niveau de Lac O. La transcription et la traduction de l'ARN polymérase peut donc avoir lieu et permet alors la transcription et la traduction de la protéine recombinante. C'est par cette technique que notre mutant de la FSA et nos oxydases ont été produites.

Les deux mutations ont été réalisées en série (l'une après l'autre) et non en une seule fois de manière à enrichir la banque d'enzymes du laboratoire. Pour réaliser ces mutations, le kit Agilent QuickChange II Site-Directed Mutagenesis (KAQ) a été utilisé. La plupart des techniques utilisent de l'ADN simple brin comme point de départ contrairement

à ce kit qui permet de réaliser des mutations spécifiques sur des plasmides doubles brins ce qui est plus simple. De plus, KAQ peut être utilisé sur n'importe quel vecteur et ne nécessite pas d'étape *in vitro* de méthylation. Cette méthode est simple, en trois étapes, à partir d'un plasmide parent purifié comme décrit dans la partie matériel et méthodes (Figure 51).

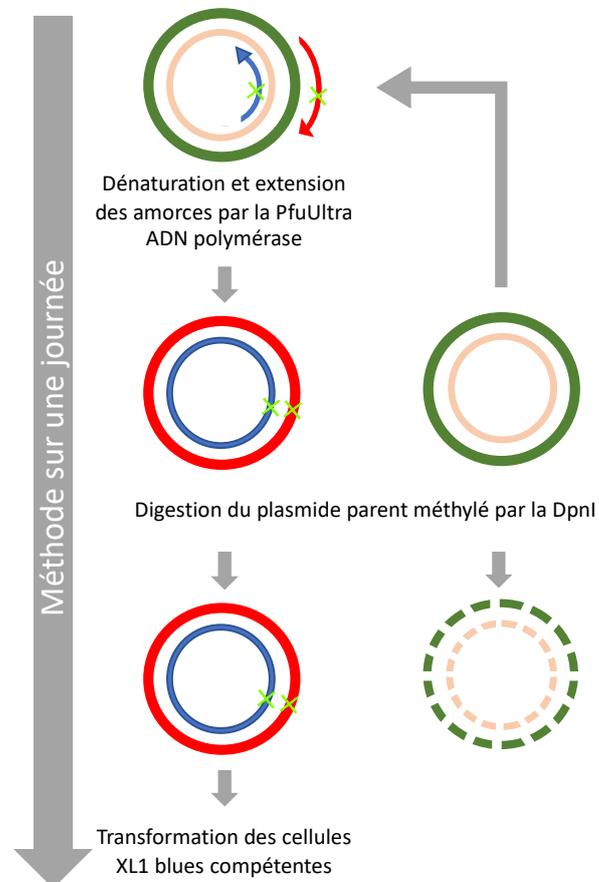


Figure 51 : procédure pour réaliser la mutation à l'aide du KAQ avec en rouge et bleu les amorces qui donneront le plasmide contenant la mutation symbolisée par une croix verte claire, en vert et beige le plasmide parent

Dans un premier temps, la mutation en position 107 pour muter la leucine en tyrosine a été réalisée, puis, dans un second temps, la muter de l'alanine 129 en glycine. Les amorces « forward » et « revers » contenant les mutations ont été construites grâce au logiciel présent sur le site www.agilent.com (Figure 52 et Figure 53).

10 20 30 40 50

.....|.....|.....|.....|.....|

ATGGAACGTGATCTGGATACTTCAGACGTTGTTGCGGTGAAGGCGCTGTC
ACGTATTTTTCCGCTGGCGGGTGTGACCACTAACCCAAGCATTATCGCCG
CGGGTAAAAAACCGCTGGATGTTGTGCTTCCGCAACTTCATGAAGCGATG
GGCGGTCAGGGGCGTCTGTTTGCCAGGTAATGGCTACCACTGCCGAAGG
GATGGTTAATGACGCGCTTAAGCTGCGTCTATTATTGCGGATATCGTGG
TGAAAGTTCGGTGACCGCCGAGGGGCTGGCAGCTATTAAGATGTTAAAA
GCGGAAGGGATTCCGACG CTG GGAACCGCGGTATATGGCGCAGCACAA
CGGAAGGGATTCCGACG TAT GGAACCGCGGTATATGG
GGGCTGCTGTCGGCGCTGGCAGGTGCGGAATATGT TGC GCCTTACGTT
AGGTGCGGAATATGT TGG GCCTTACGTT
AATCGTATTGATGCTCAGGGCGGTAGCGGCATTCAGACTGTGACCGACTT
AATCGTA
ACACCAGTTATTGAAAATGCATGCGCCGCAGGCGAAAAGTGCTGGCAGCGA
GTTTCAAACCCCGCTCAGGCGCTGGACTGCTTACTGGCAGGATGTGAA
TCAATTACTCTGCCACTGGATGTGGCACAACAGATGATTAGCTATCCGGC
GGTTGATGCCGCTGTGGCGAAGTTTGAGCAGGACTGGCAGGGAGCGTTTG
GCAGAACGTCGATTTAA

Figure 52 : séquences génétiques de la FSA_{wt} en noir et les amorces contenant les mutations positions 107Y (vert) et 129G (bleu)

10 20 30 40 50

.....|.....|.....|.....|.....|

FSA wild type	MELYLDTS	SDVVAVKALSRI	FPLAGVTTNPSII	IAAGKKPLDVVLPQLHEAM	
FSA L107Y A129G	MELYLDTS	SDVVAVKALSRI	FPLAGVTTNPSII	IAAGKKPLDVVLPQLHEAM	
FSA wild type	GGQGR	LFAQVMATTAEGMVNDALKLRSII	IADIVVKVPVTAEGLAAIKMLK		
FSA L107Y A129G	GGQGR	LFAQVMATTAEGMVNDALKLRSII	IADIVVKVPVTAEGLAAIKMLK		
FSA wild type	AEGIPT	L	GTAVYGAAQGLLSALAGA EYV	A	PYVNRIDAQGGSGIQTV
FSA L107Y A129G	AEGIPT	Y	GTAVYGAAQGLLSALAGA EYV	G	PYVNRIDAQGGSGIQTV
FSA wild type	TDLHQ	LLKMHAPQAKVLAASF	KTPRQALDCLLAGCESITLPLDVAQQMIS		
FSA L107Y A129G	TDLHQ	LLKMHAPQAKVLAASF	KTPRQALDCLLAGCESITLPLDVAQQMIS		
FSA wild type	YPAVDA	AAVAKFEQDWQGA	FGRTSI		
FSA L107Y A129G	YPAVDA	AAVAKFEQDWQGA	FGRTSI		

Figure 53 : séquences peptidiques de la FSA_{wt} et de la FSA L107Y A129G

Une fois ce double mutant créé, le problème de l'inhibition en présence du D-glycéraldéhyde s'est posé. En effet, il n'est pas possible de synthétiser préférentiellement l'un des deux énantiomères par oxydation chimique du glycérol dans les conditions testées par l'UCCS (Unité de Catalyse et Chimie du Solide). De plus, lors de l'ajout du D-glycéraldéhyde, même à très faible concentration (1 mM), dans le milieu réactionnel, aucune

activité n'est observée, même en multipliant par deux la quantité d'enzyme ou le temps de réaction.

Pour pallier cette inhibition, deux stratégies ont été menées :

- muter la FSA L107Y A129G
- réaliser spécifiquement la synthèse du L-glycéraldéhyde à partir du glycérol grâce à une oxydoréductase issue de la biodiversité

D'autres stratégies auraient pu être envisagées comme consommer le D-glycéraldéhyde notamment par l'utilisation d'une enzyme de type isomérase ou bien une autre lyase. Cependant l'utilisation d'une enzyme de type aldolase ou transcétolase nécessiterait l'ajout d'un substrat supplémentaire dans le milieu. De plus il est possible que le glycolaldéhyde soit substrat de aussi ce qui rendrait encore plus complexe la gestion des conditions réactionnelles.

La recherche d'une oxydoréductase sera présentée ultérieurement. Nous commencerons avec la partie concernant la FSA.

5. Recherche d'une 3^{ème} mutation pour lever l'inhibition de la FSA L107Y A129G par le D-glycéraldéhyde

5.1. Stratégie

L'objectif est de mettre en évidence une position différente, dans le site actif de notre mutant, du D et L-glycéraldéhyde et ainsi déterminer une 3^{ème} mutation qui lèverait l'inhibition. En effet, il est possible que le D-glycéraldéhyde ait des interactions plus fortes que le L-glycéraldéhyde dans le site actif ce qui empêcherait la réaction. Deux solutions s'offraient à nous pour tenter de lever cette inhibition par mutagénèse de la FSA L107Y A129G :

- la mutagénèse dirigée en s'appuyant au préalable sur des outils bio-informatiques
- la mutagénèse aléatoire

Nous avons décidé d'utiliser la première option qui consiste à se baser sur la bio-informatique et plus précisément sur des calculs de docking (amarrage moléculaire) pour déterminer la nouvelle mutation à effectuer. Le docking est une méthode qui prédit

l'orientation d'une molécule par rapport à une autre (ou un site actif d'enzyme dans notre cas) ainsi que les interactions. Cette approche a été utilisée afin de trouver la position du glycéraldéhyde, la plus « probable » d'un point de vue énergétique, dans le site actif de l'enzyme.

Cette technique est fréquemment utilisée en pharmacologie pour réduire le nombre d'essais expérimentaux lors de la recherche de nouveaux médicaments. La mutagenèse aléatoire n'a pas été utilisée pour des questions de temps car même si cette méthode permet de créer rapidement un grand nombre de mutants, le travail de criblage qui suit est long et fastidieux.

Dans un premier temps le fichier informatique (.PDB) de notre double mutant L107Y A129G a été créé en s'appuyant sur celui de la FSA_{wt}. Dans un second temps, nous avons réalisé l'étude de docking des deux énantiomères pris séparément. Leur position dans le site actif ainsi que leur proximité avec les différents acides aminés ont été étudiées.

5.2. Création du fichier contenant les coordonnées spatiales du mutant L107Y A129G grâce au logiciel SCWRL4

La détermination de la conformation des chaînes latérales est très importante dans la prédiction de structure des protéines. De nombreuses méthodes rapides et efficaces existent (SCWRL, SCWRL3...) mais elles sont principalement liées aux bases de rotamers (isomères de conformation), venant de la conformation des chaînes latérales de protéines connues par des méthodes statistiques de clustering. Plus récemment, le logiciel SCWRL4^[122] a été développé et permet de réaliser des calculs plus précis et rapides en se basant sur une nouvelle base de données de rotamers. Ce nouveau programme est résumé dans la Figure 54.

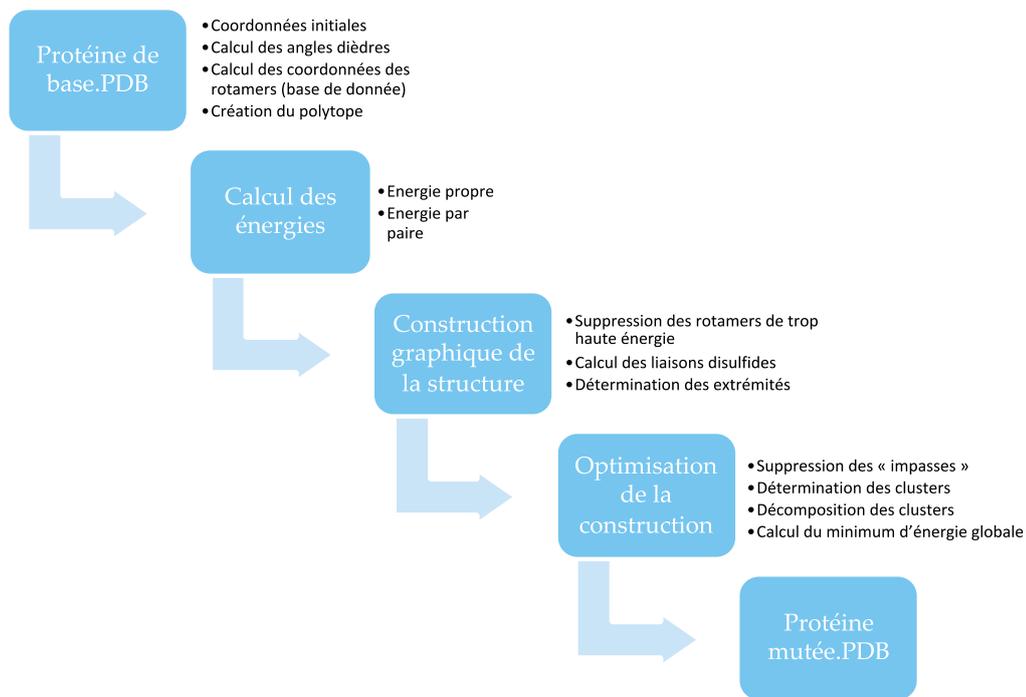


Figure 54 : différentes étapes de l'utilisation du logiciel SCWRL4 pour la création du fichier .PDB de la FSA L107Y A129G

Il utilise :

- une nouvelle base de données de structures de rotamers basée sur les calculs de densités de noyaux
- une moyenne de toutes les conformations sur leur position dans la base de données des rotamers
- une étude de la direction des liaisons hydrogènes et de leur fonction
- une étude du potentiel des interactions Van Der Waals atome-atome
- une détection de collision en utilisant des polytopes (figure présentant un grand nombre de symétries) distincts
- trois algorithmes de décomposition pour résoudre les problèmes combinatoires
- une optimisation de tous les paramètres par détermination des interactions graphiques avec le cristal en utilisant des opérateurs de symétrie de groupes cristallographiques définis

Une fois le fichier créé, une minimisation de la structure totale de la protéine avec l'aide du logiciel NAMED^[123] (Scalable Molecular Dynamics) a été réalisée. Ce logiciel permet de réaliser des dynamiques moléculaires sur des molécules de haut poids moléculaire. NAMED est basé sur le langage Charm++ qui permet d'effectuer de très nombreuses tâches

simultanément en diminuant les temps de calculs. Une minimisation par étape a été effectuée sur 200000 pas ainsi qu'une dynamique moléculaire dans une matrice de molécules d'eau à température ambiante pour équilibrer la protéine. Le même travail a été réalisé pour la création des fichiers pour les molécules de D-glycéraldéhyde et L-glycéraldéhyde. Tout ce travail a été effectué en collaboration avec le Professeur Frédéric AFFOUARD de l'UFR de Physique de l'Université de Lille et Natalia CORREIRA maître de conférences, de l'Unité Matériaux et Transformations (UMET-UMR8207) de l'Université de Lille.

5.3. Études de Docking

Les études de docking du D-glycéraldéhyde et L-glycéraldéhyde sur la FSA L107Y A129G ont été réalisées avec le logiciel autodock^[124]. La protéine, dont la structure a été minimisée dans une matrice de molécules d'eau est sélectionnée comme rigide (aucun degré de liberté). C'est à dire qu'elle restera fixe durant le docking, seuls les ligands seront flexibles et auront un grand degré de liberté de mouvement. L'Algorithme Génétique Lamarckien^[125] (LGA) a été sélectionné pour réaliser les docking. Celui-ci se base sur un gradient optimisé et recherche les minimums d'énergies locaux. Le LGA a été défini comme le plus efficace pour réaliser des études de docking entre enzymes et ligands grâce à une nouvelle fonction d'énergie libre empirique introduite lors de la recherche des plus basses énergies. Le site actif de la FSA étant connu, la première étape fut de créer une matrice de potentiel autour de celui-ci pour déterminer l'endroit où le docking doit être réalisé sur la protéine. Les coordonnées de la boîte sont (x.y.z) (16,075.-9,531.26562) et sa taille 40Å.40Å.40Å (Figure 55).

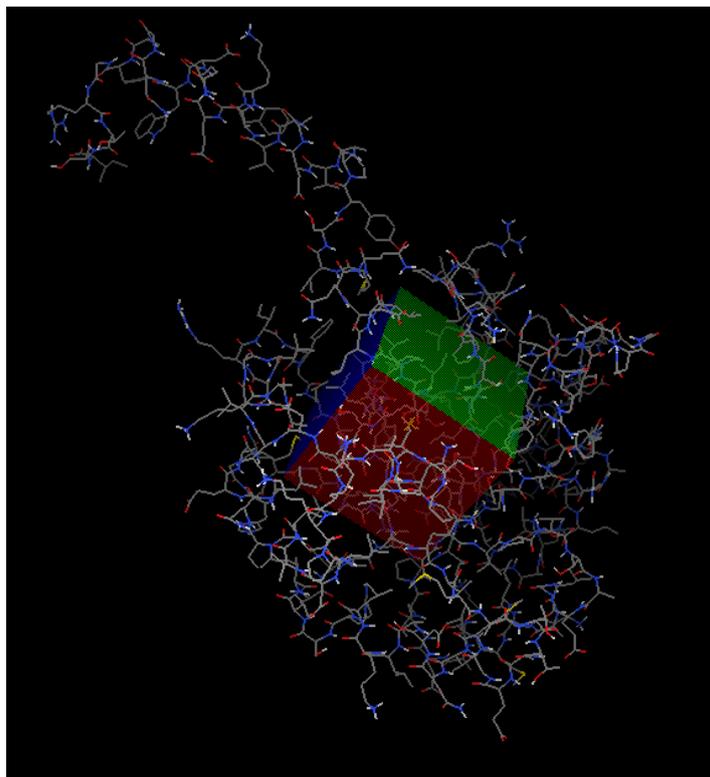


Figure 55 : position de la boîte d'amarrage du glycéraldéhyde (ligand) dans le site actif de la FSA L107Y A129G

La matrice a été définie autour du site actif en prenant soin d'inclure les acides aminés intervenant dans la réaction :

- la lysine 85 (formation de l'iminium)
- la tyrosine 107 (une mutation effectuée)
- la thréonine 109 (rôle dans la formation de l'iminium)

Ces interactions nous serviront de base pour valider ou non notre docking. En plus de ces interactions nous justifierons aussi le choix d'un modèle par rapport à l'autre en fonction du positionnement dans le site et plus particulièrement l'orientation du glycéraldéhyde dans la poche de liaisons hydrogènes.

Les calculs ont été réalisés sur 1000 et 2000 pas (positions générées). Le logiciel regroupe ensuite en 3 clusters les 2000 positions générées en se basant sur les proximités spatiales. Enfin le candidat représentatif de chaque cluster est sélectionné statistiquement. Les résultats étant très similaires seuls ceux pour 2000 pas ont été sélectionnés.

5.3.1. Avec le L-glycéraldéhyde

Sur les trois représentants par cluster que le logiciel a sélectionné, le numéro 2 ne met pas en relation le L-glycéraldéhyde et la lysine 85 de la FSA, il a donc naturellement été éliminé. Les valeurs d'énergie sont de $-3,36 \text{ kcal.mol}^{-1}$ pour le cluster 1 et $-2,77 \text{ kcal.mol}^{-1}$ pour le cluster 3. Ces valeurs confortent la pertinence de notre docking car elles sont très proches. A noter que les énergies sont calculées par addition des énergies de torsion, des interactions de Van der Waals, hydrogènes, de désolvatation et électrostatiques. Une valeur de « constante d'inhibition » est aussi indiquée, 3,42 mM pour le cluster 1 et 9,30 mM pour le cluster 3. Cependant, la pertinence de ces valeurs reste à prouver car très peu d'informations sont disponibles sur cette étape. Nous savons juste que ces calculs sont basés sur des relations connues entre ligands et protéines. La position de ces deux représentants au niveau de la protéine est également correcte comme nous le montre la Figure 56 puisque les deux molécules sont situées dans le site actif.

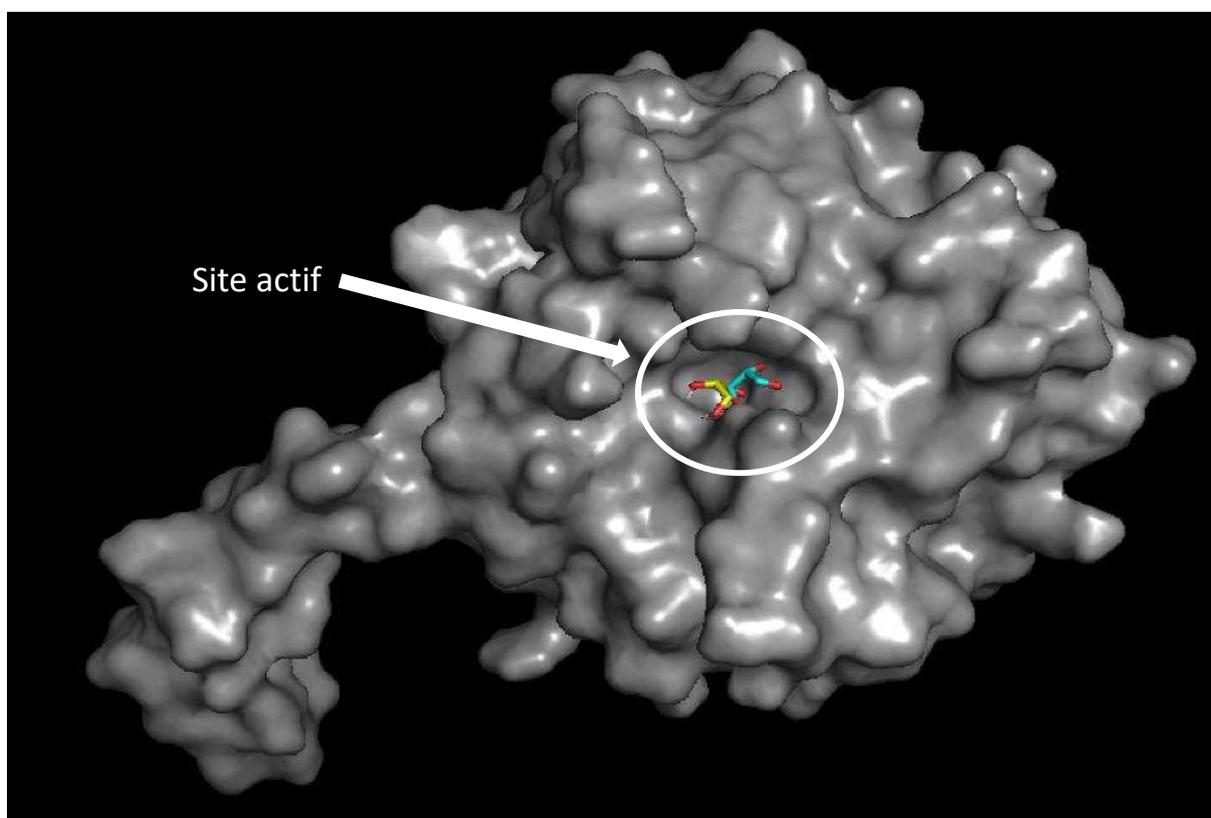


Figure 56 : position du L-glycéraldéhyde des clusters 1 (jaune) et 3 (bleu) après docking avec en rouge les fonctions alcools

On remarque tout de même à première vue que les deux molécules ne sont pas tout à fait orientées de la même manière dans le site actif. En zoomant sur la poche du site

actif, il est alors possible de visualiser les différentes interactions ainsi que la position dans la poche d'interaction hydrogène du site actif (Figure 57 a et b).

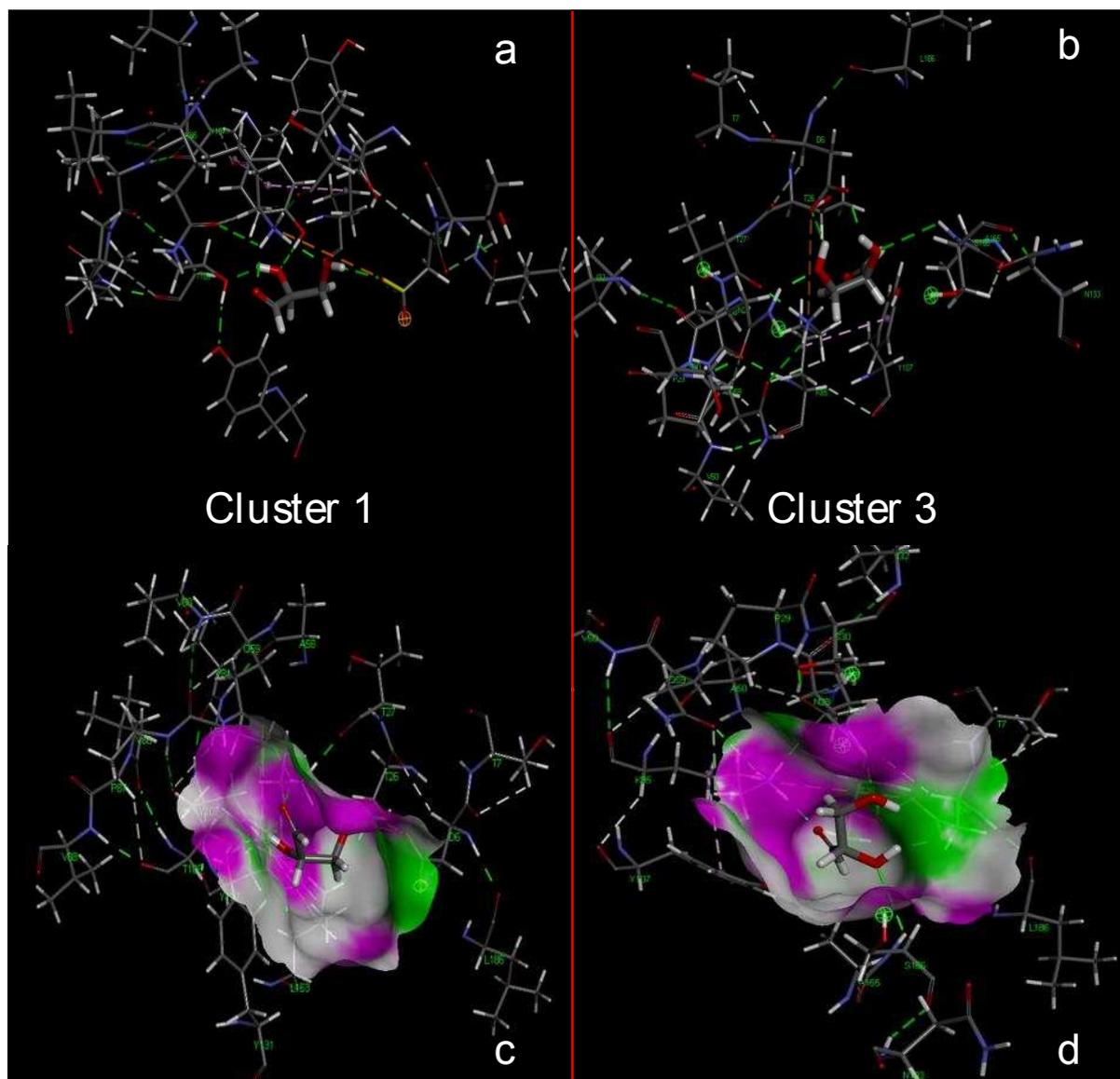


Figure 57. a et b : interactions glycéraldéhyde/FSA avec en orange les interactions de charges, en blanc les interactions entre hydrogènes de carbones et oxygènes, en vert les interactions hydrogènes et en sphère avec les molécules d'eau. **c et d :** interactions entre le ligand et la poche de liaisons hydrogène de l'enzyme avec en rose l'aspect donneur et en vert l'aspect accepteur de la poche.

Les interactions qui prédominent dans les deux clusters sont les liaisons hydrogènes. De plus, même si les molécules de L-glycéraldéhyde paraissent avoir une position très différente dans le site actif (Figure 57 c et d), tout semble cohérent puisque les alcools sont orientés du côté donneur et l'aldéhyde côté accepteur de liaisons hydrogènes dans le site actif de

l'enzyme. Par conséquent, les deux positions paraissent plausibles de par les interactions déterminées.

En s'intéressant aux acides aminés du site actif qui sont en interactions avec le ligand (Figure 58), des différences entre les représentants du cluster 1 et 3 peuvent être observées.

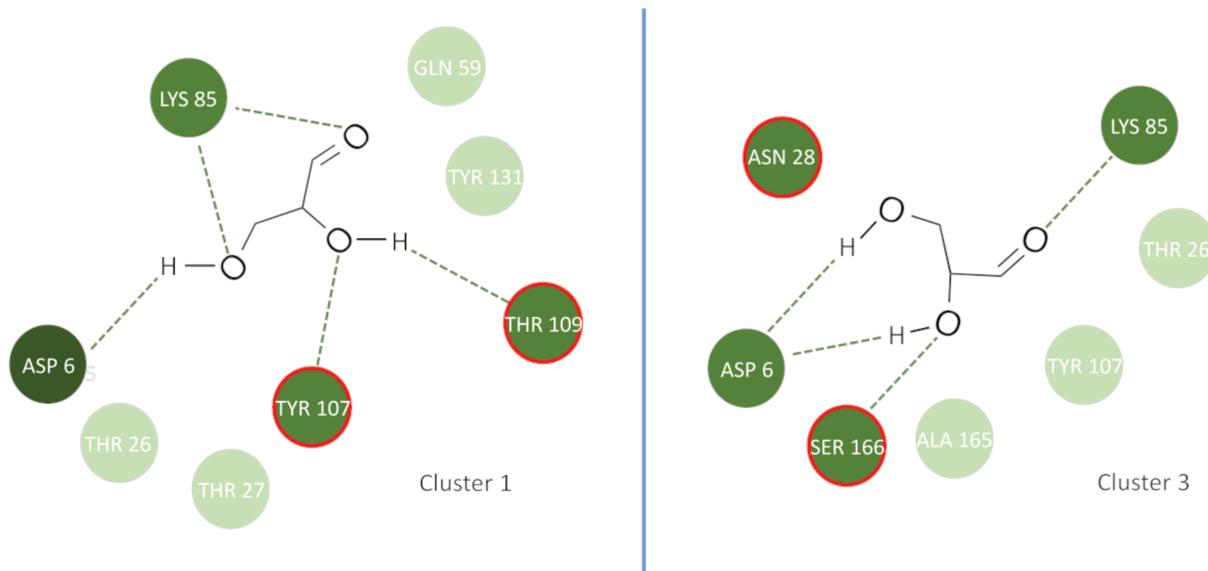


Figure 58 : acides aminés du site actif de la FSA mutant impliqués dans les interactions avec le L-glycéraldéhyde des clusters 1 et 3. En vert clair les acides aminés en interactions de Van der Waals, en vert foncé les interactions hydrogènes et entourés en rouges les acides aminés différents d'un cluster à l'autre. Les acides aminés non liés par pointillés sont trop éloignés pour interagir directement, l'interaction se fait via une molécule d'eau

Pour les deux clusters, l'interaction entre l'aldéhyde et la lysine 85 est présente. Cependant, en ce qui concerne la fonction alcool secondaire, dans les deux cas c'est un alcool aliphatique qui interagit, soit la thréonine 109 soit la sérine 166. De plus, l'acide aspartique 6 interagit avec l'alcool primaire pour les deux clusters. Enfin, l'interaction avec la tyrosine 107 (lysine avant mutation) du représentant du cluster 1 nous pousse à utiliser ce modèle dans la suite de l'étude pour comparer avec la position du D-glycéraldéhyde. En effet, étant donné que la mutation en position 107 donne à la FSA la capacité de cliver le L-glycéraldéhyde, et que le docking montre que cet acide aminé est en interaction avec le L-glycéraldéhyde, notre choix de continuer l'étude avec le représentant du cluster 1 est justifié. L'absence d'interactions avec la mutation en position 129 (alanine devenue glycine) est logique puisqu'elle avait pour but d'augmenter l'espace dans le site actif et de diminuer les interactions.

5.3.2. Avec le D-glycéraldéhyde

La même démarche que celle décrite ci-dessus a été menée avec le D-glycéraldéhyde. Les 3 représentants des clusters ont une interaction avec la lysine 85, ce qui n'était pas le cas avec le L-glycéraldéhyde, et se trouvent dans le site actif de l'enzyme (Figure 59).

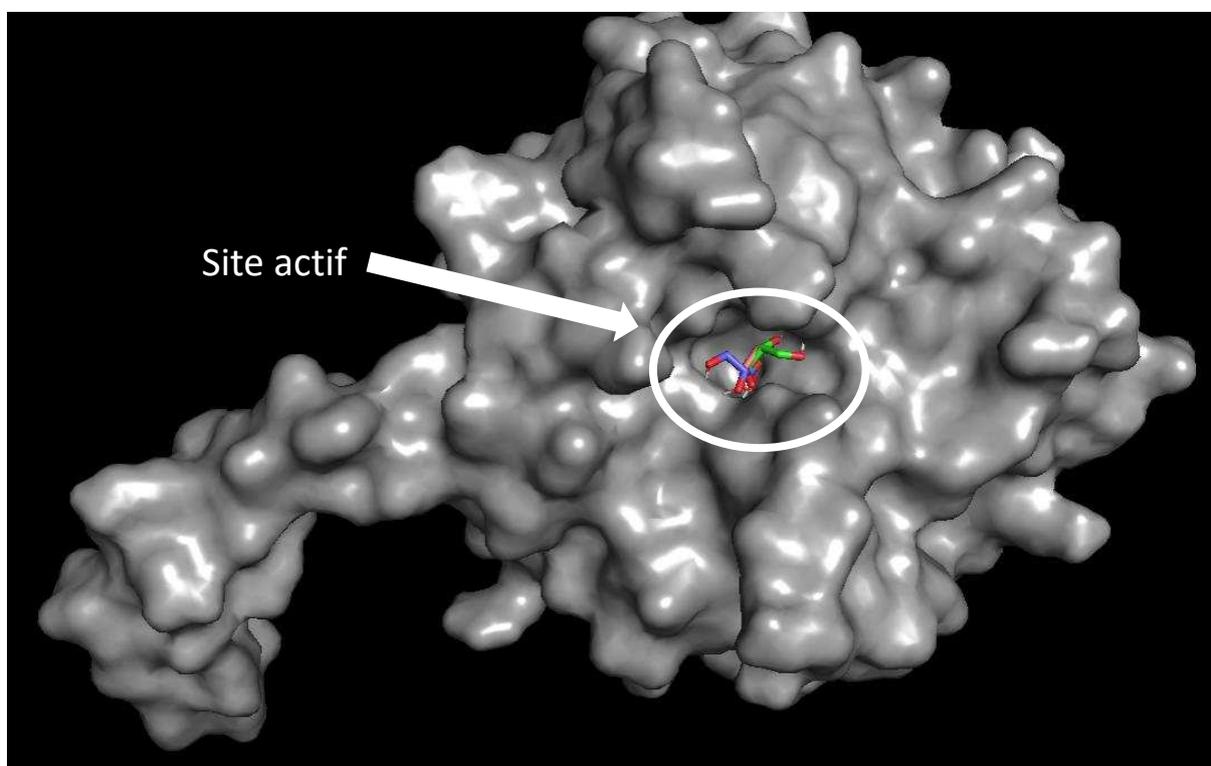


Figure 59 : position du D-glycéraldéhyde des clusters 1 (violet), 2 (orange) et 3 (vert) dans le site actif après docking, les fonctions alcools sont représentées en rouge

Les valeurs d'énergies calculées sont de $-3,23 \text{ kcal.mol}^{-1}$ pour le cluster 1, de $-3,17 \text{ kcal.mol}^{-1}$ pour le cluster 2 et $-2,86 \text{ kcal.mol}^{-1}$ pour le cluster 3. Ces valeurs sont approximativement les mêmes que celle obtenues avec le docking du L-glycéraldéhyde ($-3,36 \text{ kcal.mol}^{-1}$ et $-2,77 \text{ kcal.mol}^{-1}$). Les constantes d'inhibition ont également été calculées et sont de l'ordre de 4,5 mM tout comme pour le L-glycéraldéhyde (de l'ordre du millimolaire). Deux des clusters donnent des positions du ligand assez similaires et un cluster donne une position différente. Les représentants des clusters 2 (orange) et 3 (vert) sont superposés alors que celui du cluster 1 (violet) a une orientation différente.

Comme pour le L-glycéraldéhyde, les interactions prédominantes sont les interactions hydrogènes et elles semblent à première vue plus nombreuses (Figure 60).

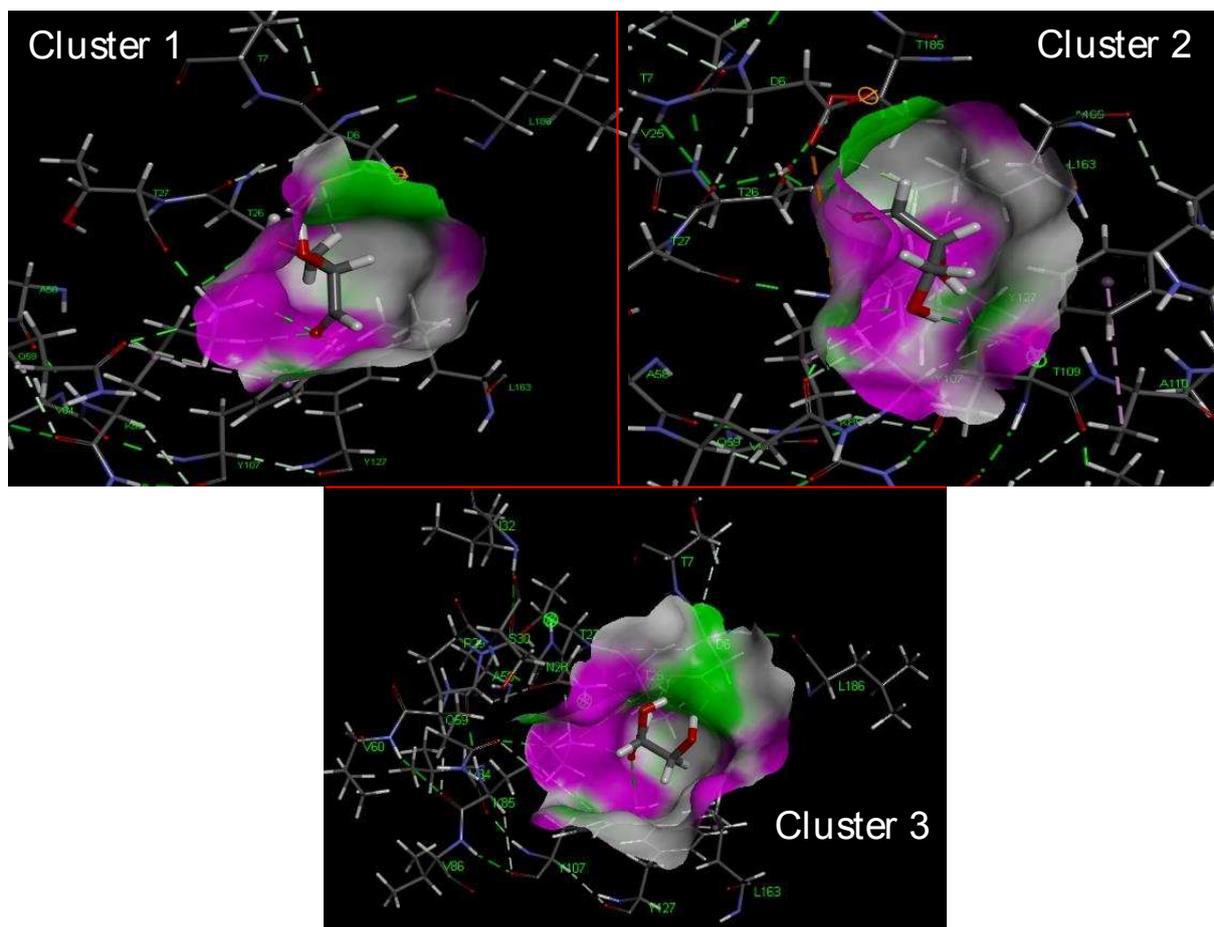


Figure 61 : interactions entre le D-glycéraldéhyde et la poche de liaisons hydrogènes de la FSA avec en rose l’aspect donneur et en vert l’aspect accepteur de la poche

Pour les clusters 2 et 3, la position est différente mais les alcools (donneurs de liaisons hydrogènes) sont orientés du côté accepteur et l’aldéhyde (accepteur de liaisons hydrogènes) du côté donneur de la poche d’interactions hydrogènes de l’enzyme ce qui est plus logique que pour le cluster 1.

Concernant les interactions avec les acides aminés du site actif, la lysine 85 est bien en interaction avec l’aldéhyde ce qui conforte la « viabilité » de notre docking (Figure 62).

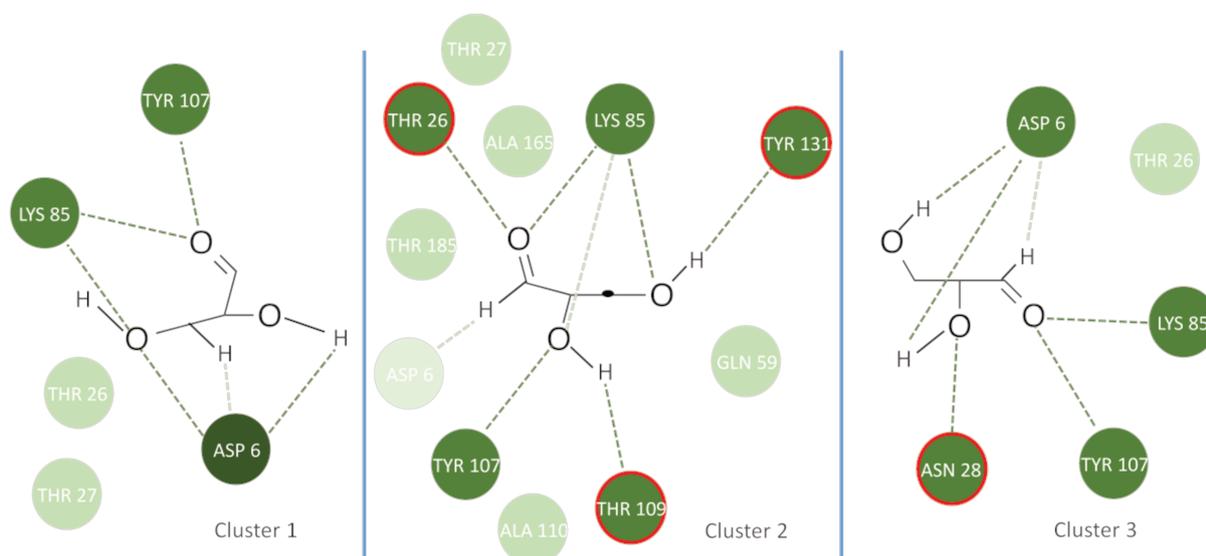


Figure 62 : acides aminés du site actif de la FSA impliqués dans les interactions avec le D-glycéraldéhyde des clusters 1,2 et 3. En vert clair les acides aminés en interactions de Van der Waals, en vert foncé les interactions hydrogènes et entourées en rouges les acides aminés différents d'un cluster à l'autre. Les acides aminés non liés par pointillés sont trop éloignés pour interagir directement, l'interaction se fait via une molécule d'eau

Des interactions « directes » (sans passer par des molécules d'eau) sont observées avec des acides aminés communs entre les clusters : LYS 85, TYR 107, ASP 6. Cependant, les clusters 2 et 3 présentent des différences par rapport au cluster 1. Des interactions supplémentaires avec les acides aminés 131, 26 et 109 sont présentes pour le cluster 2. Des interactions avec l'acide aminé 28 sont présentes pour le cluster 3 (Figure 62). Notre choix s'est porté sur le cluster 2. En effet, le résidu 131 est connu pour jouer un rôle dans la réaction catalysée par la FSA puisqu'il intervient dans la formation de l'énamine^[46]. D'autre part, la présence d'interactions avec la tyrosine 107, une des mutations effectuées, responsable de l'activité de clivage du L-glycéraldéhyde, conforte l'idée de sélectionner le représentant du cluster 2.

5.3.3. Comparaison des études de docking du L-glycéraldéhyde (cluster 1) et D-glycéraldéhyde (cluster 2)

La position des deux énantiomères dans le site actif est très similaire. Les deux molécules sont quasiment superposées. Cependant, leur orientation est différente (Figure 63). Sur le D-glycéraldéhyde (molécule rose), l'aldéhyde est orienté vers le haut du côté de l'acide aspartique 6, alors que sur le L-glycéraldéhyde (molécule verte) il est orienté en bas à gauche du côté de la tyrosine 107. Les calculs des énergies ainsi que les constantes d'inhibition sont quasiment identiques (Tableau 11).

	Énergies	Constantes d'inhibition
D-glycéraldéhyde	-3,17kcal.mol ⁻¹	4,50 mM
L-glycéraldéhyde	-2,77 kcal.mol ⁻¹	3,42 mM

Tableau 11 : comparaison des valeurs d'énergies et de constantes d'inhibition pour les deux représentants sélectionnés

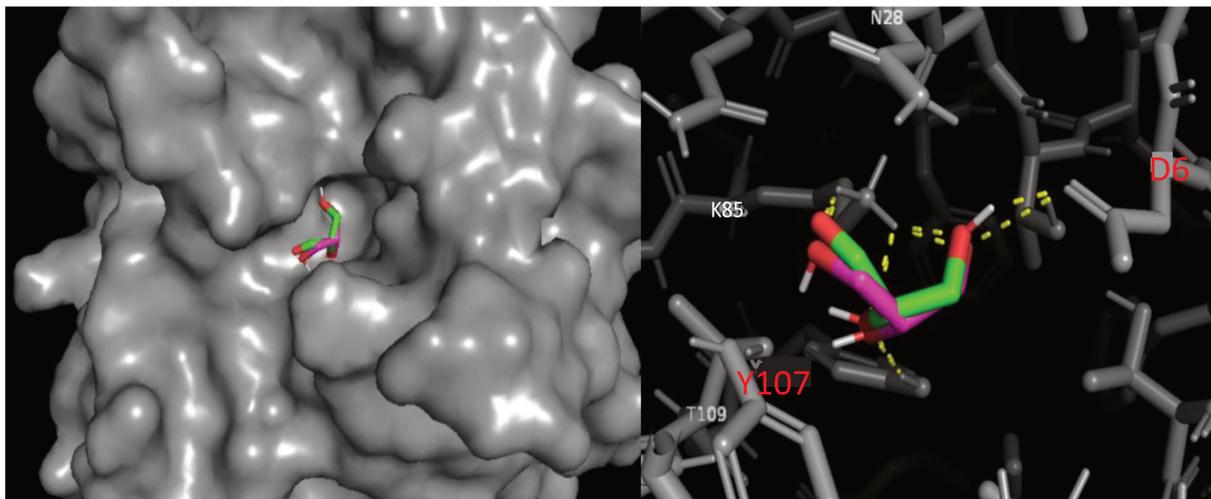


Figure 63 : superposition des deux énantiomères sélectionnés dans le site actif de la FSA L107Y A129G avec en vert le L-glycéraldéhyde et en rose le D-glycéraldéhyde

Concernant les acides aminés directement en interaction avec le D-glycéraldéhyde et le L-glycéraldéhyde, des similitudes sont visibles puisque les deux énantiomères sont en interaction avec la lysine 85, la thréonine 109 et la tyrosine 107. Cependant, le D-glycéraldéhyde a des interactions avec trois autres acides aminés, la thréonine 26, la tyrosine 131 et l'acide aspartique 6 (Figure 64).

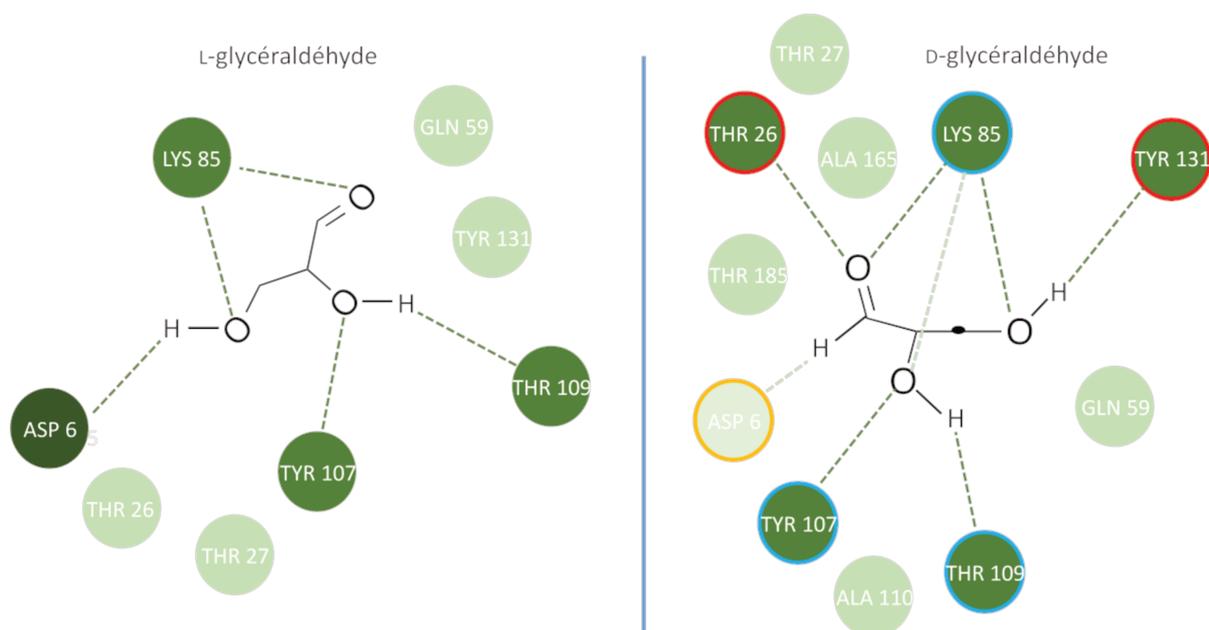


Figure 64 : comparaison des acides aminés du site actif de la FSA impliqués dans les interactions avec le D-glycéraldéhyde et le L-glycéraldéhyde avec en bleu les interactions carbonés/hydrogènes, en vert clair les interactions de Van der Waals, en vert foncé les interactions hydrogènes et entourés en rouges les acides aminés différents d'un cluster à l'autre. Les acides aminés non liés par pointillés sont trop éloignés pour interagir directement, l'interaction se fait via une molécule d'eau

La tyrosine 131 a été définie comme importante dans le mécanisme réactionnel d'aldolisation. Cependant, les acides aminés en position 131 et 26 ont été sélectionnés comme potentiellement responsables de l'inhibition de l'enzyme puisque ce sont les deux acides aminés qui créent des interactions uniquement avec le D-glycéraldéhyde. La position 6 n'a dans un premier temps pas été étudiée même si elle a été décrite comme potentiellement responsable de l'énantio-préférence. En effet nous avons pour objectif dans un premier temps de diminuer les interactions uniquement avec le D-glycéraldéhyde en réalisant une troisième mutation d'un acide aminé dans le site actif qui ne serait pas en contact avec le L-glycéraldéhyde mais uniquement avec le D-glycéraldéhyde. Cependant l'étude de la position 6 reste une perspective pour la suite à étudier en cas d'échec.

La mutation effectuée en position 107 (leucine en tyrosine) peut compenser l'action de la tyrosine 131 ce qui conforte l'idée de muter cet acide aminé pour diminuer les interactions avec le D-glycéraldéhyde. Le résidu 26 n'est lui pas décrit comme indispensable à la réaction enzymatique. Notre stratégie est de réaliser deux mutants supplémentaires, les FSA L107Y A129G Y131G et FSA T26G L107Y A129G. Une autre stratégie aurait été de réaliser ces deux mutations en remplaçant les acides aminés 131 et 26 par tous les autres acides

aminés et finir par un couplage des deux mutations mais nous n'avons pas suffisamment de temps pour ce type d'approche. Pour diminuer les interactions, l'acide aminé avec la plus petite chaîne latérale a été naturellement sélectionné, à savoir la glycine qui ne possède pas de chaîne latérale.

5.4. Création des mutants définis par docking

Les mutations ont été effectuées et validées par séquençage (Figure 65 et Figure 66).

```

          10          20          30          40          50
          |          |          |          |          |
.....|.....|.....|.....|.....|
ATGGAAGTGTATCTGGATACTTCAGACGTTGTTGCGGTGAAGGCGCTGTC
ACGTATTTTCCGCTGGCGGGTGT GAC CACTAACCCAAGCATTATCGC
      CGCTGGCGGGTGT GGG CACTAACCCAAGCA
CGCGGGTAAAAACCGCTGGATGTTGTGCTTCCGCAACTTCATGAAGCGA
TGGGCGTTCAGGGGCGTCTGTTTGCCAGGTAATGGCTACCACTGCCGAA
GGGATGTTAATGACGCGCTTAAGCTGCGTTCTATTATTGCGGATATCGT
GGTAAAAGTTCCGGTGACCGCCAGGGGCTGGCAGCTATTAAGATGTTAA
AAGCGGAAGGGATTCCGACGTATGGAACCGCGGTATATGGCGCAGCACAA
GGGCTGCTGTCGGCGCTGGCAGGTGCGGAATATGTTGGGCC TTA CGTT
      CAGGTGCGGAATATGTTGGGCC TGG CGTT
AATCGTATTGATGCTCAGGGCGGTAGCGGCATTCAGACTGTGACCGACTT
AATCGTATTG
ACACCAGTTATTGAAAATGCATGCGCCGAGGCGAAAGTGCTGGCAGCGA
GTTTCAAACCCCGCTCAGGCGCTGGACTGCTTACTGGCAGGATGTGAA
TCAATTACTCTGCCACTGGATGTGGCACAACAGATGATTAGCTATCCGGC
GGTTGATGCCGCTGTGGCGAAGTTTGGAGCAGGACTGGCAGGGAGCGTTTG
GCAGAACGTCGATTTAA

```

Figure 65 : séquence génétique de la FSA L107Y A129G en noir et des amorces contenant les mutations positions 26G (orange) et 131G (bleu)

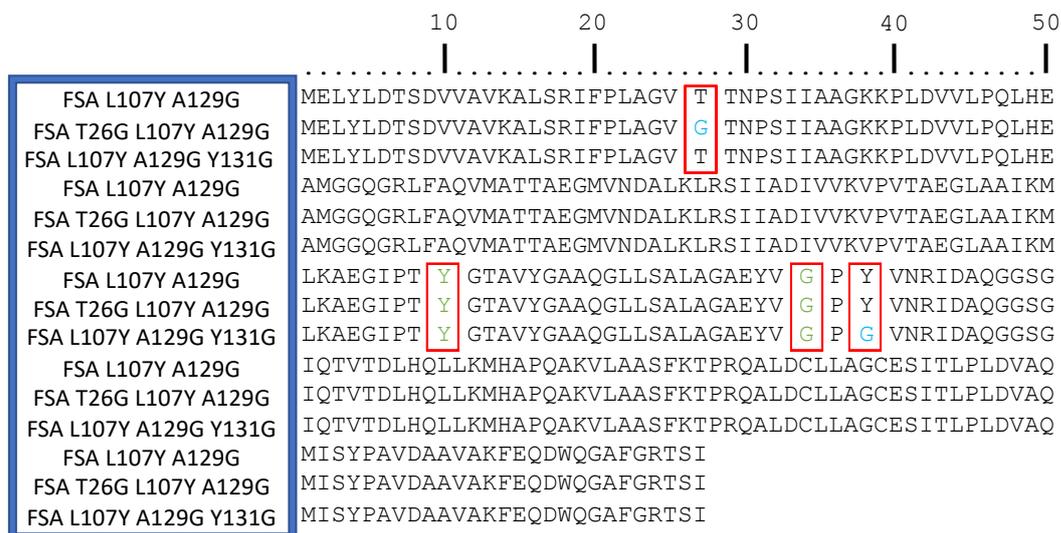


Figure 66 : séquence peptidique des FSA L107Y A129G, T26G L107Y A129G et L107Y A129G Y131G

Les enzymes ont été produites à petite échelle (20 mL) et criblées grâce au test mis au point avec l'ADH de *Saccharomyces cerevisiae* (Figure 46). La production a été vérifiée par SDS-PAGE et un dosage des protéines totales par la méthode de Bradford. L'activité de clivage du L-glycéraldéhyde par la FSA L107Y A129G et par les deux autres mutants produits a été comparée en utilisant la même quantité de protéines totales. Cependant, les mutants n'étant pas purifiés pour le criblage, leur activité enzymatique est probablement sous-estimée. Les résultats sont présentés sur la Figure 67.

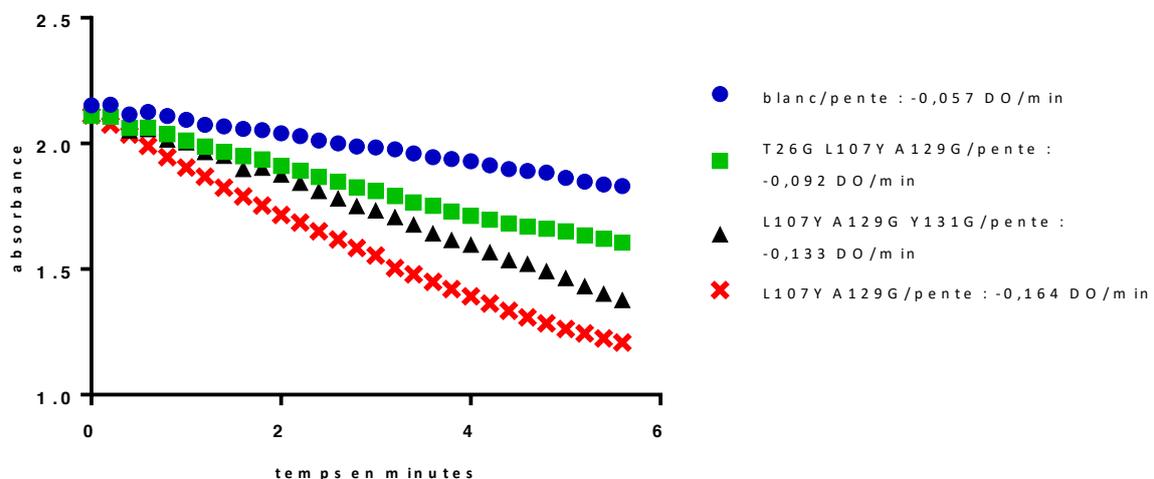


Figure 67 : criblage des mutants produits sur la réaction de clivage du L-glycéraldéhyde réalisé en plaque 96 puits de 200µL avec une concentration de L-glycéraldéhyde de 50 mM, 5 U d'ADH, 1 mg de FSA non purifiée, 0,4 mM de NADH dans du tampon Gly-gly 50 mM pH 8. Le blanc est effectué sans FSA (présenté ici car c'est pour ce blanc que la dérive est la plus importante) et sans glycéraldéhyde

Le criblage a été réalisé en triplicata mais pour des raisons de lisibilité seule la moyenne est affichée. En réalisant une régression linéaire de chaque droite, les activités ont pu être comparées. A première vue, que ce soit pour la mutation supplémentaire T26G ou Y131G, il y a une diminution de l'activité enzymatique. Cependant, les deux mutants semblent avoir gardé une activité ce qui est pour le Y131G très étonnant car la tyrosine 131 a été déterminée comme cruciale pour l'activité d'aldolisation. Deux raisons peuvent expliquer ce résultat. Soit cet acide aminé n'a pas un rôle majeur dans le cas de la rétroaldolisation, ceci serait surprenant puisque théoriquement se sont les mêmes acides aminés du site actif de la FSA qui réagissent pour catalyser aussi bien la rétroaldolisation que l'aldolisation. Soit la mutation L107Y compense le rôle que joue la tyrosine position 131 comme nous en avons fait l'hypothèse précédemment. Une des perspectives de cette étude serait de muter la position 107 pour remettre l'acide aminé initial, qui était une leucine, et observer si l'activité de clivage du L-glycéraldéhyde est toujours présente. Si cela se produit, il serait démontré que la tyrosine 131 n'est pas indispensable au mécanisme réactionnel dans notre cas. Le mutant Y131G est plus actif que le mutant T26G, environ 1,5 fois. Ces résultats montrent que nos mutations n'ont pas affecté la réaction de clivage du L-glycéraldéhyde catalysée par la FSA L107Y A129G, ce qui est encourageant pour la suite de notre étude.

Le mélange racémique L/D-glycéraldéhyde a ensuite été testé pour voir si une activité se dégagait (Figure 68).

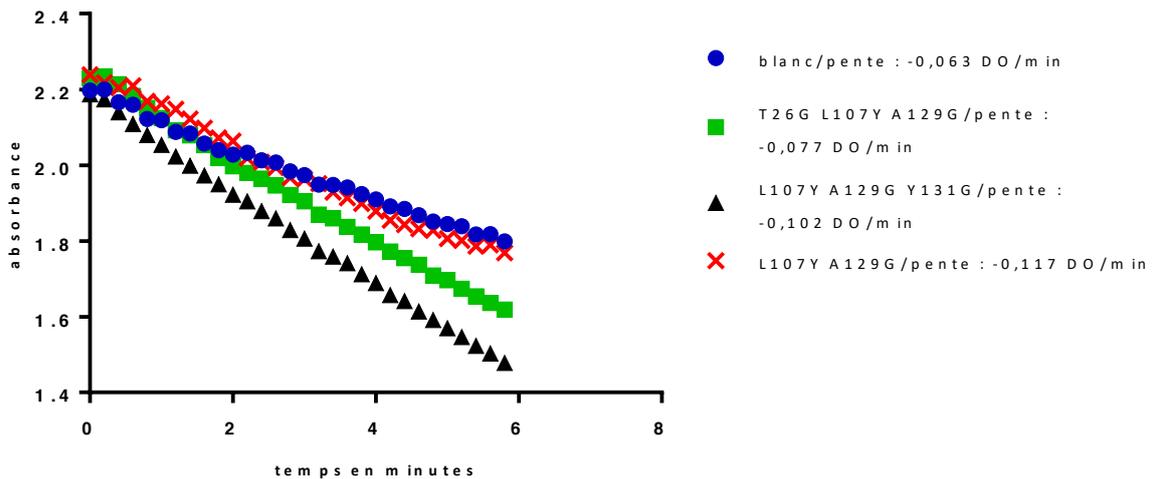


Figure 68 : criblage des mutants sur la réaction de clivage du L/D-glycéraldéhyde réalisé en plaque 96 puits de 200µL avec une concentration de D/L-glycéraldéhyde de 50 mM, 5 U d'ADH, 1 mg de FSA non purifiée, 0,4 mM de NADH dans du tampon Gly-gly 50 mM pH 8. Le blanc est effectué sans glycéraldéhyde et sans FSA (présenté ici car c'est pour ce blanc que la dérive est la plus importante)

L'activité du mutant L107Y A129G est identique à celle du blanc ce qui est logique puisqu'une inhibition en présence de D-glycéraldéhyde a été mise en évidence. Les deux mutants créés ont eux aussi une activité quasiment identique au mutant L107Y A129G. Ce résultat est étonnant au vu des activités déterminées précédemment sur le L-glycéraldéhyde puisque la mutation en position 26 avait entraîné une diminution d'activité. Même si ces expériences ont été réalisées en triplicata, ne se baser que sur ce résultat de criblage serait une erreur puisque nous travaillons sur lysat. Il est possible d'avoir des faux positifs lors de criblage sur lysat, notamment des enzymes qui dégradent le NADH. Une étude de suivi en RMN du proton de la réaction a été réalisée pour déterminer les éventuels produits formés et ainsi confirmer ou non l'activité réelle des deux nouveaux mutants.

Pour le mutant FSA T26G L107Y A129G, l'étude RMN ne montre pas de réaction de clivage du D ou L-glycéraldéhyde ni du mélange racémique. Cependant cette étude ne nous permet pas pour le moment de confirmer ou non la levée de l'inhibition. En effet, les réactions étant réversible et préférentiellement dans le sens de la synthèse, il est possible que notre FSA T26G L107Y A129G clive notre glycéraldéhyde mais catalyse beaucoup plus rapidement la réaction inverse d'aldolisation. Ceci peut expliquer la différence entre les résultats de la RMN et les cinétiques au spectrophotomètre. Pour conclure sur ce mutant il faudrait réaliser une réaction dans le sens de la synthèse en présence de D-glycéraldéhyde. Si

nous ne détectons toujours pas le produit, à ce moment il sera possible de conclure que l'inhibition n'a pas été levée.

Le mutant FSA L107Y A129G Y131G montre une activité de clivage du L-glycéraldéhyde suivant le même principe que pour le mutant FSA L107Y A129G. Le glycolaldéhyde n'est pas observé mais le produit d'aldolisation du L-glycéraldéhyde sur le glycolaldéhyde, le L-xylose est présent dans le milieu. Ce résultat est très surprenant, puisque comme évoqué dans la partie bibliographique, le résidu Y131 est indispensable à l'activité de l'enzyme. Le résidu Y131 est responsable de la reprotonation du carbanion comme décrit précédemment. Si l'on compare la distance du résidu Y131 par rapport au potentiel carbanion et la distance du résidu Y107 on remarque le 2^{ème} est beaucoup plus proche donc peut plus facilement avoir une action à ce niveau (Figure 69 : interaction hydrogène entre le L-glycéraldéhyde et le site actif du mutant de la FSA L107Y A129G avec en gris les atomes de carbone, en bleu les atomes d'azote, en rouge les atomes d'oxygène et en blanc les atomes d'hydrogène).

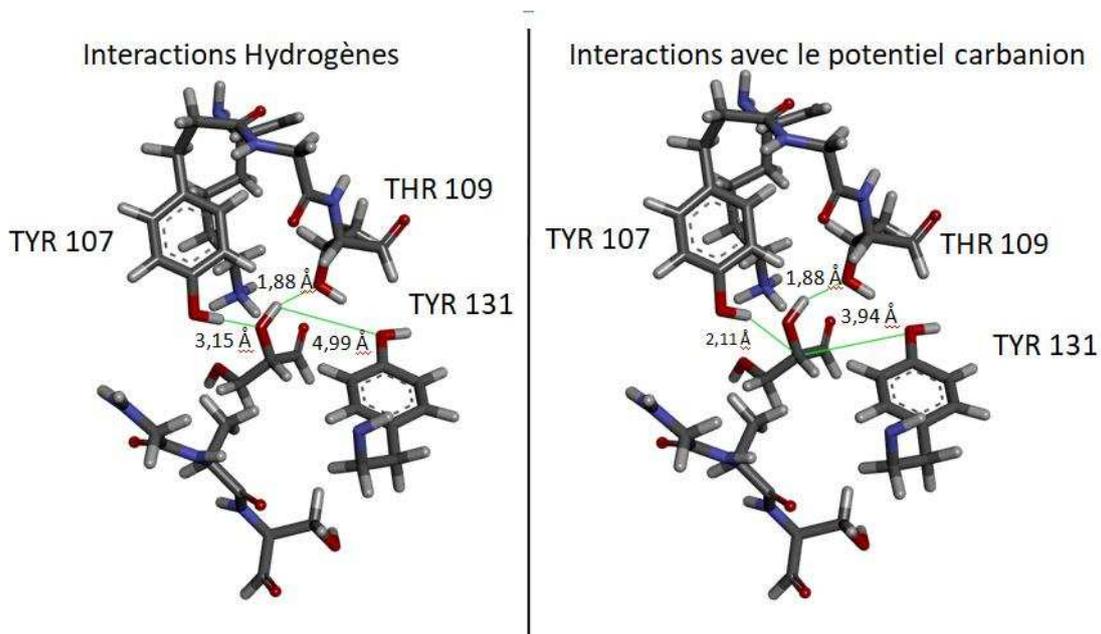


Figure 69 : interaction hydrogène entre le L-glycéraldéhyde et le site actif du mutant de la FSA L107Y A129G avec en gris les atomes de carbone, en bleu les atomes d'azote, en rouge les atomes d'oxygène et en blanc les atomes d'hydrogène

Dans un premier temps, on peut voir que la distance entre les alcools de la TYR 107 et de la THR 109 et du L-glycéraldéhyde sont plus faible (respectivement 3,15 Å et 1,88 Å) par rapport à la distance entre la TYR 131 et ce même alcool (4,99 Å) (Figure 69). Ceci laisse à penser que le glycéraldéhyde est déplacé dans le site actif en direction des résidus 107 et

109 ce qui va rendre plus aisé la reprotonation du carbanion après le clivage (Figure 15). En effet comme nous le montre les distances entre les résidus et le carbanion, la TYR 107 est beaucoup plus proche (2,11 Å) que le résidu TYR 131 (3,94 Å). Une liaison hydrogène étant comprise entre 1,2 Å et 2,5 Å il est donc plus probable statistiquement que ce soit le résidu TYR 107 qui reprotonne le carbanion et non la TYR 131.

Le fait de combiner cette mutation position 131 et les deux autres positions 107 et 129 pourrait remettre en question le caractère indispensable de ces acides aminés dans la réaction d'aldolisation catalysée par la FSA. Pour valider ce résultat, il serait intéressant de muter une par une les positions 107 et 129 pour revenir à l'acide aminé de départ. Ainsi si aucune réaction n'est observée, l'hypothèse d'un changement des acides aminés catalysant la réaction de la FSA lorsque les mutations L107Y A129G et Y131G sont présentes serait validée. Il est tout à fait possible que la tyrosine 107 compense l'action qu'avait la tyrosine 131. Une étude RMN a aussi été réalisée en présence D-glycéraldéhyde ou du mélange racémique, cependant aucuns signaux différents de ceux du glycéraldéhyde n'ont été détectés. Cette fois-ci il est par contre possible de conclure avec certitude que nous n'avons pas levé l'inhibition de notre enzyme puisque nous observons la synthèse de L-xylose en présence de L-glycéraldéhyde et seulement le glycéraldéhyde en présence du mélange racémique.

Pour le moment il est difficile de conclure sur le mutant FSA T26G L107Y A129G. Il n'est à l'heure actuelle pas possible de dire si l'inhibition a été levée ou non. Il faut approfondir l'étude de ce triple mutant et tester d'autres acides aminés en position 26 ou 131. Nous avons choisis dans un premier temps, pour diminuer au maximum les interactions remplacé la tyrosine en position 131 et la thréonine en position 26 par un glycine mais il est possible que ce changement soit trop différent et peut-être qu'une chaîne latérale juste moins longue serait suffisante. La position 6 aussi peut être un sujet d'étude intéressant et devra être envisagée par la suite. La pertinence de notre docking peut aussi être mise en cause. La réalisation d'une dynamique moléculaire doit aussi être étudiée puisque qu'elle permettrait d'observer le déplacement du glycéraldéhyde dans le site actif. Cependant cette méthode est très chronophage et demande une expertise plus poussée.

II. Recherche d'une glycérol oxydase dans la biodiversité

1. Introduction

La deuxième stratégie consiste à trouver une oxydase capable de synthétiser sélectivement le L-glycéraldéhyde à partir du glycérol. Pour cela nous avons décidé de développer au sein du laboratoire une approche de clustering comparable à celle du GENOSCOPE CEA d'Evry^[126,127]. L'idée est de rechercher des enzymes dans la biodiversité en se référant à la base de données Uniprot. Les enzymes présentes dans cette base « non vérifiées » (annotées automatiquement) sont généralement annotées en fonction de la proximité avec des clusters connus. Ces annotations peuvent être fausses notamment dans les familles d'enzymes très étendues et diverses (oxydo-réductases par exemple) et c'est sur ce biais que repose le programme que nous avons créé. Enfin, lorsque des nouvelles enzymes sont répertoriées, tous les substrats ne sont pas forcément testés donc rien ne dit qu'une enzyme décrite pour catalyser une réaction ne pourra pas être aussi active sur d'autres substrats. Un programme bio-informatique a donc été développé en se basant sur ces constatations.

2. Création du programme bio-informatique

2.1. Stratégie

L'approche qui a été développée repose sur un postulat : l'activité enzymatique dépend de la structure tridimensionnelle de la protéine qui elle-même découle de la structure primaire linéaire en acides aminés. Deux enzymes ayant des séquences primaires proches auront donc une activité semblable. C'est la première fois qu'une telle démarche est mise en œuvre au sein de l'Institut Charles Viollette. Toutes les étapes du programme ont été travaillées et optimisées. Dans notre cas, la réaction d'oxydation du glycérol en L-glycéraldéhyde (Figure 70) est très peu décrite dans la littérature et souvent cette activité est catalysée par des enzymes fongiques. Ce qui pose quelques problèmes car les temps de culture des champignons sont généralement longs et il est difficile d'exprimer une enzyme

fongique dans une bactérie puisque ce n'est pas toujours le même code génétique. D'autre part, sachant que notre objectif est de synthétiser le glyoxal à partir du glycérol, si une enzyme s'avère active sur le glycérol, elle peut éventuellement aussi oxyder le glycolaldéhyde en glyoxal puisque la structure des deux molécules est relativement proche. Il faut bien entendu garder aussi à l'esprit que l'oxydase pourra peut-être catalyser l'oxydation des deux alcools primaires afin de donner l'hydroxypropanedial. Enfin, nous avons pris en compte que la réaction d'oxydation soit catalysée sans cofacteurs. En effet, dans un contexte de potentiel développement industriel il est difficile d'envisager l'utilisation de ces molécules souvent très onéreuses.

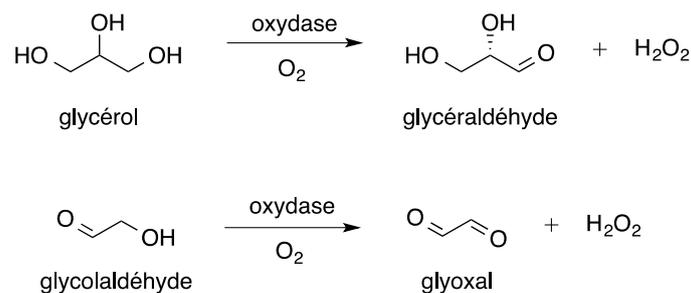


Figure 70 : oxydation du glycérol en L-glycéraldéhyde et du glycolaldéhyde en glyoxal

La Figure 71 illustre les différentes étapes qui peuvent être divisées en deux grandes parties, une première partie bio-informatique pour le développement d'un pipeline d'analyses et une seconde partie pour la réalisation des étapes de biologie moléculaire (amplification des gènes, clonage) et le criblage de l'activité enzymatique.

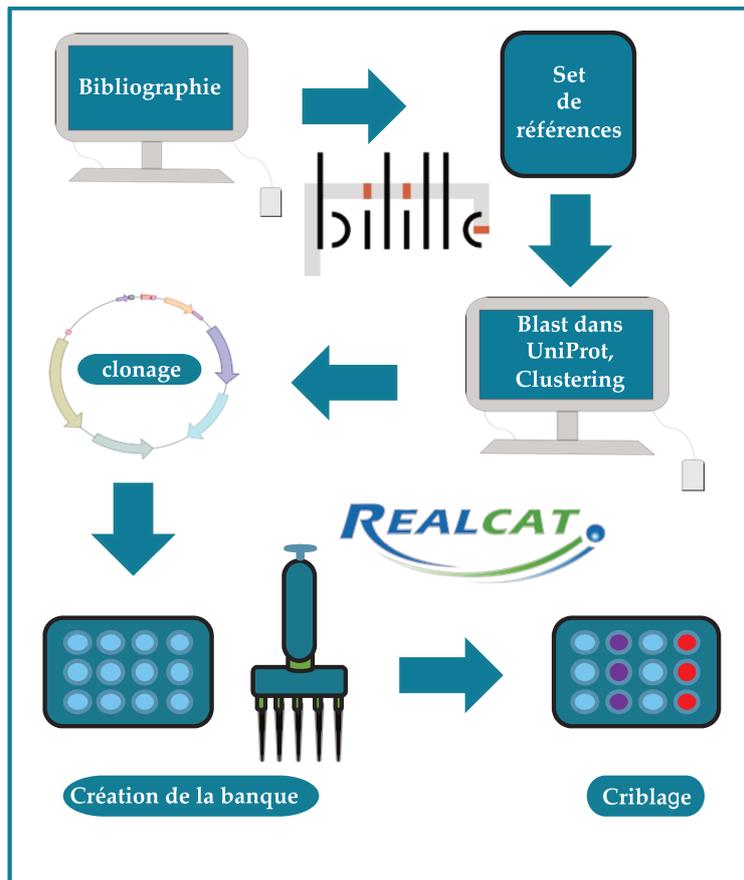


Figure 71 : stratégie de recherche de nouvelles oxydases dans la biodiversité

Le programme bio-informatique va se dérouler en différentes étapes :

- Rechercher dans la littérature des enzymes possédant des activités d'oxydation d'alcools primaires à courtes chaînes (2-4 carbones) pour créer un set de références de séquences primaires protéiques
- Générer un grand nombre de données par Blast dans toute la base de données Uniprot
- Regrouper des séquences similaires en clusters pour limiter le nombre de séquences à traiter (appelé étape de clustering)
- Cloner les gènes sélectionnés dans des vecteurs de surexpression et transformer des bactéries compétentes pour la production des enzymes
- Exprimer les protéines recombinantes et cribler l'activité enzymatique

Pour la réalisation du programme bio-informatique, nous avons répondu à un appel à projet de la plateforme de bio-informatique de Lille, appelée « Bilille », et obtenu l'aide de madame Isabelle GUIGON, ingénieure en bio-informatique, et de madame Hélène TOUZET, chercheuse CNRS et directrice de l'équipe BONSAI dans le laboratoire CRYSTAL de

l'Université Lille. L'intérêt de ce programme bio-informatique est d'automatiser les étapes mais aussi de pouvoir modifier tous les paramètres de Blast et de clustering à notre guise. Le programme comprend :

- un outil de Blast similaire à celui d'Uniprot dont tous les paramètres peuvent être modifiés, allant du nombre de candidats au pourcentage d'identité sur une longueur précise
- un filtre permettant d'éliminer les doublons, même si cela est peu probable, dans le cas où le set de références sélectionné contient une grande diversité
- MMseqs2 pour réaliser l'étape de clustering^[128] (Many-against-Many-searching) dont le pourcentage d'identité au sein de chaque cluster peut être modifié (MMseqs2 sera détaillé par la suite)
- un filtre pour ne garder que les enzymes bactériennes
- un filtre pour supprimer les clusters de petite taille (valeur modulable)
- un dernier filtre statistique pour extraire le candidat le plus représentatif du cluster, assuré par un module de MMseqs2.

A noter que toutes ces analyses peuvent s'effectuer soit étape par étape, soit en appliquant directement la totalité du pipeline. Le programme ne contient pas d'interface graphique, il est utilisable depuis le terminal d'un ordinateur par connexion à une machine virtuelle hébergée sur le Cloud de Bilille.

Les étapes de biologie moléculaire et de criblage des activités enzymatiques ont d'abord été développées manuellement au laboratoire, afin de déterminer les conditions optimales de chaque étape. Une automatisation du procédé a été réalisée sur les robots de la plate-forme EQUIPEX REALCAT, avec la collaboration de monsieur Egon HEUSON, ingénieur de recherche. C'est la première fois qu'une telle démarche est réalisée au sein de l'Institut Charles Viollette. Je présenterai essentiellement la stratégie développée manuellement ainsi que les résultats obtenus à chaque étape.

2.2. Création du set de références de séquences protéiques primaires

Le set de références a été élaboré de façon à contenir une dizaine d'enzymes catalysant des réactions d'oxydation proches de celle souhaitée, ainsi que des oxydases n'utilisant pas de cofacteurs. Onze enzymes ont été sélectionnées, ces enzymes sont présentes dans onze microorganismes différents (Tableau 12).

Nom Uniprot	Organisme	Référence
P93762	<i>Chondrus crispus</i>	[129]
P04841	<i>Pichia angusta</i>	[130]
A5VWI7	<i>Pseudomonas putida</i>	[131]
F2QW10	<i>Komagataella phaffi</i>	[132]
Q4WZB4	<i>Neosartorya fumigata</i>	[133]
C5NTD8	<i>Paenibacillus sp AIU311</i>	[120]
Q47NJ0	<i>Thermobifida fusca</i>	[134]
Q9ZBU1	<i>Streptomyces coelicolor</i>	[135]
A0LST6	<i>Acidothermus cellulolyticus</i>	[136]
Q7SIC2	<i>Aspergillus japonicus</i>	[137]
POCS93	<i>Gibberella zeae</i>	[138]

Tableau 12 : enzymes du set de références

Le taux d'identité entre chaque séquence a été vérifié. Un taux trop élevé diminue la diversité puisqu'il entraînerait des recoupements. Pour cela le programme en ligne gratuit Muscle (Multiple Sequence Comparaison by Log Expectation) a été utilisé dans le but de générer une matrice d'identité (Tableau 13).

Nom Uniprot	P93762	P04841	A5VWI7	F2QWIO	Q4WZB4	C5NTD8	Q47NJ0	Q9ZBU1	A0LST6	Q7SIC2	POCS93
P93762											
P04841	9										
A5VWI7	9	26									
F2QWIO	7	8	9								
Q4WZB4	11	8	8	10							
C5NTD8	13	9	10	9	9						
Q47NJ0	10	11	10	14	12	13					
Q9ZBU1	12	11	11	8	14	14	33				
A0LST6	12	13	13	11	14	13	28	49			
Q7SIC2	10	11	10	7	12	11	11	14	14		
POCS93	8	8	9	11	11	13	8	11	12	18	

Tableau 13 : matrice d'identité des enzymes du set de références avec en jaune clair les faibles taux d'identité et en orange les forts taux d'identité

Le pourcentage moyen d'identité du set de références est de 13 % avec une valeur maximale de 49%, ce qui prouve que les enzymes sont vraiment différentes. Un arbre d'identité a été créé (Figure 72). Cette représentation du tableau ci-dessus permet de voir plus rapidement que la diversité au sein de notre set de références est élevée puisque toutes les branches ne partent pas du même nœud.

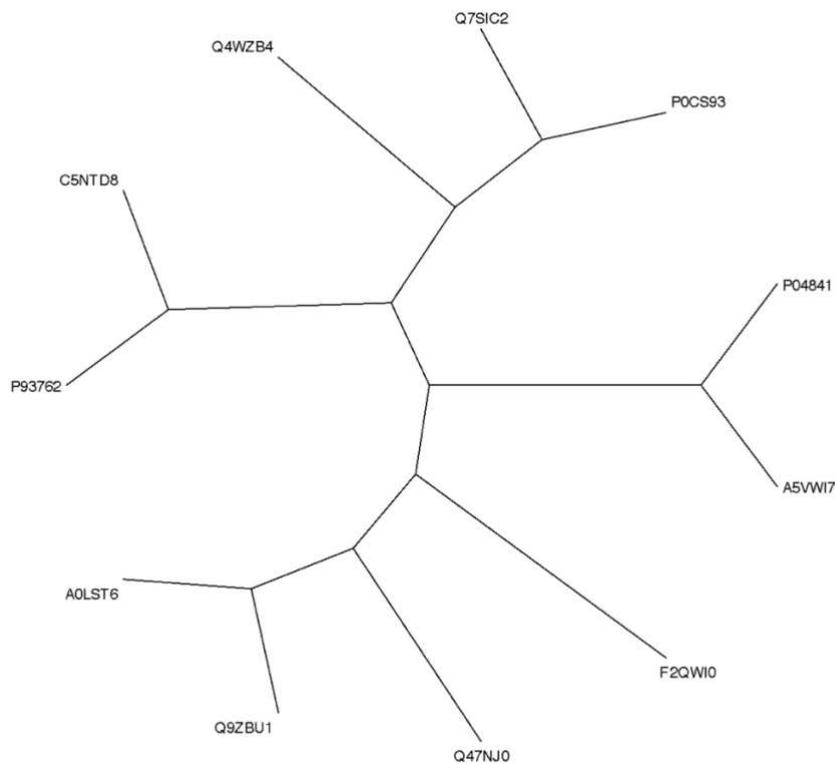


Figure 72 : arbre d'identité du set de références

2.3. Création informatique de la banque d'enzymes

Les séquences du set de références ont été alignées avec toute la base de données Uniprot (étape de Blast). Plus de deux cent mille séquences ont été générées en utilisant des paramètres « classiques », tels qu'avoir 30 % d'identité sur 80 % de la longueur de la séquence. Il faut noter que ce pourcentage d'identité est supérieur au taux d'identité moyen de la matrice ce qui va limiter les doublons. Cependant, une telle quantité de gènes à cloner n'était pas envisageable à l'échelle d'un laboratoire. Par mesure de précaution, un filtre a été appliqué pour supprimer les doublons. Ensuite l'étape de clustering a été réalisée afin de regrouper les séquences présentant une homologie de séquences d'au moins 50 %. Cette valeur a été sélectionnée pour limiter le nombre de clusters. Ainsi, quatre-vingt-cinq clusters ont été créés, avec quatre-vingt-quatre candidats représentatifs.

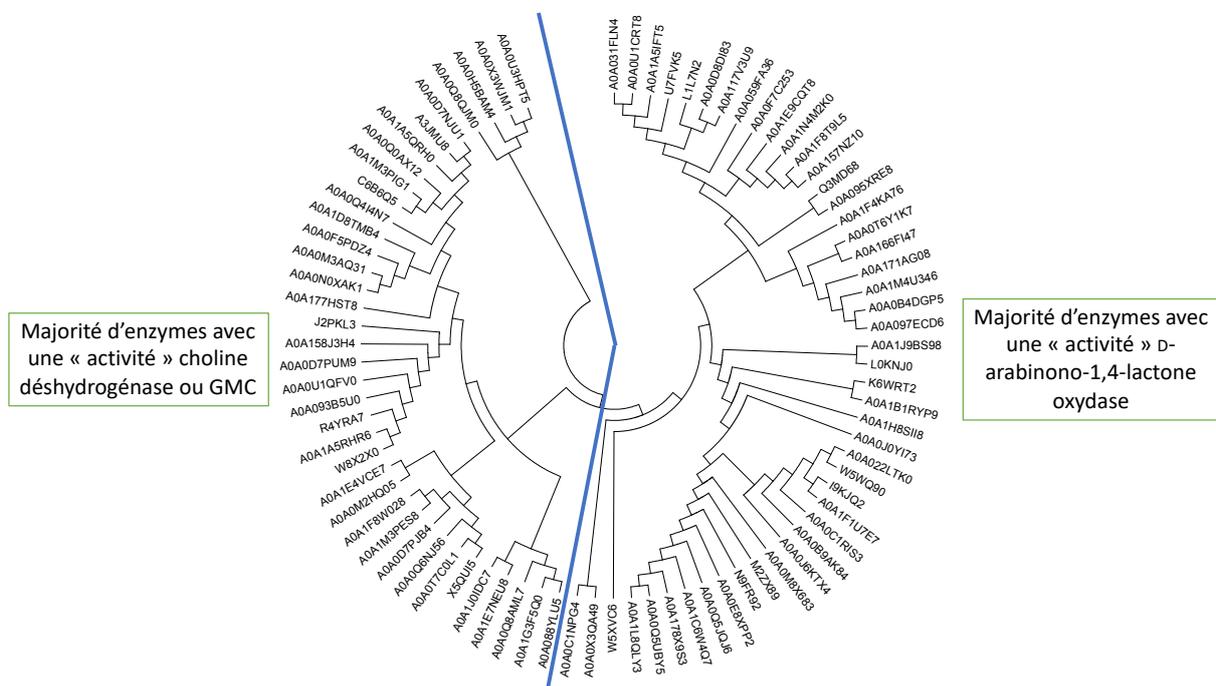


Figure 73 : arbre phylogénique des 85 représentants de clusters, GMC : glucose-méthanol-choline oxydoréductase

Un arbre phylogénique a été constitué à partir de l'ensemble des candidats sélectionnés. La quasi-totalité des enzymes sont annotées comme pouvant se lier au FAD. Deux groupes d'annotations ressortent (Figure 73). La partie gauche de l'arbre correspond à une majorité de choline déshydrogénases et de glucose-méthanol-choline oxydoréductases. La partie droite est représentée par une majorité de D-arabino-1,4-lactone oxydases. Des enzymes annotées super oxyde dismutases et ligases apparaissent dans les deux groupes, ce qui ne

correspond pas du tout à « l'activité » des autres enzymes de cet arbre. Ceci nous illustre parfaitement les limites d'annotations automatiques de la base de données Uniprot soulevées précédemment.

L'étape suivante est de rechercher les séquences génétiques de chaque enzyme ainsi qu'une société qui commercialise les microorganismes contenant les gènes sélectionnés. Seulement 45 microorganismes étaient commerciaux chez DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen) et ATCC (American Type Culture Collection). Les conditions de culture de ces bactéries et leur numéro sont référencés dans le Tableau 14.

Bactérie	Nom Uniprot	Nom DSMZ ou ATCC	Milieu	Température en °C	Gram
<i>Pseudonocardia autotrophica</i>	A0A1L8QLY3	DSMZ 43083	1	28	+
<i>Ruegeria intermedia</i>	A0A1M4U346	DSMZ 23285	2	28	-
<i>Mycobacterium abscessus</i>	A0A1N4M2K0	DSMZ 44196	3	28	+
<i>Anabaena variabilis</i>	Q3MD68	ATCC 29413	4	26	-
<i>Blastochloris viridis</i>	A0A0H5BAM4	DSMZ 133	5	25	-
<i>Streptomyces sp. F-3</i>	A0A171AG08	DSMZ 40783	1	28	+
<i>Microbacterium sp. Root166</i>	A0A0Q8AML7	DSMZ 102459	6	28	+
<i>Gordonia namibiensis</i>	K6WRT2	DSMZ 44568	7	28	+
<i>Pseudarthrobacter phenanthrenivorans</i>	A0A0B4DGP5	DSMZ 18606	6	28	+
<i>Microbacterium sp. Leaf151</i>	A0A0Q5JQJ6	DSMZ 102660	8	28	+
<i>Rhizobium tibeticum</i>	A0A1H8SII8	DSMZ 21102	9	28	-
<i>Kocuria palustris</i>	A0A0X3QA49	DSMZ 11925	6	30	+
<i>Micromonospora peucetia</i>	A0A1C6W4Q7	DSMZ 43363	1	28	+

<i>Brevibacterium linens</i>	A0A0B9AK84	DSMZ 20425	10	30	+
<i>Pseudomonas sp.</i> <i>RIT288</i>	A0A031FLN4	ATCC 17654	11	37	-
<i>Burkholderia ubonensis</i>	A0A1J9BS98	DSMZ 17311	12	30	-
<i>Thalassomonas viridans</i>	A0A0D8DI83	DSMZ 13753	2	28	-
<i>Streptomyces griseus</i> <i>subsp. griseus</i>	A0A0X3WJM1	DSMZ 40236	1	28	-
<i>Planococcus sp. L10.15</i>	A0A1B1RYP9	DSMZ 101994	2	15	+
<i>Hyphomonas jannaschiana</i> VP2	A0A059FA36	ATCC 33882	2	28	-
<i>Prauserella sp. Am3</i>	A0A0C1NPG4	DSMZ 43194	1	28	+
<i>Microbacterium hydrocarbonoxydans</i>	A0A0M2HQ05	DSMZ 16089	6	28	+
<i>Lysinibacillus sp. LK3</i>	A0A0J6KTX4	DSMZ 28	9	30	+
<i>Brevibacillus laterosporus</i>	A0A0F7C253	DSMZ 8787	9	30	+
<i>Acinetobacter lwoffii</i>	N9FR92	DSMZ 2403	12	30	-
<i>Corynebacterium falsenii</i>	W5WQ90	DSMZ 44353	7	37	+
<i>Mesorhizobium loti</i>	A0A1A5IFT5	DSMZ 2626	13	26	-
<i>Sphingobium sp.</i> <i>Leaf26</i>	A0A0Q4I4N7	DSMZ 102723	14	28	-
<i>Pseudomonas syringae</i> pv. <i>cilantro</i>	A0A0N0XAK1	DSMZ 10604	15	26	-
<i>Streptomyces jeddahensis</i>	A0A177HST8	DSMZ 101878	1	28	+
<i>Sphingomonas sp.</i> <i>Root710</i>	A0A0Q8QJM0	DSMZ 102538	14	28	-
<i>Mesorhizobium sp.</i>	A0A0Q6NJ56	DSMZ 102449	13	30	-

<i>Root102</i>					
<i>Pseudomonas sp.</i> <i>HMSC066B11</i>	A0A1E9CQT8	DSMZ 22644	7	30	-
<i>Achromobacter sp.</i>	A0A0U1QFV0	DSMZ 30026	9	30	-
<i>Oleispira antarctica</i> <i>RB-8</i>	R4YRA7	DSMZ 14852	2 + 1 % pyruvate	15	-
<i>Pseudomonas</i> <i>amygdali pv. tabaci</i>	A0A0Q0AX12	DSMZ 7298	9 + 1 % glucose	27	-
<i>Mesorhizobium loti</i>	A0A1A5QRH0	DSMZ 2626	13	26	-
<i>Corynebacterium casei</i>	W5XVC6	DSMZ 44701	7	30	+
<i>Streptomyces</i> <i>aureofaciens</i>	A0A1E7NEU8	DSMZ 40127	1	28	+
<i>Yersinia</i> <i>pseudotuberculosis</i>	A0A0E8XPP2	DSMZ 8992	16	37	-
<i>Burkholderia</i> <i>pseudomallei</i> <i>berberine</i>	A0A088YLU5	ATCC 700388	9	28	-
<i>Flavobacterium sp.</i> <i>Leaf359</i>	A0A0Q5UBY5	DSMZ 102605	14	28	-
<i>Amycolatopsis lurida</i> <i>NRRL 2430</i>	A0A093B5U0	DSMZ 43134	1	28	+
<i>Castellaniella</i> <i>defragrans 65Phen</i>	W8X2X0	DSMZ 12143	9	30	-
<i>Streptomyces</i> <i>ipomoeae 91-03</i>	L1L7N2	DSMZ 40383	1	28	+

Tableau 14 : référencement des bactéries dont les gènes vont être clonés

Au travers de ce tableau, nous pouvons observer que nous avons sélectionné des candidats très divers puisque les milieux de culture ainsi que les conditions de culture sont très différentes.

3. Création de la banque d'oxydases

3.1. Préparation des microorganismes et extraction de l'ADN

L'ensemble des étapes concernant la préparation et l'extraction de leur ADN sont illustrées sur la Figure 74.

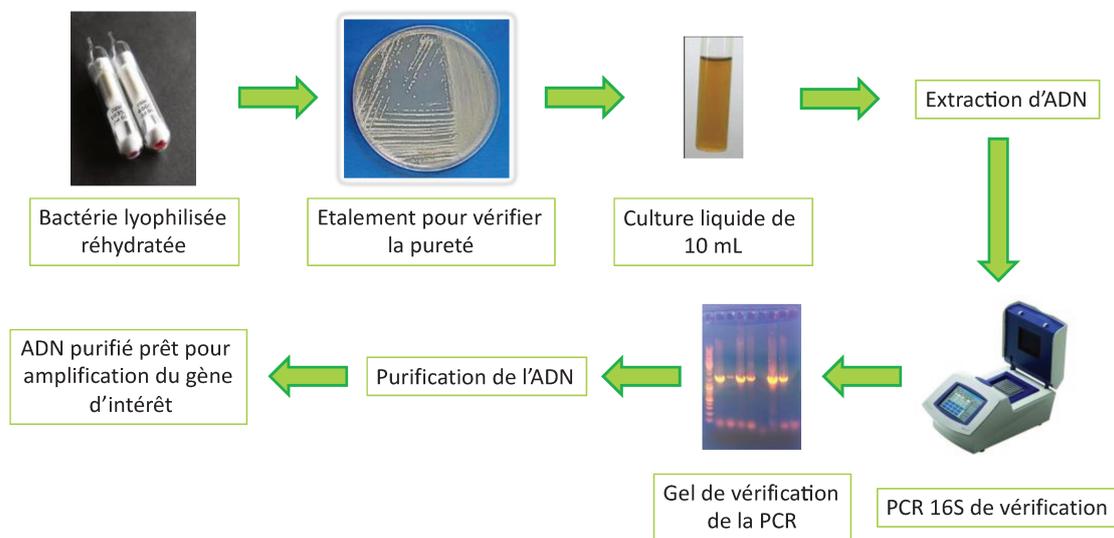


Figure 74 : procédure suivie depuis la réception des bactéries lyophilisées jusqu'à l'obtention des ADNs bactériens purs

Les bactéries ont été cultivées dans les conditions appropriées fournies par ATTC ou DSZM. Une fois réalisé, l'ADN est extrait, dosé et sa présence est vérifiée par PCR 16S. Une électrophorèse est réalisée pour vérifier la présence et la taille du produit amplifié. La PCR 16S consiste à utiliser des amorces universelles pour amplifier le gène codant pour la sous-unité 16S de l'ARN ribosomal. Ce gène est constitué d'une région conservée (amorces 16S) ainsi que d'autres portions de séquences propres à un groupe de bactéries (séquence et identification). Cependant, il faut noter que pour certaines bactéries la PCR 16S ne fonctionne pas pour des raisons de température d'hybridation ou de non spécificité. Dans notre cas seulement 2 ADNs n'ont pas pu être amplifiés grâce aux amorces 16S. Étonnement, les deux amplifications 16S qui n'ont pas fonctionné ne correspondent pas aux concentrations d'ADN les plus faibles dosées par absorbance UV (260 nm) au moyen d'un spectrophotomètre NanoDrop LITE. Les ADNs ont ensuite été purifiés grâce à un kit de purification. La quantité d'ADN est variable d'un microorganisme à l'autre (Tableau 15). Cependant, l'étape suivante de PCR nécessaire pour amplifier les gènes d'intérêt ne nécessite pas de grandes quantités d'ADN, quelques nano grammes suffisent.

Bactérie	Quantité d'ADN en ng.µL⁻¹	Réponse PCR 16S
<i>Pseudonocardia autotrophica</i>	169	Oui
<i>Ruegeria intermedia</i>	210	Oui
<i>Mycobacterium abscessus</i>	11	Oui
<i>Anabaena variabilis</i>	178	Non
<i>Blastochloris viridis</i>	2	Oui
<i>Streptomyces sp. F-3</i>	28	Oui
<i>Microbacterium sp. Root166</i>	243	Oui
<i>Gordonia namibiensis</i>	42	Oui
<i>Pseudarthrobacter phenanthrenivorans</i>	16	Non
<i>Microbacterium sp. Leaf151</i>	170	Oui
<i>Rhizobium tibeticum</i>	258	Oui
<i>Kocuria palustris</i>	105	Oui
<i>Micromonospora peucetia</i>	52	Oui
<i>Brevibacterium linens</i>	88	Oui
<i>Pseudomonas sp. RIT288</i>	748	Oui
<i>Burkholderia ubonensis</i>	1851	Oui
<i>Thalassomonas viridans</i>	35	Oui
<i>Streptomyces griseus subsp. griseus</i>	991	Oui

<i>Planococcus sp.</i> L10.15	53	Oui
<i>Hyphomonas jannaschiana</i> VP2	428	Oui
<i>Prauserella sp.</i> Am3	20	Oui
<i>Microbacterium hydrocarbonoxydans</i>	97	Oui
<i>Lysinibacillus sp.</i> LK3	191	Oui
<i>Brevibacillus laterosporus</i>	45	Oui
<i>Acinetobacter lwoffii</i>	821	Oui
<i>Corynebacterium falsenii</i>	28	Oui
<i>Mesorhizobium loti</i>	291	Oui
<i>Sphingobium sp.</i> Leaf26	422	Oui
<i>Pseudomonas syringae</i> pv. <i>cilantro</i>	84	Oui
<i>Streptomyces jeddahensis</i>	95	Oui
<i>Sphingomonas sp.</i> Root710	73	Oui
<i>Mesorhizobium sp.</i> Root102	171	Oui
<i>Pseudomonas sp.</i> HMSC066B11	887	Oui
<i>Achromobacter sp.</i>	688	Oui
<i>Oleispira antarctica</i> RB-8	86	Oui
<i>Pseudomonas amygdali</i> pv. <i>tabaci</i>	53	Oui

<i>Mesorhizobium loti</i>	299	Oui
<i>Corynebacterium casei</i>	75	Oui
<i>Streptomyces aureofaciens</i>	130	Oui
<i>Yersinia pseudotuberculosis</i>	201	Oui
<i>Burkholderia pseudomallei berberine</i>	1024	Oui
<i>Flavobacterium sp. Leaf359</i>	256	Oui
<i>Amycolatopsis lurida</i> NRRL 2430	88	Oui
<i>Castellaniella defragrans 65Phen</i>	1433	Oui
<i>Streptomyces ipomoeae 91-03</i>	10	Oui

Tableau 15 : quantité d'ADN extraite à partir des bactéries cultivées et corrélation avec la PCR 16S

A noter que l'amplification de tous les gènes a été testée par la suite même en l'absence de réponse lors de la PCR 16S puisque dans certain cas cette PCR ne fonctionne pas.

3.2. Amplification des gènes d'intérêt par PCR

Une fois les ADNs purifiés obtenus, des PCRs ont été réalisées, avec les amorces spécifiques aux gènes d'intérêt contenant les extrémités LIC (cohésives et complémentaires des extrémités du vecteur), dans différentes conditions afin d'amplifier les gènes codant pour les enzymes d'intérêt. Les amorces ont été définies en 5' à partir du codon codant pour la première méthionine de la protéine et en 3' jusqu'au codon codant pour l'arrêt de la traduction. Le logiciel Primer3 (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>) a été utilisé pour le design de ces amorces. Cependant, certaines amorces ont été créées manuellement car le logiciel n'indiquait pas de solution. Les amorces spécifiques obtenues ont été prolongées en position 5' par une séquence

spécifique pour permettre l'insertion du gène dans le vecteur. Aux amorces forward est ajoutée « l'extrémité LIC » AGAAGGAGATATAACTATG et aux amorces reverse est ajoutée « l'extrémité LIC » GGAGATGGGAAGTCATTA.

Plusieurs polymérases ont été testées et un gradient de températures entre 52 et 62°C a été réalisé pour optimiser les conditions de la PCR. Des électrophorèses en gel d'agarose ont permis de mettre en évidence la présence des produits PCR pour plus de la moitié des cibles (Tableau 16).

Bactérie	Amorce forward Amorce Reverse	Numéro d'accèsion Uniprot	Taille du gène en paires de bases
<i>Pseudonocardia autotrophica</i>	GTGACGGACCGGTGGCGCAACTGGG TCAGGTCGTCGACGTGCACGGCACTC	A0A1L8QLY3	1311
<i>Ruegeria intermedia</i>	ATGTGGCAGAAGTGGGCAGGCAACC TCATGGGCGAAACATCCGGGCCAGC	A0A1M4U346	1248
<i>Mycobacterium abscessus</i>	ATGCTGGTGCCTCCCGGTCCGAAG TCAGGGGCTGACCGGCCCGAGCAGC	A0A1N4M2K0	1239
<i>Anabaena variabilis</i>	ATGAAATCATCTCTGTGGCAACGG TTACGAATTACTTTGCCGTGAAGC	Q3MD68	813
<i>Blastochloris viridis</i>	ATGAACAGACGCGATACTCTGCAAAC TTAGGCGACCGCCTTGCGATAGTTGC	A0A0H5BAM4	717
<i>Streptomyces sp. F-3</i>	GTGACCAACTGGGCGGGCAACATCG TCACGGGGCCAGGAGGTCCGCGAGG	A0A171AG08	1233
<i>Microbacterium sp. Root166</i>	GTGACCCGCAACTGGGCGGGCACG TCAGCGCGGCTCGCGGGCCCCGATC	A0A0Q8AML7	1242
<i>Gordonia namibiensis</i>	GTGAGCGGGACGTCGACCACTGCCG CTAGTCGCCAGCACCCGCCGAGG	K6WRT2	1323
<i>Pseudarthrobacter phenanthrenivorans</i>	ATGAAGAACTGGGCAGGAAACCTCG TTAACGGCGGGCCGGCACGCCAGG	A0A0B4DGP5	1251
<i>Microbacterium sp.</i>	ATGGGAAGCGTGATCCACAGACGCG	A0A0Q5JQJ6	1257

<i>Leaf151</i>	CTACAGGCCTAGCGTCCGCCGAGG		
<i>Rhizobium tibeticum</i>	ATGACGCGAGTTCCAACGACCAATTG CTAGTTTTTCGCTCGTACCAAATATGAC	A0A1H8SII8	1254
<i>Kocuria palustris</i>	ATGACCGCAGAGGATGGACTGCCCTC TCAGAGGCCAGCAGGCCGCGGGCGG	A0A0X3QA49	1353
<i>Micromonospora peucetia</i>	GTGGGACAGTGGTTCGTGAACTGGTC TCACGCCTCCTCGAGCAGTCGGGAC	A0A1C6W4Q7	1218
<i>Brevibacterium linens</i>	GTGAGAGAAGTGCAGTCAGGGCAGG TCAGCGACCAGCTACACCGAACAGC	A0A0B9AK84	1359
<i>Pseudomonas sp. RIT288</i>	ATGAGCATGGATCTGACACGGCGTC TCAGCTCACCCCAGAATCGATGCC	A0A031FLN4	1395
<i>Burkholderia ubonensis</i>	ATGTGGCGCAACTGGTCCGGATACG CTATACACCGAACAGATCGCGCAAG	A0A1J9BS98	1287
<i>Thalassomonas viridans</i>	ATGAAATTTGAACGTCTTAACCTTAATC CTATGAGAGGCCAGGACGCTTTTAATATG	A0A0D8DI83	1434
<i>Streptomyces griseus subsp. griseus</i>	ATGGACGAGCAGCAGACGCAGGGCAG TCAGTCGCCCAGCACGCGCCGAGG	A0A0X3WJM1	1350
<i>Planococcus sp. L10.15</i>	ATGATGAAGGTGGACAAAGGAAAGAC TTAACGTGGCGGGCGACTAGCTCTTTC	A0A1B1RYP9	1332
<i>Hyphomonas jannaschiana VP2</i>	ATGAGCACGACACGCCGGGCCATAC TCAGGCATCGAAGTCTACTCCGAAC	A0A059FA36	1425
<i>Prauserella sp. Am3</i>	GTGACCNAACCGCCACTTTCCGCGC TCAGCGAAGGAGCGGGCGCAGGTG	A0A0C1NPG4	1263
<i>Microbacterium hydrocarbonoxydans</i>	GTGCAGGCGGCCGCTGCTCAGGGGC TCACGTCCCCAGGATGCGATCGAGG	A0A0M2HQ05	1212
<i>Lysinibacillus sp. LK3</i>	ATGTTTTCTATAGAAAAATGGCGAAATG TTACAATAGAAACAACCTTTTTTAAG	A0A0J6KTX4	1323
<i>Brevibacillus laterosporus</i>	ATGAGAACAGTAGCAATGCAACCATG CTATACTCCAAACAGCCGTGACAAG	A0A0F7C253	1308

<i>Acinetobacter lwoffii</i>	ATGAAAATACAACAGGCGAATGCAG CTATTCAGACAAAAATAACTGCTG	N9FR92	1284
<i>Corynebacterium falsenii</i>	ATGAGCAGTCTGAAAAGCACACAG TTAGGCCAAGCCGAACACCCGGTCC	W5WQ90	1335
<i>Mesorhizobium loti</i>	ATGCCGCACGCGGAAAGTTACGATTAC TCATTGCACTGCCCTGCTTTCTTTC	A0A1A5IFT5	1632
<i>Sphingobium sp. Leaf26</i>	ATGAGTGAATGACTATATCGTGATC TCAGGCGGCGATATTGGATCGAAGG	A0A0Q4I4N7	1626
<i>Pseudomonas syringae pv. cilantro</i>	ATGACATACGACTACATCATCGCTG TCACGCATCCTGCAAATAAGATC	A0A0NOXAK1	1608
<i>Streptomyces jeddahensis</i>	ATGGCTCCCCTGCAATACGACTTCG CTACCCCGCCTGCTTCTGGGCGTC	A0A177HST8	1671
<i>Sphingomonas sp. Root710</i>	ATGCGATCAAGGCGGATATCCAGAC TCAGGCGCTGAGGACGCGGCCGAGG	A0A0Q8QJM0	1701
<i>Mesorhizobium sp. Root102</i>	ATGCCCGACTTTATCGTTGTAGGCG TCAAATCCGATTTCCCTTCAGAAAATC	A0A0Q6NJ56	1587
<i>Pseudomonas sp. HMSC066B11</i>	ATGACTTCCACTGCCC GCCGTTTCGTTT TCAGGCGCCG GCCAGAGCCTGCACGG	A0A1E9CQT8	1680
<i>Achromobacter sp.</i>	ATGACCCAGGCAAAGTACGAATTCATC TTAGTTTCCCGCCTGGATCAGGCG	A0A0U1QFV0	1611
<i>Oleispira antarctica RB-8</i>	ATGACAATTGAATATGACTATATC TCAATAGTAATAAGACTTACCCAG	R4YRA7	1614
<i>Pseudomonas amygdali pv. tabaci</i>	ATGAGCGATTATGACTACATCATTG TCAGGCTTGTTTCATCGATAGTCTTC	A0A0Q0AX12	2085
<i>Mesorhizobium loti (2)</i>	ATGCAGACCTATGACTTCATTATCG TCATTTGTGGTCTCCAGAACCAG	A0A1A5QRH0	1617
<i>Corynebacterium casei</i>	ATGAACACCACCACGATTGAAACGC TTAGGCGGGAAGTTCCACGCCGAGG	W5XVC6	1701
<i>Streptomyces</i>	ATGCGACACGAATTCTCAGGGCGG	A0A1E7NEU8	1521

<i>aureofaciens</i>	TCAGGCCCGGAGCGGCCGGATCGAC		
<i>Yersinia pseudotuberculosis</i>	ATGAAGATTATTGATAAAAATGTC TTAACGACCCGCGAGTATACCCCACC	A0A0E8XPP2	2058
<i>Burkholderia pseudomallei berberine</i>	ATGTCCTATGTGTACGTCGACGGTG TCAACCAAGCGGCGGCGTCAGCGGAATC	A0A088YLU5	1662
<i>Flavobacterium sp. Leaf359</i>	ATGGGCAATCCTGACGATGTAGCAG TTATGCTTTAGGAGCAGAAGCAGC	A0A0Q5UBY5	663
<i>Amycolatopsis lurida NRRL 2430</i>	GTGGCCGATCAGGACTCCTTCGAC TCAGCCTCGGCCCGGATGAGGTC	A0A093B5U0	1584
<i>Castellaniella defragrans 65Phen</i>	ATGCAGGAGGAGTTGACTACGTCATC TCACCCGGCGCCGTCCGCCAGGATC	W8X2X0	1605
<i>Streptomyces ipomoeae 91-03</i>	ATGACCGACGCCGACCGACGGCACC CTAATCCCGGGCCGGCTCCACGGAC	L1L7N2	1608

Tableau 16 : récapitulatif des gènes amplifiés et caractéristiques d'amplification, avec en premier l'amorce forward et en second l'amorce reverse. Aux amorces forward est rajoutée l'extrémité LIC AGAAGGAGATATAACTATG et aux amorces reverse est rajoutée l'extrémité LIC GGAGATGGGAAGTCATTA

Plus de la moitié des gènes ciblés ont été amplifiés avec succès (27/45). Tous les gènes inférieurs à 1000 bp sont amplifiés. Ce sont principalement les gènes de plus de 1000 bp qui posent problème. Cependant, ce rendement de seulement 60 % d'amplifications réussies ne s'explique pas uniquement par la longueur des gènes. Certains gènes présentent une composition nucléotidique avec un taux de GC supérieurs à 75 %^[139-141]. Un taux de GC élevé affecte négativement la capacité d'un gène à être amplifié en raison d'une stabilité accrue (3 interactions hydrogènes au lieu de 2). Du DMSO (5, 10 et 15 %) a été ajouté lors de nos PCR. Ce solvant permet de déstabiliser l'ADN lors de l'étape de dénaturation en se liant aux bases cytosines. Cependant, aucun résultat positif n'a été obtenu grâce à cette méthode. D'autres méthodes existent, telles que augmenter le taux de chlorure de magnésium, ajouter de la bêtaïne^[142] ou bien utiliser des kits spécifiques mais nous n'avons pas pu les tester par manque de temps. Une fois les produits amplifiés, ils sont purifiés grâce à un kit spécifique de purification de produits PCR.

3.3. Clonage des gènes amplifiés

Le Tableau 17 présente les 27 enzymes dont les gènes codants ont été amplifiés et purifiés.

Numéro	Nom Uniprot	Taille du gène en paires de bases
1	A0A1M4U346	1248
2	A0A1N4M2K0	1239
3	A0A1H8SII8	1254
4	A0A1B1RYP9	1332
5	A0A059FA36	1425
6	A0A0C1NPG4	1263
7	A0A0Q4I4N7	1626
8	A0A0Q6NJ56	1587
9	W5XVC6	1701
10	A0A093B5U0	1584
11	A0A0Q5UBY5	663
12	A0A0E8XPP2	2058
13	A0A1E9CQT8	1680
14	A0A0B4DGP5	1251
15	A0A0B9AK84	1359
16	A0A1A5QRH0	1617
17	L1L7N2	1608
18	A0A0Q8QJM0	1701
19	A0A0M2HQ05	1212
20	A0A0F7C253	1308
21	W5WQ90	1335

22	A0A0Q8AML7	1242
23	A0A0H5BAM4	717
24	K6WRT2	1323
25	A0A0X3QA49	1353
26	A0A177HST8	1671
27	A0A088YLU5	1662

Tableau 17 : gènes amplifiés avec succès

Le clonage des (Open Reading Frame) ORF a été réalisé par la méthode LIC^[143] (Ligation Independent Cloning). Cette méthode ne nécessite pas d'étape de digestion enzymatique suivie d'une étape de ligation par rapport à d'autres méthodes comme Gloden Gate Assembly ou Gateway Cloning^[144,145]. Cette méthode utilise l'activité 3' exonucléasique de la T4 Polymérase pour créer des extrémités complémentaires entre le vecteur et l'insert. La Figure 75 montre le déroulement des étapes du clonage.

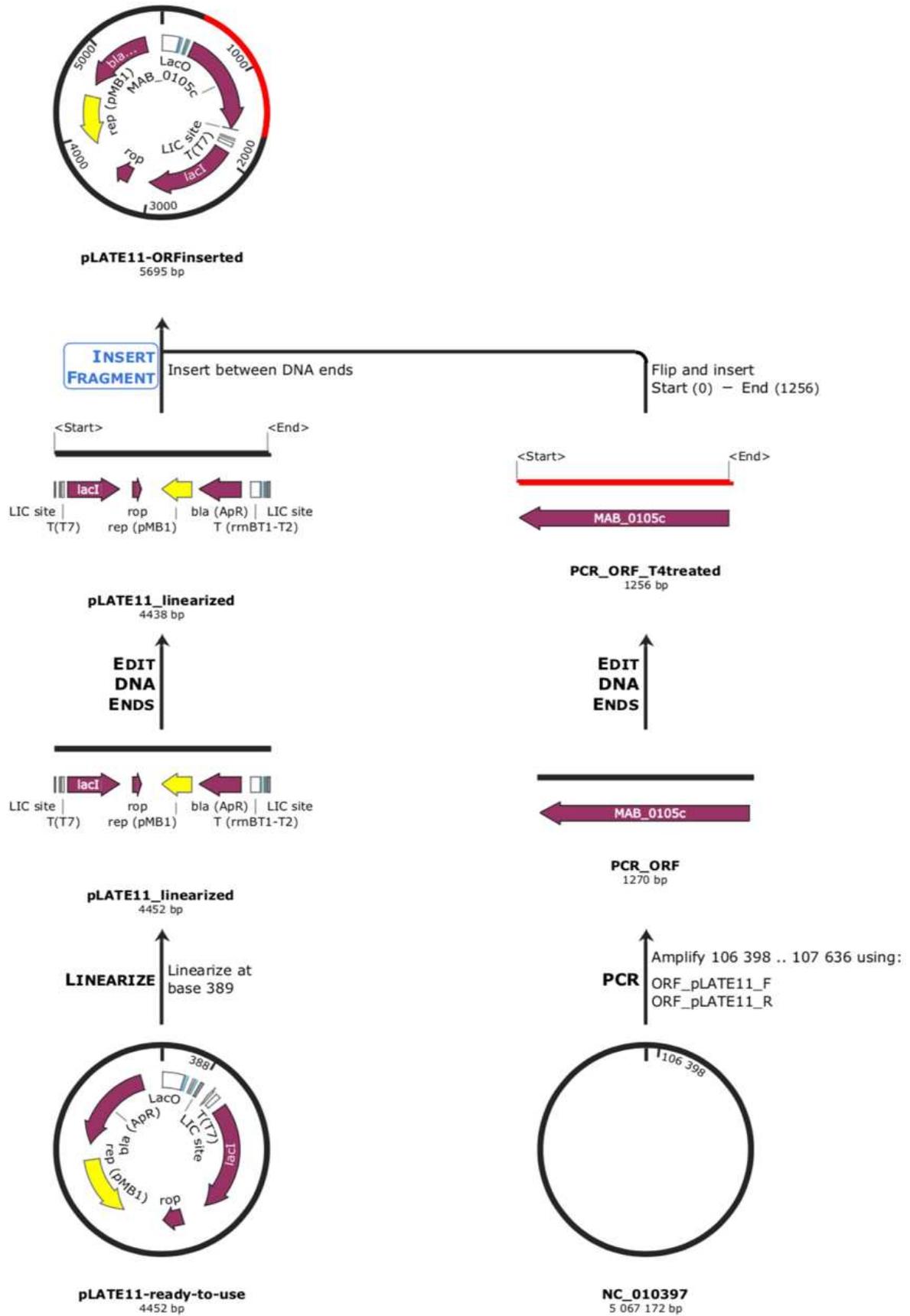


Figure 75 : principe du clonage d'un ORF dans le vecteur d'expression pLATE11 selon la méthode LIC

Du dGTP est incorporé dans le milieu réactionnel pour stopper le processus exonucléase à la première cytosine rencontrée sur l'insert au niveau de l'extension 5' des amorces. Le vecteur est soit acheté déjà linéarisé et prêt à intégrer l'insert, soit obtenu à partir d'un plasmide préparé au laboratoire dont la linéarisation est effectuée par l'enzyme de restriction Bsal. Les extrémités du vecteur sont rendues cohésives et complémentaires des extrémités de l'insert grâce à l'activité exonucléasique de la T4 polymérase et l'incorporation de dCTP. La simple complémentarité des extrémités cohésives du vecteur et de l'insert suffit à circulariser la construction qui est alors directement utilisable pour la transformation bactérienne.

3.4. Transformation et expression des protéines

3.4.1. Obtention de colonies transformées non productrices

Dans un premier temps le produit issu du clonage a été directement utilisé pour transformer les cellules compétentes *E.coli* BL21 DE3 pLys selon les instructions du fournisseur. Ces cellules sont aptes à la production des enzymes. Cependant, tous les tests réalisés ont montré que dans ces conditions très peu voire aucune colonie n'est présente sur boîte de pétri. Dès lors, l'obtention de colonies permettant la production de protéines recombinantes a été réalisée en deux étapes. Un premier sous clonage a été réalisé dans les cellules compétentes *E.coli* JM109 (Promega), qui elles ne sont pas productrices de protéines, pour produire et purifier le plasmide. Des colonies PCR avec les amorces spécifiques du vecteur LIC ont été réalisées pour vérifier la présence de l'insert, puis séquencé pour vérifier la construction. 13 gènes ont été clonés avec succès, un gène a été tronqué de la séquence codant des 20 acides aminés en position N terminale. Nous n'avons pas d'explications pour les gènes que nous n'avons pas réussi à insérer dans le plasmide. Ce n'est pas un problème de taille de gènes puisque des gènes de toutes tailles (de 700 bp à 1700 bp) ont été clonés avec succès. Peut-être un problème de quantités de produits PCR lors de la « ligation » ou lors de la transformation des cellules compétentes.

3.4.2. Obtention de colonies transformées productrices de protéines recombinantes

Dans un second temps, le plasmide purifié a été utilisé pour transformer les bactéries *E.coli* BL21 DE3 pLys capables de surexprimer les gènes d'intérêt et de produire les enzymes correspondantes (Figure 76).

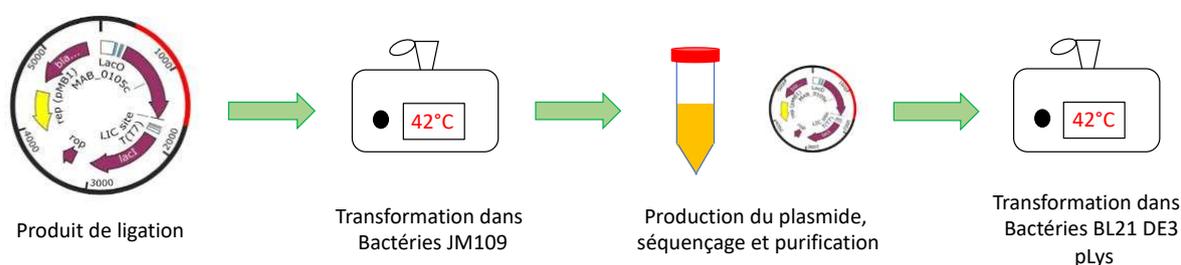


Figure 76 : étapes du processus de transformation

Parmi les 27 constructions plasmidiques réalisées, 14 constructions n'ont pas permis d'obtenir des colonies transformées malgré plusieurs essais. Les 13 transformants obtenus sont ensuite mis en production.

3.4.3. Surexpression et production des protéines recombinantes

La production des 13 (2/3/5/6/11/14/16/19/20/21/23/24/26 tableau 17) enzymes a été réalisée grâce à une induction par IPTG (Figure 50). Différentes conditions de culture et de production de protéines ont été testées à partir des préconisations du fabricant. D'autres tests ont été effectués en faisant varier la température d'incubation, la vitesse d'agitation ou la concentration de l'inducteur IPTG. Cependant aucune surexpression n'a été détectée.

Les séquences nucléotidiques incorporées dans le vecteur d'expression proviennent de microorganismes éloignés de *E.coli* (Tableau 14). Un biais d'usage de codon peut-être la raison de l'absence de surexpression. La traduction a pu être limitée ou rendue impossible par la présence, dans la séquence codante, de codons dont les ARNt sont faiblement présents ou absents chez *E.coli*^[146]. Cette hypothèse a été vérifiée par l'utilisation du programme informatique en ligne ATGme^[147] (<http://atgme.org/>) permettant la détection de codons rares d'*E.coli* dans une séquence nucléotidique. La présence de codons rares dans les séquences que nous cherchons à exprimer a été montrée par ces analyses. Il n'est pas possible d'affirmer avec certitude que les enzymes n'ont pas été produites. Nous pouvons

juste affirmer qu'elles n'ont pas été surexprimées. La production aurait pu être vérifiée en découpant des bandes sur notre gel SDS-Page et en réalisant une analyse en Nano-LC-MS (nanoélectrospray ionization-liquid chromatography-mass spectrometry) de ces bandes. Cependant, comme nous voulons utiliser l'oxydase pour réaliser de la synthèse il nous en faut en grande quantité. Nous avons donc décidé de ne pas nous attarder sur cette méthode qui de plus aurait dû être réalisée par un autre laboratoire.

Dès lors, plusieurs alternatives pour permettre la traduction des ARNm correspondants aux enzymes à produire ont été envisagées : notamment l'optimisation des séquences afin d'éliminer la présence de ces codons rares, ou l'utilisation de la souche *d'E.coli* ROSETTA 2 (DE3) (Novagen) dans laquelle le plasmide pRARE comportant les gènes codants les ARNt rares *d'E.coli* a été introduit^[148]. Pour des raisons évidentes de temps et de nombres de codons à optimiser, la première alternative n'a pas été envisagée. Des nouvelles transformations en utilisant la souche *d'E.coli* ROSETTA 2 (DE3) ont donc été effectuées.

De nombreuses conditions de production de nos 13 potentielles oxydases ont été testées :

- variation de la quantité d'IPTG lors de l'induction de la production (de 0,5 à 5 mM)
- variation de l'agitation (150, 200, 250 rpm)
- variation de la température après induction de la production (30°C et 37°C)

Seulement 3 enzymes ont été correctement exprimées comme nous le montre la Figure 77, les enzymes numéros 11, 16 et 26 qui possèdent toutes les trois des tailles différentes.

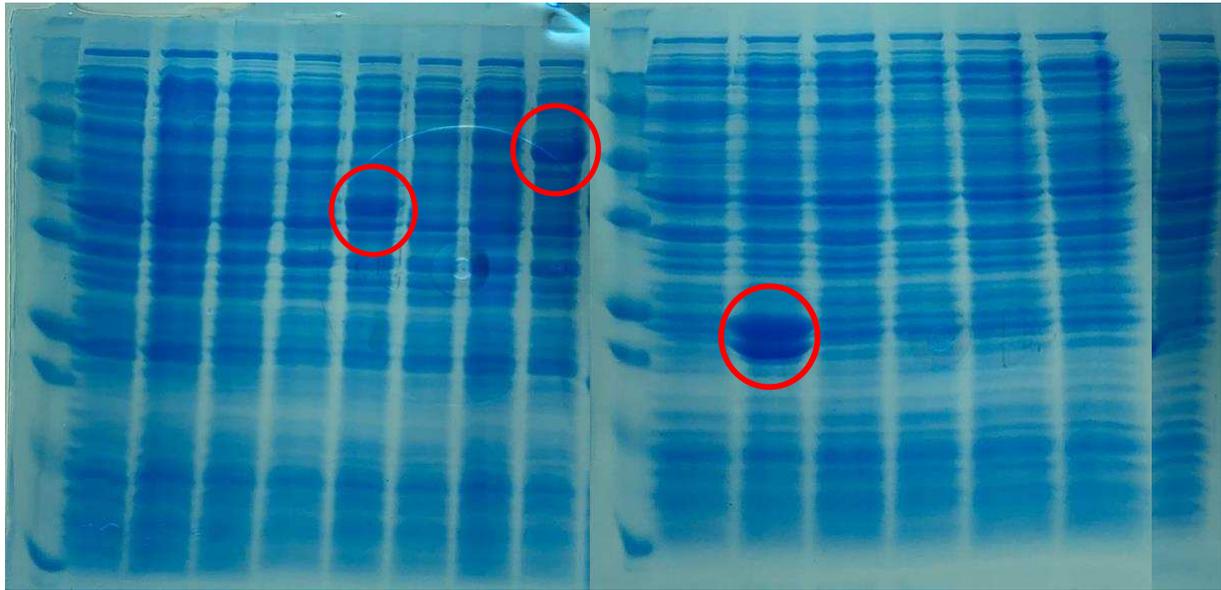


Figure 77 : gel SDS-page pour la vérification de la production des protéines avec, en partant de la gauche, le marqueur de taille et les lysats de production des enzymes 2/5/6/14/21/23/24/26, un deuxième marqueur de taille puis les lysats de production des enzymes 3/11/16/19/20/témoins non induits. Entourées en rouge les bandes correspondantes aux enzymes produites.

A l'heure actuelle, nous ne connaissons pas la raison pour laquelle les autres oxydases ne sont pas produites. Le problème peut venir de la transcription. Cependant, les gènes codants les ARNt rares sont présents dans la bactérie productrice. Une dégradation des enzymes ou des ARN peut aussi être envisagée. Enfin, les conditions de culture peuvent aussi être enzyme dépendante puisque dans notre cas, les 3 enzymes ne sont produites que lorsque l'agitation du milieu de culture est fixée à 250 rpm.

Ces trois enzymes produites ont été criblées en dosant l' H_2O_2 libéré lors de la réaction d'oxydation du glycérol en glycéraldéhyde (Figure 78).

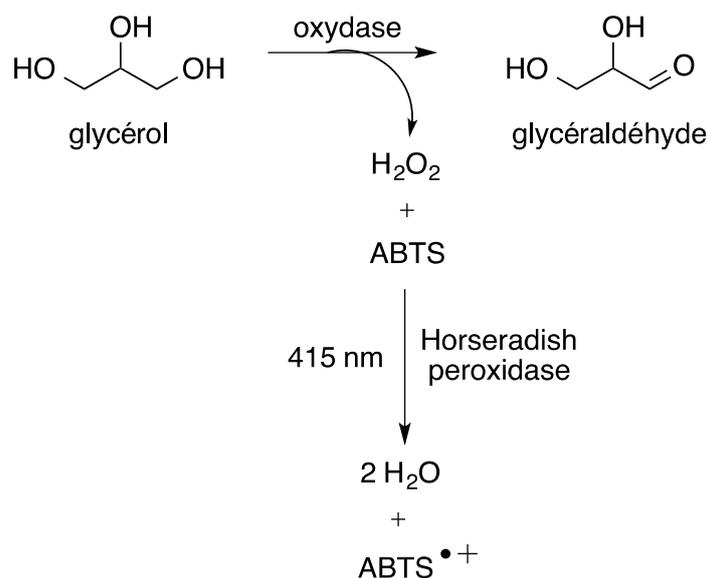


Figure 78 : criblage des oxydases

Pour cela, la peroxydase de raifort (Horseradish Peroxidase ou HRP) ainsi que l'acide 2,2'-azino-bis(3-éthylbenzothiazoline-6-sulfonique (ABTS) ont été utilisés. Cette peroxydase utilise une molécule de H_2O_2 produite lors de la réaction d'oxydation du glycérol pour oxyder l'ABTS ce qui permet d'observer une cinétique d'oxydation à 415 nm. Aucune des trois oxydases testées ne montre d'activité. Cependant la pertinence de ce test sur des lysats laisse à désirer car les bactéries produisent naturellement des catalases qui dismutent le H_2O_2 . Une étude RMN a été réalisée pour observer la formation de glycéraldéhyde mais les données obtenues ne sont pas interprétables en raison de la faible résolution de la RMN 400 MHz.

D'autres études pour la production des 10 autres enzymes doivent être menées. Il est possible que ces enzymes soient toxiques pour la bactérie et que celle-ci les dégrade au cours du temps. Cependant, l'hypothèse la plus plausible reste un problème au niveau de la traduction ou bien des conditions de production. Il serait intéressant aussi de tester les 3 oxydases surexprimée sur différents substrats pour définir leur spectre d'activité (glycolaldéhyde ou polyols).

Conclusion Générale et perspectives

Dans un contexte de développement de la bioraffinerie oléagineuse, mon travail de thèse avait pour objectif de transformer le glycérol par catalyse enzymatique en une molécule plate-forme, le glyoxal. La voie de synthèse choisie consistait tout d'abord à oxyder le glycérol en glycéraldéhyde, puis cliver le glycéraldéhyde en glycolaldéhyde et enfin oxyder le glycolaldéhyde obtenu en glyoxal (Figure 79).

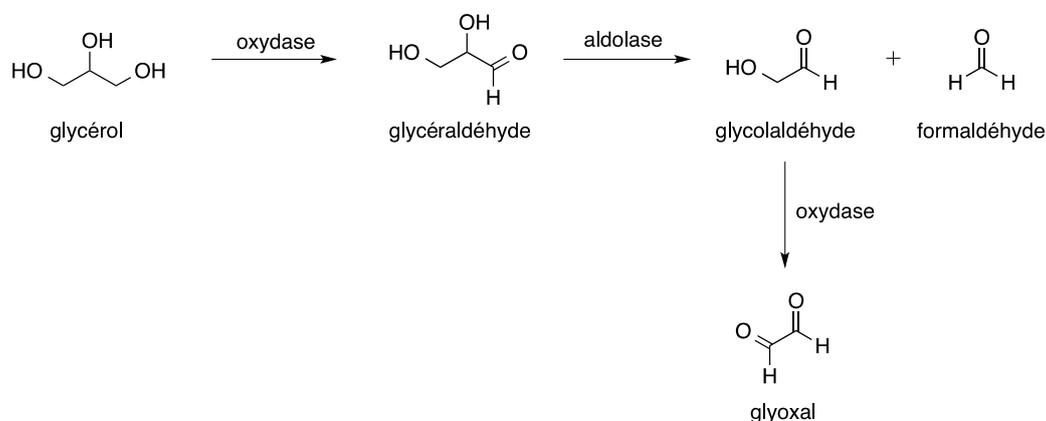


Figure 79 : Voie envisagée pour la valorisation du glycérol par catalyse enzymatique

Une recherche bibliographique a été effectuée pour trouver une aldolase capable de cliver le glycéraldéhyde. La fructose-6-phosphate aldolase a été sélectionnée comme point de départ puisque c'est l'une des seules enzymes pouvant catalyser des réactions d'aldolisation entre deux aldéhydes. En ce qui concerne la recherche de l'oxydase, nous avons adopté une méthodologie de recherche de nouvelles activités enzymatiques dans la biodiversité en utilisant certains outils de la bio-informatique. C'est la première fois que de telles approches sont développées au sein de l'Institut Charles Violette.

La fructose-6-phosphate aldolase sous sa forme native ne catalyse pas la réaction de clivage du glycéraldéhyde. Une collaboration avec l'Institut de Chimie de Clermont-Ferrand a été réalisée dans le but de mettre au point un test de criblage et de tester les mutants de cette enzyme en leur possession. Un mutant de la FSA, le L107Y A129G, a été révélé comme catalysant la réaction de clivage uniquement du L-glycéraldéhyde. Cependant, une inhibition de cette enzyme a été mise en évidence en présence du D-glycéraldéhyde. Deux approches ont alors été développées pour lever cette inhibition. Une première approche a été de créer une nouvelle mutation dans le site actif du mutant FSA L107Y A129G en s'appuyant au préalable sur une étude de docking pour déterminer l'acide aminé à

muter. Une seconde approche bio-informatique, a été développée pour rechercher une oxydase dans la biodiversité capable de ne produire spécifiquement que le L-glycéraldéhyde.

Concernant la première approche, le fichier .PDB de la FSA L107Y A129G a d'abord été créé en utilisant le logiciel SCWRL4 avec la collaboration du professeur Frédéric AFFOUARD et de la maîtresse de conférences Natalia CORREIRA de l'UFR de Physique de l'Université de Lille, membre de l'Unité Matériaux et Transformations (UMET). Ensuite, la nouvelle structure obtenue a été minimisée à l'aide du logiciel NAMED dans le but d'obtenir la conformation du site actif de plus basse énergie. Enfin une étude de dynamique moléculaire, permettant des calculs plus long mais plus précis, dans une matrice de molécules d'eau a été effectuée afin d'augmenter au maximum la pertinence du modèle créé. Une étude de docking du L-glycéraldéhyde et du D-glycéraldéhyde dans le site actif de la FSA L107Y A129G a été réalisée pour mettre en évidence les interactions entre les deux molécules et les acides aminés du site actif de l'enzyme. Des interactions plus nombreuses sont été observées avec le D-glycéraldéhyde : les deux acides aminés Y131 et T26 ont été mis en évidence comme interagissant avec le D-glycéraldéhyde mais pas avec le L-glycéraldéhyde, ce qui pourrait expliquer le phénomène d'inhibition de la FSA L107Y A129G par le D-glycéraldéhyde. Par conséquent, nous avons étudié l'influence de la mutation de ces deux positions sur l'activité enzymatique de la FSA L107Y A129G en présence de D-glycéraldéhyde. Les deux mutations sont été réalisées par mutagenèse dirigée et les deux acides aminés ont été remplacés par une glycine. La mutation de la thréonine position 26 en glycine a entraîné la perte de l'activité de clivage du L-glycéraldéhyde de l'enzyme à première vue. Cependant, il est difficile de conclure sur ce triple mutant T26G L107Y A129G puisque la FSA catalyse préférentiellement les réactions d'oxydations. Il est possible que la FSA clive le glycéraldéhyde mais catalyse instantanément l'aldolisation du glycolaldéhyde et sur le formaldéhyde et non la synthèse du xylose.

La mutation de la tyrosine position 131 en glycine a révélé une activité de clivage du L-glycéraldéhyde mais pas du D-glycéraldéhyde. Nous pouvons donc conclure sur ce triple mutant L107Y A129G Y131G que l'inhibition de l'enzyme n'est pas levée en présence de D-glycéraldéhyde. Ce résultat est tout de même très intéressant puisque jusqu'à maintenant la tyrosine en position 131 a toujours été considérée comme indispensable à l'activité catalytique de la FSA. Il est possible que la proximité spatiale de la tyrosine 107 compense

l'action de la tyrosine 131. Il est possible que la réalisation des trois mutations L107Y A129G Y131G ait une synergie et modifie la réactivité du site actif de l'enzyme. Pour valider cette théorie, il faudrait muter un par un les acides aminés en position 107 et 129 en leur acide aminé natif (Y107L et G129A). Cela permettrait de mettre en évidence une éventuelle synergie entre les trois mutations et leur influence sur la réaction de clivage ou d'aldolisation catalysée par la FSA. La réalisation d'une étude de dynamique moléculaire pourrait aider à comprendre le déplacement du glycéraldéhyde dans le site actif et apporterait certainement des pistes supplémentaires sur l'inhibition de l'enzyme. L'étude de la position 6 (acide aspartique pur la FSA_{wt}) doit être réalisée puisque cet acide aminé reste relativement énigmatique dans la bibliographie. Etant décrit comme responsable de l'énantio-préférence de la FSA il serait intéressant d'étudier différentes mutations en cette position afin de lever l'inhibition.

La seconde approche étudiée a été de trouver une oxydase catalysant spécifiquement l'oxydation du glycérol en L-glycéraldéhyde. Cette réaction d'oxydation est très peu décrite dans la littérature et fait, la plupart du temps, intervenir des enzymes fongiques, dont le processus de production est plus complexe et long. Une approche bio-informatique a alors été développée pour rechercher cette oxydase dans la biodiversité. La séquence protéique primaire en acides aminés est responsable de la structure tertiaire d'une protéine et donc de son activité. Si des enzymes ont des séquences primaires proches, il est plausible qu'elles catalysent les mêmes réactions. C'est en se basant sur ce postulat que nous avons créé un programme bio-informatique permettant d'automatiser toutes les étapes en partenariat avec la plateforme bio-informatique de Lille « Bilille ». Un travail de recherche bibliographique d'oxydases catalysant l'oxydation d'alcools primaires à courtes chaînes en aldéhydes a été effectué. 11 enzymes ont été retenues pour la création d'un set de références. Un blast de ce set de références sur toute la base de données Uniprot a été réalisé pour générer un grand nombre d'enzymes possédant une structure primaire proche de celles du set de références. Enfin des clusters d'identité ont été créés à partir de ces données et un candidat représentatif de chaque cluster a été sélectionné. Sur les 85 candidats représentatifs, seules 45 enzymes provenaient de bactéries commerciales. Les 45 bactéries ont été achetées et les gènes codant pour les 45 enzymes retrouvées sur NCBI (National center for Biotechnology Information). Les microorganismes ont été produits dans

leur milieu de culture et les ADN ont été extraits et purifiés. L'amplification des gènes d'intérêt a été réalisée par PCR en utilisant des amorces spécifiques contenant des extrémités LIC. 27 gènes ont été amplifiés avec succès et 13 ont été clonés dans le vecteur pLATE11. Les cellules *E.coli* ROSETTA 2 (DE3) possédant un plasmide pRARE2 comportant les gènes codant pour les ARNt rares de *E.coli* ont été transformées et utilisées pour la production des enzymes. 3 enzymes ont été surexprimées et criblées par dosage du H₂O₂ formé lors de la réaction d'oxydation du glycérol (Figure 80) et par RMN du proton. Cependant, aucune de ces trois enzymes ne présentent avec certitude une activité d'oxydation du glycérol en L-glycéraldéhyde.

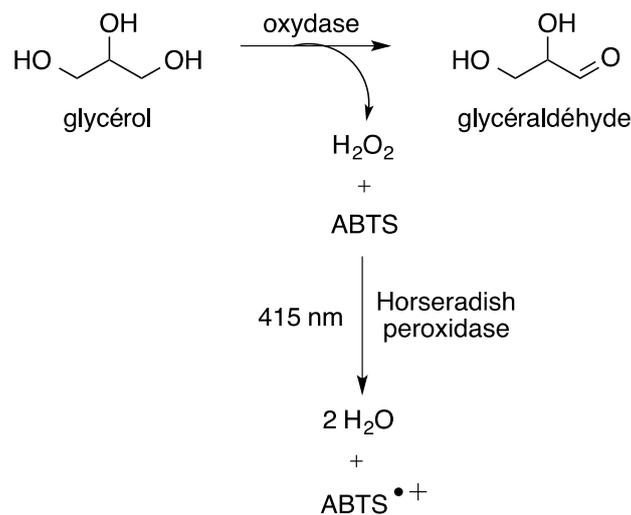


Figure 80 : criblage des oxydases

D'autres études pour la surexpression des 10 autres enzymes doivent être menées. Il est possible que ces enzymes soient toxiques pour la bactérie et que celle-ci les dégrade au cours du temps. Cependant, l'hypothèse la plus plausible reste un problème au niveau de la traduction ou bien des conditions de production. Il est aussi envisagé de réaliser des études de Nano-LC-MS (nanoélectrospray ionization-liquid chromatography-mass spectrometry) afin de vérifier si les enzymes ont été produites mais pas surexprimées. Enfin, il faut aussi essayer d'autres substrats avec nos 3 oxydases surexprimées afin de définir leur spectre d'activité.

Ce travail a permis d'amener les prémices d'une nouvelle approche pour la recherche de nouvelles enzymes au sein de l'Institut Charles Viollette. En effet, la création du programme bio-informatique et la mise au point des méthodes de biologie moléculaire manuellement vont permettre à la plateforme de criblage haut débit REALCAT de

développer des projets similaires et ainsi contribuer à la découverte de nouvelles activités enzymatiques.

Matériels et Méthodes

1. Généralités

1.1. Réactifs

Les Réactifs chimiques utilisés ont été commandés chez la société Sigma Aldrich, Sigma-Fluka, Sigma Aldrich Co. (St. Louis, MO). Les kits de biologie moléculaire sont fournis par la société Thermofischer, Promega et Agilent, les produits pour la préparation des gels SDS-PAGE par la société BIORAD.

1.2. Appareils pour analyses

Les mesures d'absorbance UV-vis en macro-cuve sont effectuées avec un spectrophotomètre SAFAS UVmc. Les mesures en micro-plaque 96 puits sont réalisées sur un lecteur de micro-plaque BioTek ELx808. Les PCRs sont effectuées sur une machine Thermofisher Scientific modèle ProFlex 3x32 puits. Les mesures de concentrations d'ADN sont faites avec le NanoDrop Lite de chez ThermoFischer scientifique. Les spectres RMN sont réalisés sur des appareils Bruker à 300 Mhz et 400 Mhz.

2. Protocoles standards pour les cultures bactériennes

2.1. Protocole pour la culture des souches bactériennes

A 10 mL de milieu approprié contenant la concentration adéquat d'antibiotique (ampicilline $0,1 \text{ mg.mL}^{-1}$ ou chloramphénicol $34 \text{ }\mu\text{g.mL}^{-1}$) est ajouté 1 μL de cryotube contenant la bactérie dans des conditions stériles. Le mélange est ensuite incubé à 37°C sous une agitation de 250 rpm pendant une nuit.

En cas de production de protéine recombinante, une préculture est effectuée dans les conditions ci-dessus sur la nuit. La préculture est ensuite diluée 50 fois dans un milieu frais identique à celui de la préculture. Au bout d'1h30 (DO de 0,5 à 600 nm), la production de la bactérie est induite par ajout d'IPTG (0,5 mM final) et la température est diminuée à 30°C . La croissance bactérienne et la production de la protéine recombinante sont stoppées par centrifugation du milieu après 8 H. Le culot cellulaire obtenu est ensuite lavé par une solution tampon contenant du TRIS 50 mM et du NaCl 50 mM à pH 7,5. Les cellules sont conservées au congélateur à -20°C jusqu'à utilisation.

2.2. Purification de la FSA par résine IMAC

Les souches contenant les différentes FSA sont tout d'abord cultivées suivant le protocole ci-dessus. Puis la biomasse obtenue est lysée sous ultrasons. Ces enzymes étant exprimées avec une étiquette 6-His à leur extrémité N-terminale, la chromatographie d'affinité en présence de cations métalliques (IMAC) a été utilisée, méthode usuelle pour la purification des enzymes. L'extrait cellulaire (lysat) est déposé sur une colonne (Figure 81 a) contenant une résine chargée en Ni^{2+} , chélatés par l'acide nitriloacétique (résine Ni-NTA). 14 mL de résine sont utilisés pour 5 g de culot cellulaire. L'enzyme se fixe à la résine par liaisons de coordination entre l'étiquette 6-His et les cations Ni^{2+} (Figure 81 b), tandis que les autres protéines ne sont pas retenues. L'enzyme est ensuite éluée grâce à l'ajout d'une solution concentrée d'imidazole qui permet de décrocher la protéine par compétition sur le site de fixation. Enfin l'imidazole est extrait par dialyse et l'enzyme peut ensuite être utilisée lyophilisée.

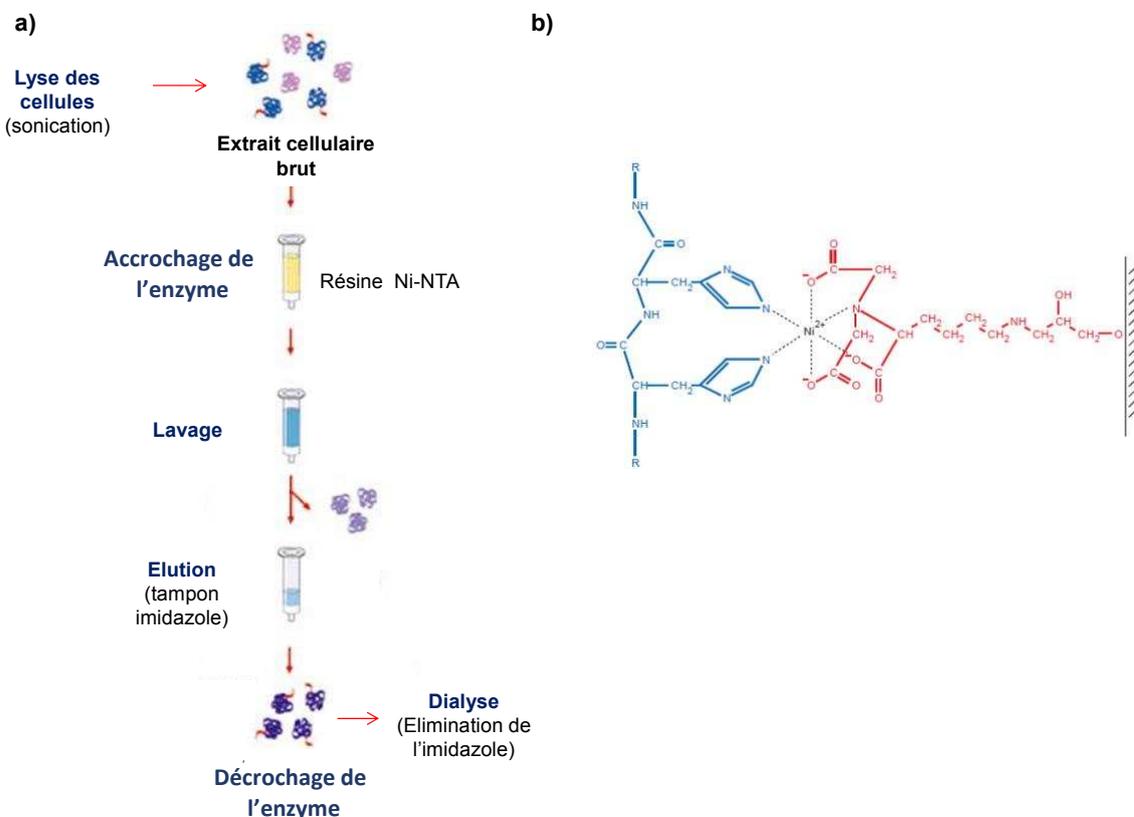


Figure 81 : purification d'enzymes sur résine IMAC, a : étapes de purification, b : interaction entre l'étiquette 6-His et la résine Ni-NTA

2.3. Production et purification de la FSA par choc thermique

Le protocole décrit par la suite doit se dérouler en milieu stérile. A 10 mL de milieu LB/Ampicilline ($0,1 \text{ mg.mL}^{-1}$) est ajouté 1 μL de cryotube contenant la FSA wt ce qui va constituer la préculture. Le mélange est ensuite incubé à 30°C sous une agitation de 150 rpm pendant une nuit. Le lendemain, 10 mL de préculture sont transférés dans 1 L de milieu LB/Ampicilline ($0,1 \text{ mg.mL}^{-1}$). Le milieu de culture est ensuite placé à 30°C sous une agitation de 250 rpm. Au bout de 2 h 30 min, la DO à 600 nm de la culture atteint 0,5-0,6. L'IPTG (0,5 mM final) est ajouté afin d'induire la production de l'enzyme. Après 16 h, le milieu est centrifugé à 10000 rpm pendant 30 min à 4°C . Le culot est lavé 2 fois avec du tampon Tris pH 7,5 (Tris 50 mM, NaCl 50 mM) et le culot est conservé par la suite au congélateur.

Le culot est resuspendu dans du tampon Glygly 50 mM pH 8 (environ 10 mL par gramme de cellule). Les cellules sont ensuite placées dans un bain de glace et lysées par sonication grâce à 2 cycles de 5 min à 50 %. Le milieu est ensuite centrifugé à 15000 rpm pendant 20 min pour faire culoter les débris cellulaires. Le surnageant est ensuite placé dans un bain marie à 70°C pendant 45 min pour faire précipiter les enzymes non thermostables. Le milieu est centrifugé à 15000 rpm pendant 20 min une dernière fois et l'activité de l'enzyme est dosée suivant le protocole décrit page 137. Un dosage de la concentration en protéines est ensuite réalisé par la méthode de Bradford.

2.4. Dosage des protéines par le test de Bradford

La quantité de protéines totales peut être déterminée par un test de Bradford. Le bleu de Coomassie G250 (coloration brune) qui est présent dans la solution de Bradford se colore en bleu en raison de sa complexation avec les résidus aromatiques des protéines. La BSA (sérum albumine bovine) est utilisée comme protéine de référence. Une droite de calibration est réalisée pour différentes concentrations de BSA dans le même tampon que la protéine d'intérêt ($0,1$ à 1 mg.mL^{-1}) en mélangeant 980 μL de solution de Bradford et 20 μL de chacune des solutions. Une cuve « blanc » est réalisée avec 20 μL de tampon. Les cuves sont incubées à l'abri de la lumière pendant 10 min et la DO est mesurée à 595 nm.

2.5. Protocole standard SDS-PAGE pour la vérification de la production des protéines recombinantes

L'électrophorèse en condition dénaturante a pour but de séparer des protéines chargées grâce à un champ électrique. Les protéines sont dans un premier temps dénaturées pour casser les ponts disulfures afin de leur faire perdre leur structure tridimensionnelle.

Préparation du gel de concentration (4 %) :

Solution	Volume en mL
Acrylamide-bisacrylamide (30 % / 0,8 %)	0,43
4X Tris-HCl 0,5 M pH 6,8	0,83
H ₂ O	2,03
Persulfate d'ammonium 10 %	0,017
N,N,N',N'-tétraméthylènediamine (TEMED)	0,0033

Préparation du gel de séparation (12,5 %) :

Solution	Volume en mL
Acrylamide-bisacrylamide (30 % / 0,8 %)	2,08
4X Tris/HCl 1,5 M pH 8,8	1,25
H ₂ O	1,75
Persulfate d'ammonium 10 %	0,017
N,N,N',N'-tétraméthylènediamine (TEMED)	0,0033

Préparation de 1 Litre de tampon de migration 10X :

Produit	Masse en g
Tris-base	30,3 g
Glycine	144 g
Dodécylsulfate de sodium (SDS)	10 g

Préparation du tampon de dénaturation Laemmli :

Produit	Concentration
Tris/HCl pH 6,8	62,5 mM
SDS	2 %
Glycérol	10 %
β -mercaptoéthanol	5 %
Bleu de bromophénol	0,02 %

Il est important pour les deux gels d'ajouter le TEMED en dernier car c'est le produit qui va démarrer la polymérisation du gel.

Le matériel est tout d'abord lavé à l'éthanol, principalement les plaques en verre qui contiendront les gels. Le montage est effectué et le gel de séparation est coulé sans emprisonner de bulles d'air jusqu'à environ 3 cm du bord. De l'eau est ajoutée pour avoir une ligne de base linéaire. La polymérisation peut être contrôlée avec le reliquat de gel non utilisé. Une fois le gel de séparation polymérisé, l'eau est retirée et le gel de concentration est coulé en présence du peigne en remplissant au maximum. Après polymérisation le gel est prêt pour le dépôt des échantillons et placé dans la cuve d'électrophorèse rempli de tampon de migration.

Les échantillons sont par la suite dilués dans le tampon de dénaturation (V/V) et incubés à 95°C pendant 10 min. Le dépôt des échantillons dénaturés est effectué et le générateur démarré à 100 mA pendant 15 min pour que les protéines pénètrent bien dans le gel de concentration puis 200 mA jusqu'à ce que le front de migration atteigne la fin du gel. Après la migration, le gel est démoulé. Il est placé dans une cuve contenant de l'INSTANTBLUE (colorant des protéines), les bandes sont visibles 15 min après.

2.6. Composition des différents milieux de culture

Les milieux gélosés sont préparés en ajoutant aux milieux liquides 15 g d'agar-agar. Les milieux sont stérilisés 15 min à 121°C.

(1) Milieu DSMZ 214 (65 + amidon) pour 1 L

Glucose : 4 g

Extrait de levure : 4 g

Extrait de malt : 10 g

Amidon : 20 g

pH 7,2

(2) Milieu DSMZ 514 pour 1 L

Peptone métabolisé par des bactéries : 5 g

Extrait de levure métabolisé par des bactéries : 1 g

Citrate de Fe(III) : 0,1 g

NaCl : 19,45 g

MgCl₂ (anhydre) : 5,9 g

Na₂SO₄ : 3,24 g

CaCl₂ : 1,8 g

KCl : 0,55 g

NaHCO₃ : 0,16 g

KBr : 80 mg

SrCl₂ : 34 mg

H₃BO₃ : 22 mg

Silicate de sodium : 4 mg

NaF : 2,4 mg

NH₄NO₃ : 1,6 mg

Na₂HPO₄ : 8 mg

pH 7,6

(3) Milieu DSMZ 219 pour 1 L

Extrait de levure : 2 g

Protéose de peptone numéro 3 : 2 g

Peptone de caséine métabolisé par la trypsine : 2 g

$\text{Na}_2\text{HPO}_4, 12\text{H}_2\text{O}$: 2,5 g

KH_2PO_4 : 1 g

Citrate de sodium : 1,5 g

$\text{MgSO}_4, 7\text{H}_2\text{O}$: 0,6 g

Glycérol : 50 mL

Tween : 0,5 g

pH 7

(4) Milieu 616

NaNO_3 : 1,5 g

KH_2PO_4 : 40 mg

$\text{MgSO}_4, 7\text{H}_2\text{O}$: 75 mg

$\text{CaCl}_2, 2\text{H}_2\text{O}$: 36 mg

Acide citrique : 6 mg

Citrate d'ammonium ferrique : 6 mg

EDTA : 1 mg

Na_2CO_3 : 20 mg

Trace de métaux, mixe A5 : 1 mL

pH 7,1

(5) Milieu ATCC 543 pour 1 L

Extrait de levure : 0,5 g

Succinate de sodium : 2,5 g

sulfate d'ammonium : 1,25 g

Citrate de Fe(III) : 3 mg

KH_2PO_4 : 0,6 g

MgSO₄, 7H₂O : 0,2 g

K₂HPO₄ : 0,9 g

CaCl₂, 2H₂O : 70 mg

EDTA : 2 mg

pH 7

(6) Milieu DSMZ 92 pour 1 L

TSB : 30 g

Extrait de levure : 3 g

pH 7,1

(7) Milieu DSMZ 535 pour 1 L

TSB : 30 g

pH 7,3

(8) Milieu DSMZ 531 pour 1 L

Polypeptone : 10 g

Extrait de levure : 2 g

MgSO₄, 7H₂O : 1 g

pH 7

(9) Milieu DSMZ 1 pour 1 L

Peptone : 5 g

Extrait de viande : 3 g

pH 7

(10) Milieu DSMZ 53 pour 1 L

Extrait de levure : 5 g

Glucose : 5 g

NaCl : 5 g

pH 7,3

(11) Milieu DSMZ 3 pour 1 L

Mélange de nutriments BD cat 234000 : 8 g

Extrait de bœuf : 3 g

Peptone : 5 g

pH 6,8

(12) Milieu DSMZ 220 pour 1 L

Peptone de caséine : 15 g

Peptone de farine de soja : 5 g

NaCl : 5 g

pH 7,3

(13) Milieu DSMZ 98 pour 1 L

Extrait de levure 1 g

Mannitol : 10 g

Extrait de sol : 200 mL

pH 7,2

(14) Milieu DSMZ 830 pour 1 L

Extrait de levure : 0,5 g

Protéose de peptone (Difco 3) : 0,5 g

Acides aminés de caséine : 0,5 g

Glucose : 0,5 g

Amidon soluble : 0,5 g

Pyruvate de sodium : 0,3 g

K₂HPO₄ : 0,3 g

MgSO₄, 7H₂O : 50 mg

pH 7,2

(15) Milieu DSMZ 54 pour 1 L

Glucose : 20 g

Extrait de levure : 10 g

CaCO₃ : 20 g

pH 7

(16) Milieu DSMZ 215 pour 1 L

Mélange de cœur et cervelle Difco : 37 g

pH 7

(17) Milieu LB pour 1 L

Tryptone : 10 g

NaCl : 5 g

Extrait de levure : 5 g

pH 7

(18) Milieu SOC pour 1 L

Hydrolysate de caséine : 10 g

NaCl : 5 g

Extrait de levure : 5 g

pH 7,5

le milieu est ensuite autoclavé puis sont ajoutés :

MgCl₂ (1 M) : 12,5 mL

MgSO₄ (1 M) : 12,5 mL

Glucose (2 M) : 10 mL

3. Procédures standards de biologie moléculaire

3.1. Extraction de plasmides grâce au GeneJET Plasmid Miniprep Kit de chez ThermoScientific

A partir des culots congelés produits comme décrit en 2.1 (10 mL de milieu), 250 µL de solution de resuspension sont ajoutés. Le mélange est ensuite agité pour homogénéisation. Puis, 250 µL de tampon de lyse sont ajoutés et le mélange est agité rapidement jusqu'à obtention d'une solution visqueuse. 350 µL de tampon de neutralisation sont ensuite ajoutés, un précipité blanc apparaît. Le milieu est centrifugé 10 min à 15000 rpm pour faire culoter les débris cellulaires et protéiques. Le surnageant est transféré dans le Gene Jet Spin Column puis centrifugé 1 min à 15000 rpm. Le filtrat est jeté et 500 µL de tampon de lavage sont ajoutés sur la colonne et centrifugés pendant 1 min à 15000 rpm. L'étape de lavage est répétée une seconde fois. Une dernière centrifugation est effectuée pendant 1 min à 15000 rpm afin de faire passer tout le tampon de lavage dans la colonne. Le récipient ayant contenu toutes les eaux de filtration est ensuite changé par un Eppendorf neuf stérile et 50 µL de tampon d'élution sont placés sur la membrane de filtration sans la toucher. Le mélange est laissé incubé 2 min à température ambiante pour que le tampon d'élution pénètre correctement et doucement dans la membrane puis centrifugé pendant 2 min à 15000 rpm. Le filtrat récupéré contient le plasmide. Il est à noter que les données précises de la composition des solutions de resuspension, du tampon de lyse, de la solution de neutralisation et de la solution de lavage ne sont pas fournis par le ThermoScientific.

3.2. Protocole pour gel d'électrophorèse ADN

Pour un gel à 0,8 %, 0,6 g d'agarose sont dilués dans 60 mL de tampon TAE (Tris Acétate, EDTA) 0,5 X et chauffés au micro-onde pendant 3 min pour solubiliser l'agarose. Ensuite 5 µL de solution Gelred sont ajoutés et le gel est coulé. Une fois que le gel a pris, il est placé dans l'appareil d'électrophorèse contenant le tampon de migration (TAE 1X) et les dépôts sont réalisés :

- 4 µL pour le marqueur de taille
- un mélange de 1 µL de tampon de charge et 4 µL de solution de plasmide pour les échantillons.

La migration est effectuée à 100 V jusqu'à ce que le front de migration atteigne le bas du gel. Le gel est ensuite révélé sous UV.

Composition du tampon TAE 1X pour 1 L :

- TRIS base 4,84 g
- 1,142 mL d'acide acétique glacial 100 %
- 2 mL d'EDTA 0,5 M
- 600 mL d'eau sont ajoutés et le pH est ajusté à 8

Une fois les produits dissous, le volume est complété à 1 L.

3.3. Procédure générale pour la mutagenèse dirigée avec le kit Agilent Quick Change II Mutagenesis

La mutagenèse dirigée des gènes codant les protéines recombinantes est réalisée au moyen du kit quikchange II (Agilent). Deux tubes à PCR sont réalisés en parallèle, l'un contenant une réaction contrôle de la mutagenèse avec des amorces et un plasmide fourni par Agilent.

Dans le tube contrôle, sont introduits dans cet ordre :

- 5 µL de tampon de réaction 10 x
- 2 µL (10 ng) de pWhitescript 4-5-kb contrôle (le plasmide contient le gène d'expression lac Z induit par l'IPTG)
- 1,25 µL (125 ng) d'amorce de contrôle 1
- 1,25 µL (125 ng) d'amorce de contrôle 2
- 1 µL de dNT mix
- 38,5 µL d'eau distillée pour arriver à un volume final de 50 µL

Dans le tube mutagenèse, sont introduits dans cet ordre :

- 5 µL de tampon de réaction 10 x
- 1 µL (100 ng) de plasmide extrait
- 1,25 µL (125 ng) d'amorce 1 contenant la mutation créée sur le site d'Agilent
- 1,25 µL (125 ng) d'amorce 2 contenant la mutation créée sur le site d'Agilent
- 1 µL de dNT mix

- 39,5 μL d'eau distillée pour arriver à un volume final de 50 μL

Une fois les deux tubes prêts, 1 μL de PfuUltra HF DNA polymérase ($2,5 \text{ U} \cdot \mu\text{L}^{-1}$) est ajouté dans chaque tube et la PCR est démarrée :

- 1 cycle de 30 s à 95°C
- 18 cycles de 30 s à 95°C , 1 min à 55°C et 5 min à 68°C

Une fois la PCR terminée, 1 μL de Dnp I est introduit pour digérer le plasmide parent méthylé. Le milieu est agité vigoureusement 1 min et placé à 37°C pendant 1 heure. Après la digestion, les tubes sont placés dans la glace ainsi que trois Eppendorfs vides qui serviront à réaliser la transformation. A partir de cette étape tout se fait en conditions stériles. Les cellules compétentes XL1-Blue sont décongelées (les cellules sont très fragiles). 1 μL du tube contrôle est introduit dans 50 μL de cellules compétentes, idem pour le tube mutagénèse et 1 μL de plasmide pUC18 de contrôle est introduit aussi dans 50 μL de cellules compétentes afin de réaliser un contrôle de transformation (le plasmide contient le gène d'expression lac Z induit par l'IPTG). Un contact est effectué pendant 30 min dans la glace puis un choc thermique est réalisé pour transformer les cellules compétentes à 42°C pendant 45-50 s. Une fois le choc thermique effectué, les milieux réactionnels sont placés dans la glace 2min. Après ce temps de repos, 500 μL de milieu SOC préalablement incubés à 42°C sont introduits pour stimuler les cellules compétentes. Les milieux réactionnels sont ensuite placés sous agitation (250 rpm) 2 heures à 37°C . En parallèle, des boîtes de pétri sont préparées pour isoler des colonies et vérifier les contrôles :

- ✓ 2 Boîtes de culture simple :
 - 20 mL de LB
 - Ampicilline $0,1 \text{ mg} \cdot \text{mL}^{-1}$
- ✓ 2 Boîtes de contrôle :
 - 20 mL de LB
 - Ampicilline $0,1 \text{ mg} \cdot \text{mL}^{-1}$
 - Xgal $0,08 \text{ mg} \cdot \text{mL}^{-1}$
 - IPTG 20 mM

Pour le contrôle de la mutation et des cultures simples, étaler 250 μL sur chaque boîte de milieu de culture. Pour le contrôle transformation, étaler un mélange de 5 μL de milieu de

culture et 200 µL de SOC (milieu super optimal avec répression catabolique). Les boîtes sont ensuite incubées 1 à 2 jours à 37°C. Les boîtes « contrôles » doivent contenir au moins 80 % de colonies bleues pour le contrôle mutation et 98 % pour le contrôle transformation. Cette couleur bleue correspond à l'expression du gène lac Z dégradant le Xgal (couleur bleue). Une fois les colonies visibles, 20 colonies de chacune des boîtes de cultures simple sont repris dans des boîtes contenant le même milieu pour un isolement et incubée à 37°C pendant une nuit. Enfin des cultures liquides sont lancées et les plasmides sont extraits pour séquençage suivant les protocoles décrits précédemment.

Composition du milieu SOC :

- 0,5 % d'extrait de levure
- 2 % de tryptone
- 10 mM de NaCl
- 2,5 mM de KCl
- 10 mM de MgSO₄
- 10 mM MgCl₂
- 20 mM de glucose

Le pH est ajusté à 7,5.

3.4. Extraction d'ADN génomique à l'aide du kit DNA Purification de Promega

Le culot est repris dans le tampon de lyse du noyau et incubé à 80°C pendant 5 min. Le mélange est laissé à température ambiante et 3 µL de RNase sont ajoutés, s'en suit une incubation de 1 h à 37°C. 200 µL de tampon pour faire précipiter les protéines sont ajoutés. Les échantillons sont placés dans la glace 5 min puis centrifugés à 13000 rpm 3 min. Le surnageant est ensuite transféré dans 600 µL d'isopropanol (précipitation de l'ADN) puis centrifugé à 13000 rpm 3 min. Le culot d'ADN est ensuite lavé à l'éthanol et mis à sec. L'ADN est repris dans l'eau ultra pure et stocké à 4°C.

3.5. Procédure standard pour l'Amplification en Chaîne par Polymérase (PCR)

Deux Taq-polymérases ont été utilisées : la DreamTaq ainsi que la Taq-phusion de chez Thermo Fischer Scientific. Des gradients de température ont été réalisés pour les 45 gènes à cloner entre 55 et 63°C. Dans un tube PCR de volume final 50 µL sont introduits :

- 2,5 µL de chaque amorce
- 25 µL de solution Taq-phusion ou DreamTaq
- 5 µL d'ADN (150 ng)
- 15 µL H₂O milliQ

Les tubes PCR sont ensuite placés dans l'appareil PCR. Pour la DreamTaq, le programme est :

- 95°C pendant 3 min pour la première dénaturation
- 35 cycles comportant 30 s à 95°C (deuxième dénaturation), 30 s au T_m retenu et 2 min à 72°C pour l'élongation (1 min/kB).
- 72°C pendant 10 min d'élongation finale

Pour la Taq-phusion le programme PCR est :

- 98°C pendant 30 s pour la première dénaturation
- 30 cycles comportant 10 s à 98°C (deuxième dénaturation), 30 s à 65°C et 1 min à 72°C (30 s/kB)
- 72°C pendant 10 min d'élongation finale

3.6. Procédure standard pour le LIC clonage

Le LIC clonage se réalise à partir des produits PCR purifiés. Les produits PCR sont purifiés sur gel d'agarose 0,8 %. La migration est effectuée à 100 V pour une bonne séparation des différents résidus et les bandes qui correspondent aux gènes d'intérêt sont découpées et purifiées à l'aide du kit d'extraction de gel (Thermo Fischer Scientific). Pour 100 mg de gel sont ajoutés 100 µL de tampon de liaison. Le mélange est ensuite placé dans un bain marie à 50-60°C pour dissoudre le gel pendant 10 min. Une fois le gel dissous, le tampon de liaison contenant l'ADN est déposé sur les colonnes fournies dans le kit et centrifugé 1 min à 13 000 rpm. L'ADN reste accroché à la membrane et le tampon de liaison

qui lui passe dans le tube de collecte est jeté. 700 μL de tampon de lavage sont introduits et une nouvelle centrifugation est effectuée dans les mêmes conditions que la précédente. Puis, une centrifugation supplémentaire de 1 min est réalisée pour bien retirer tout le tampon de lavage. 20 μL de tampon d'éluion sont ensuite déposés au centre de la membrane et le produit PCR purifié est récupéré après 2 min de centrifugation.

Une fois le produit PCR purifié, le LIC clonage peut être réalisé. Dans un tube PCR de 10 μL de volume final sont introduits :

- 2 μL de tampon LIC
- 7 μL de produit PCR purifié
- 1 μL de T4 ADN polymérase (1 U)

Le mélange est agité 3-5 s et incubé à température ambiante pendant 5 min. La réaction est stoppée par ajout de 0,6 μL de EDTA 0,5 M. 1 μL de plasmide pLATE11 (fourni dans le kit) est introduit directement dans le résultat de la digestion précédente et le mélange est incubé 5 min à température ambiante. La réaction est complète au bout de 5 min mais le fait de laisser plus longtemps ne joue pas sur la stabilité ou l'efficacité.

Enfin, 3 μL de cette préparation sont mis en contact avec 50 μL de cellules compétentes BL21 DE3 dans un tube polypropylène de 15 mL. Le contact est effectué pendant 30 min dans la glace puis un choc thermique est réalisé pour transformer les cellules compétentes à 42°C pendant 60 s. Une fois le choc thermique effectué, les milieux réactionnels sont placés dans la glace pendant 2min. Après ce temps de repos, 950 μL de milieu SOC préalablement incubé dans la glace sont introduits pour stimuler les cellules compétentes. Les milieux réactionnels sont ensuite placés sous agitation (250 rpm) pendant 2 heures à 37°C. Une fois ce temps écoulé, 500 μL sont transférés dans 10 mL de milieu LB à 0,1 $\text{mg}\cdot\text{mL}^{-1}$ d'ampicilline et 100 μL sont étalés sur boîte de pétri contenant le même milieu. Les bactéries sont ensuite incubées toute la nuit à 37°C.

3.7. Dosage de l'ADN (plasmides inclus) par NanoDrop

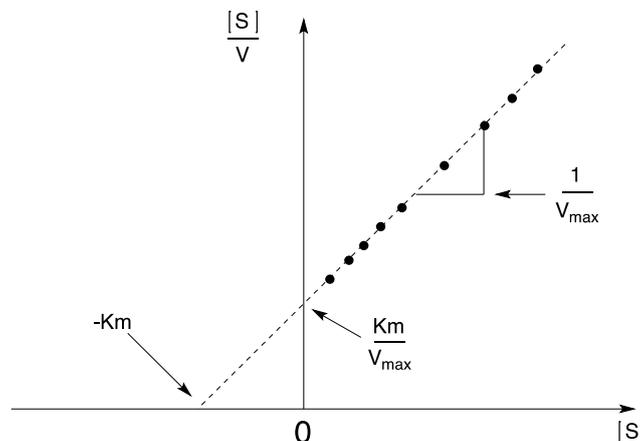
Le NanoDrop LITE spectrophotometer Thermo Scientific est ensuite utilisé pour déterminer la quantité d'ADN double brins (génomique ou plasmidique) obtenus après extraction par mesure de l'absorbance à 260 nm. Le blanc est effectué 2 fois avec 1 μL de tampon d'éluion. Les mesures sont ensuite effectuées avec 1 μL de solution de plasmides

(le NanoDrop LITE spectrophotometer est rincé entre chaque mesure par 1 μL d'eau distillée).

4. Protocoles d'enzymologie

4.1. Détermination des constantes cinétiques par la méthode de Hanes-Woolf

Les constantes cinétiques d'une enzyme pour un substrat sont déterminées en faisant varier la concentration $[S]$ de ce même substrat et en mesurant la vitesse V_i de la réaction correspondante, le second substrat (s'il y en a un) reste à une concentration fixe. La méthode de Hanes-Woolf a été utilisée pour le calcul des constantes cinétiques. Elle consiste à tracer $[S]/V$ en fonction de $[S]$. Si la cinétique est de type Michaelis-Menten, la courbe est une droite d'équation $y = ax+b$:



Les constantes sont donc égales à :

$$K_M = -b/a \text{ et } V_{max} = 1/a$$

Les constantes cinétiques de la FSA L107Y A129G ont été déterminées en faisant varier les concentrations en L-glycéraldéhyde de 1 à 50 mM en microplaques 96 puits. Le volume réactionnel est de 200 μL . Le tampon utilisé est du Gly-gly 50 mM, pH 8 et la quantité de FSA lyophilisée est ramenée à 0,6 mg par puits.

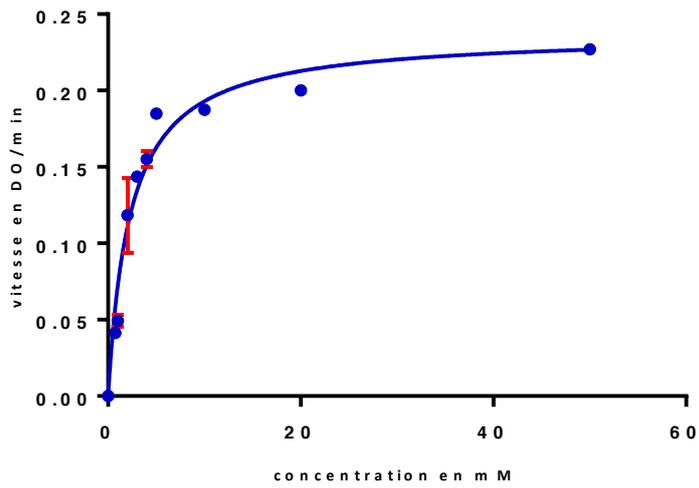


Figure 82 : représentation graphique des vitesses initiales en fonction des concentrations en substrat :
 $V=f(\text{concentration})$

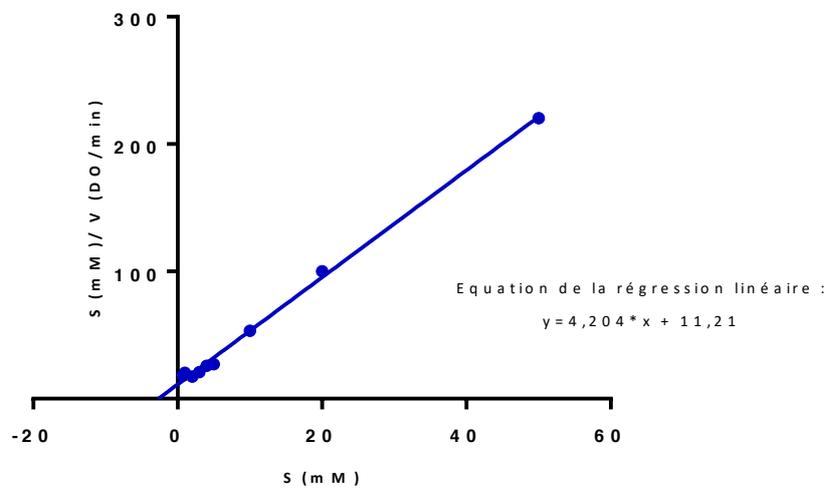


Figure 83 : représentation de Hanes-Woolf : $S/V = f(S)$

Les résultats sont fournis (Tableau 10).

4.2. Test de criblage des fructose-6-phosphate aldolases

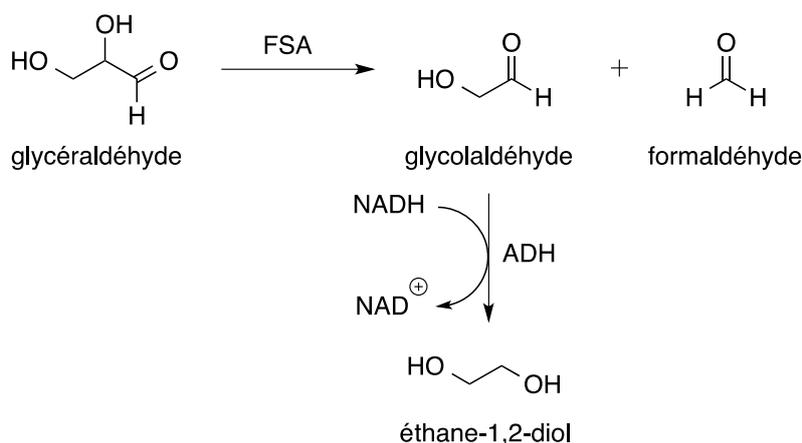


Figure 84 : méthode de criblage et de détermination des constantes cinétiques de la FSA L107Y A129G

La réaction est suivie au spectrophotomètre à 340 nm. Dans un puits de 200 μL de volume final sont introduits :

- 6 μL NADH d'une solution à 11 $\text{mg}\cdot\text{mL}^{-1}$
- 10 μL L-glycéraldéhyde d'une solution à 45 $\text{mg}\cdot\text{mL}^{-1}$
- 10 μL ADH d'une solution à 500 $\text{U}\cdot\text{mL}^{-1}$
- 114 μL de tampon Glygly 50 mM pH 8
- 48 μL FSA (25 $\text{mg}\cdot\text{mL}^{-1}$)

4.3. Synthèse et caractérisation RMN du L-xylose

La réaction est effectuée dans 1 mL de tampon Glygly 50 mM pH 8. 10 mg de FSA L107Y A129G sont introduits dans le milieu réactionnel contenant 4,5 mg de L-glycéraldéhyde (50 mM). Au bout de 24 h, 400 μL du milieu réactionnel sont prélevés et mélangés à 200 μL de solvant RMN D_2O . Le spectre RMN est identique à la littérature^[55] :

RMN ^1H (400 MHz, D_2O) $\delta = 5,17$ ppm (d, $J = 3,6$ Hz, 1H, H-1- β), $\delta = 4,55$ ppm (d, $J = 7,9$ Hz, 1H, H-1- α), $\delta = 3,90$ ppm (dd, $J = 11,5/5,5$ Hz, 1H, H-5- α), $\delta = 3,65$ ppm (s, 2H, H-5/5'- β), $\delta = 3,60$ ppm (m, 1H, H-3- β), $\delta = 3,58$ ppm (m, 1H, H-4- α), $\delta = 3,56$ ppm (m, 1H, H-4- β), $\delta = 3,49$ ppm (dd, $J = 9/3,6$ Hz, 1H, H-2- β), $\delta = 3,4$ ppm (t, $J = 9,2$ Hz, 1H, H-3- α), $\delta = 3,3$ ppm (t, $J = 11$ Hz, 1H, H-5'- α), $\delta = 3,19$ ppm (dd, $J = 9,4/7$ Hz, 1H, H-2- α).

4.4. Criblage des oxydases

Le criblage des oxydases a été réalisé par dosage de l' H_2O_2 formé par la réaction d'oxydation (Figure 85)

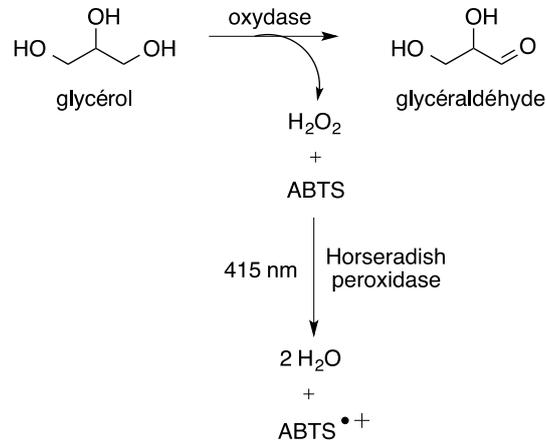


Figure 85 : criblage des oxydases

La réaction est suivie au spectrophotomètre à 415 nm. Dans une cuve de 1,5 mL de volume final sont introduits :

- 1 mg de protéines issues des différents lysats (volume variable)
- 150 μL de ABTS d'une solution à 100 mM
- 10 μL d'une solution d'HRP (2 U)
- 100 μL d'une solution de glycérol à 800 mM
- Le volume est complété par un tampon phosphate K_2HPO_4 50mM pH 8,7

Annexes

Annexe 1

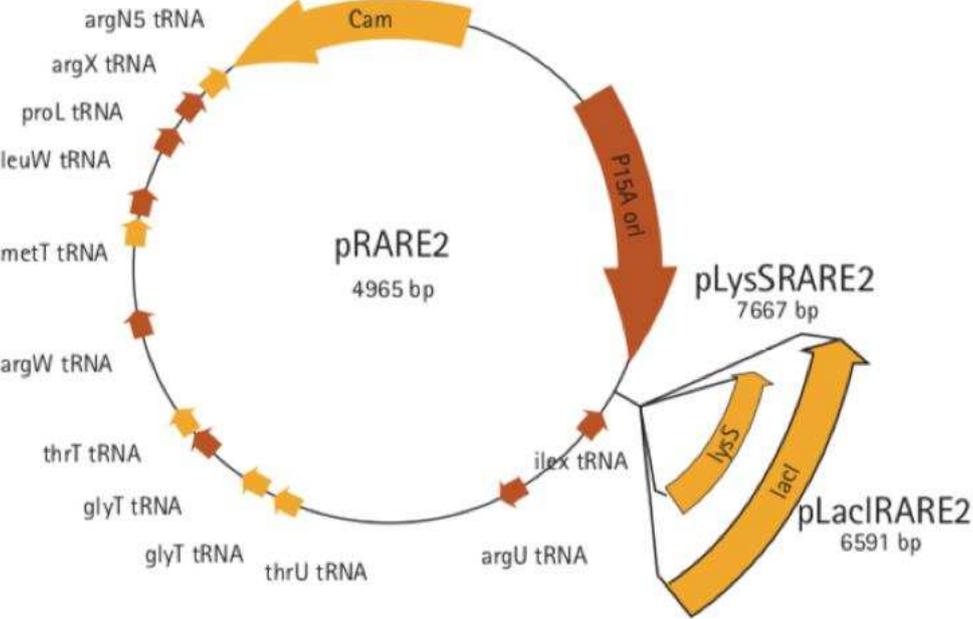


Figure 86 : carte du plasmide contenant les ARNt incluse dans les cellules compétentes Rosetta 2 (DE3)

Annexe 2

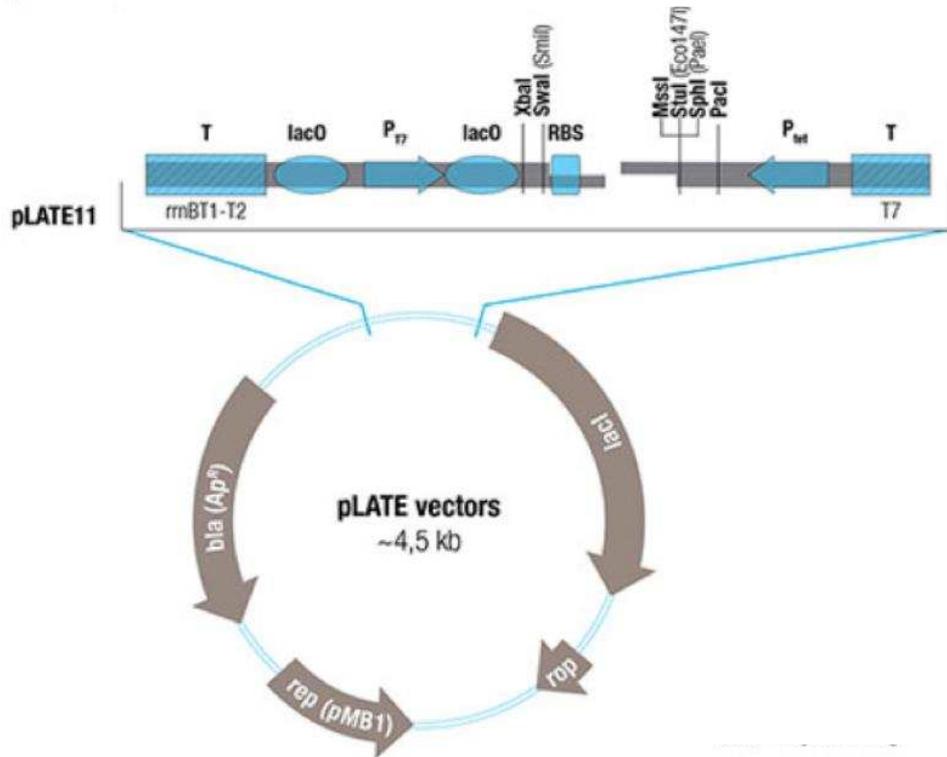


Figure 87 : carte du plasmide pLATE11

P_{T7} : promoteur de l'ARN T7 polymérase

T rrnBT1-T2 : terminateur de transcription

LacO : opérateur lac

RBS : site de fixation du ribosome

P_{tet} : promoteur P_{tet}

T T7 : terminateur T7

lacI : répresseur lac

rop : protéine rop régulatrice du nombre de copies

rep (pMB1) : origine de réplication du plasmide pMB1

bla (Ap^R) : gène β -lactamase

Références

Bibliographique

- [1] M. Pagliaro, R. Ciriminna, H. Kimura, M. Rossi, C. Della Pina, *Angew. Chem. Int. Ed.* **2007**, *46*, 4434–4440.
- [2] V. V. Ordonsky, A. Y. Khodakov, A. T. Nijhuis, J. C. Schouten, *Green Process. Synth.* **2015**, *4*, DOI 10.1515/gps-2015-0037.
- [3] J. Vandeputte, *Ol. Corps Gras Lipides* **20121-2**, 16–21.
- [4] G. Bagnato, A. Iulianelli, A. Sanna, A. Basile, *Membranes* **2017**, *7*, 17.
- [5] B. Katryniok, S. Paul, V. Bellière-Baca, P. Rey, F. Dumeignil, *Green Chem.* **2010**, *12*, 2079.
- [6] P. T. Anastas, J. C. Warner, *Green Chemistry: Theory and Practice*, Oxford Univ. Press, Oxford, **2000**.
- [7] P. Clapés, in *Org. Synth. Using Biocatal.*, Elsevier, **2016**, pp. 285–337.
- [8] J. Abdoul-Zabar, I. Sorel, V. Hélaine, F. Charmantray, T. Devamani, D. Yi, V. de Berardinis, D. Louis, P. Marlière, W.-D. Fessner, et al., *Adv. Synth. Catal.* **2013**, *355*, 116–128.
- [9] M. P. Kamble, G. D. Yadav, *Catal. Today* **2018**, *309*, 236–241.
- [10] P. Clapés, X. Garrabou, *Adv. Synth. Catal.* **2011**, *353*, 2263–2283.
- [11] M. Brovetto, D. Gamenara, P. Saenz Méndez, G. A. Seoane, *Chem. Rev.* **2011**, *111*, 4346–4403.
- [12] M. Vashishtha, M. Mishra, D. O. Shah, *Appl. Catal. Gen.* **2013**, *466*, 38–44.
- [13] D. A. Evans, J. Bartroli, T. L. Shih, *J. Am. Chem. Soc.* **1981**, *103*, 2127–2129.
- [14] A. M. Pieczonka, S. Leśniak, M. Rachwalski, *Tetrahedron Lett.* **2014**, *55*, 2373–2375.
- [15] S. M. Dean, W. A. Greenberg, C.-H. Wong, *Adv. Synth. Catal.* **2007**, *349*, 1308–1320.
- [16] B. F. Tack, P. Chapman, S. Dagley, *J. Biol. Chem.* **1972**, *247*, 6444–6449.
- [17] K. Hernandez, J. Bujons, J. Joglar, S. J. Charnock, P. Domínguez de María, W. D. Fessner, P. Clapés, *ACS Catal.* **2017**, *7*, 1707–1711.
- [18] H. Ling, G. Wang, Y. Tian, G. Liu, H. Tan, *Biochem. Biophys. Res. Commun.* **2007**, *361*, 196–201.
- [19] W. Wang, S. Y. K. Seah, *Biochemistry* **2005**, *44*, 9447–9455.

- [20] M. J. Walters, E. J. Toone, *Nat. Protoc.* **2007**, *2*, 1825–1830.
- [21] E. Racker, *J. Biol. Chem.* **1952**, *1*, 347.
- [22] P. A. Hoffee, *Arch. Biochem. Biophys.* **1968**, *126*, 795–802.
- [23] H. J. M. Gijsen, C.-H. Wong, *J. Am. Chem. Soc.* **1995**, *117*, 2947–2948.
- [24] H. J. M. Gijsen, C.-H. Wong, *J. Am. Chem. Soc.* **1995**, *117*, 7585–7591.
- [25] C.-H. Wong, E. Garcia-Junceda, L. Chen, O. Blanco, H. J. M. Gijsen, D. H. Steensma, *J. Am. Chem. Soc.* **1995**, *117*, 3333–3339.
- [26] K. Fesko, C. Reisinger, J. Steinreiber, H. Weber, M. Schürmann, H. Griengl, *J. Mol. Catal. B Enzym.* **2008**, *52–53*, 19–26.
- [27] A. Saeed, D. W. Young, *Tetrahedron* **1992**, *48*, 2507–2514.
- [28] J. Steinreiber, K. Fesko, C. Reisinger, M. Schürmann, F. van Assema, M. Wolberg, D. Mink, H. Griengl, *Tetrahedron* **2007**, *63*, 918–926.
- [29] J.-Q. Liu, T. Dairi, N. Itoh, M. Kataoka, S. Shimizu, H. Yamada, *J. Mol. Catal. B Enzym.* **2000**, *10*, 107–115.
- [30] N. Dückers, K. Baer, S. Simon, H. Gröger, W. Hummel, *Appl. Microbiol. Biotechnol.* **2010**, *88*, 409–424.
- [31] T. Kimura, V. P. Vassilev, G.-J. Shen, C.-H. Wong, *J. Am. Chem. Soc.* **1997**, *119*, 11734–11742.
- [32] J. Steinreiber, K. Fesko, C. Mayer, C. Reisinger, M. Schürmann, H. Griengl, *Tetrahedron* **2007**, *63*, 8088–8093.
- [33] W.-D. Fessner, O. Eyrisch, *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 56–58.
- [34] N. J. Turner, *Nat. Prod. Rep.* **1994**, *11*, 1.
- [35] P. R. Alefounder, S. A. Baldwin, R. N. Perham, N. J. Short, *Biochem. J.* **1989**, *257*, 529–534.
- [36] M. K. Dreyer, G. E. Schulz, *J. Mol. Biol.* **1993**, *231*, 549–553.
- [37] T. H. Chiu, D. S. Feingold, *Biochem.* **1969**, *8*, 98–108.
- [38] T. D. Machajewski, C.-H. Wong, *Angew. Chem. Int. Ed.* **2000**, *39*, 1352–1375.
- [39] T. van Herk, A. F. Hartog, H. E. Schoemaker, R. Wever, *J. Org. Chem.* **2006**, *71*,

6244–6247.

- [40] M. Schümperli, R. Pellaux, S. Panke, *Appl. Microbiol. Biotechnol.* **2007**, *75*, 33–45.
- [41] V. Laurent, E. Darii, A. Aujon, M. Debacker, J-L. Petit, V. Hélaïne, T. Liptaj, M. Breza, A. Mariage, L. Nauton, et al., *Angew. Chem. Int. Ed.* **2018**, *57*, 5467-5471.
- [42] M. Schürmann, *J. Biol. Chem.* **2001**, *276*, 11055–11061.
- [43] C. Guérard-Hélaïne, V. de Berardinis, M. Besnard-Gonnet, E. Darii, M. Debacker, A. Debard, C. Fernandes, V. Hélaïne, A. Mariage, V. Pellouin, et al., *ChemCatChem* **2015**, *7*, 1871–1879.
- [44] I. Sánchez-Moreno, L. Nauton, V. Théry, A. Pinet, J.-L. Petit, V. de Berardinis, A. K. Samland, C. Guérard-Hélaïne, M. Lemaire, *J. Mol. Catal. B Enzym.* **2012**, *84*, 9–14.
- [45] S. Thorell, M. Schürmann, G. A. Sprenger, G. Schneider, *J. Mol. Biol.* **2002**, *319*, 161–171.
- [46] V. Sautner, M. M. Friedrich, A. Lehwiss-Litzmann, K. Tittmann, *Biochemistry* **2015**, *54*, 4475–4486.
- [47] L. Stellmacher, T. Sandalova, S. Leptihn, G. Schneider, G. A. Sprenger, A. K. Samland, *ChemCatChem* **2015**, *7*, 3140–3151.
- [48] M. Schürmann, M. Schürmann, G. A. Sprenger, *J. Mol. Catal. B Enzym.* **2002**, *19–20*, 247–252.
- [49] G. A. Sprenger, U. Schörken, T. Wiegert, S. Grolle, A. A. De Graaf, S. V. Taylor, T. P. Begley, S. Bringer-Meyer, H. Sahm, *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12857–12862.
- [50] X. Garrabou, J. A. Castillo, C. Guérard-Hélaïne, T. Parella, J. Joglar, M. Lemaire, P. Clapés, *Angew. Chem. Int. Ed.* **2009**, *48*, 5521–5525.
- [51] M. Sugiyama, Z. Hong, P.-H. Liang, S. M. Dean, L. J. Whalen, W. A. Greenberg, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 14811–14817.
- [52] A. Soler, M. L. Gutiérrez, J. Bujons, T. Parella, C. Minguillon, J. Joglar, P. Clapés, *Adv. Synth. Catal.* **2015**, *357*, 1787–1807.
- [53] F. Camps Bres, C. Guérard-Hélaïne, C. Fernandes, J. A. Castillo, M. Lemaire, *Tetrahedron Asymmetry* **2013**, *24*, 1075–1081.

- [54] J. Fanton, F. Camps, J. A. Castillo, C. Guérard-Hélaine, M. Lemaire, F. Charmantray, L. Hecquet, *Eur. J. Org. Chem.* **2012**, 2012, 203–210.
- [55] J. A. Castillo, C. Guérard-Hélaine, M. Gutiérrez, X. Garrabou, M. Sancelme, M. Schürmann, T. Inoue, V. Hélaine, F. Charmantray, T. Gefflaut, et al., *Adv. Synth. Catal.* **2010**, 352, 1039–1046.
- [56] A. Szekrenyi, A. Soler, X. Garrabou, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, J. Bujons, P. Clapés, *Chem. - Eur. J.* **2014**, 20, 12572–12583.
- [57] A. Szekrenyi, X. Garrabou, T. Parella, J. Joglar, J. Bujons, P. Clapés, *Nat. Chem.* **2015**, 7, 724–729.
- [58] C. J. A. Mota, B. Peres Pinto, A. L. de Lima, *Glycerol*, Springer International Publishing, Cham, **2017**.
- [59] B. Katryniok, H. Kimura, E. Skrzyńska, J.-S. Girardon, P. Fongarland, M. Capron, R. Ducoulombier, N. Mimura, S. Paul, F. Dumeignil, *Green Chem.* **2011**, 13, 1960.
- [60] A. Beltram, I. Romero-Ocaña, J. José Delgado Jaen, T. Montini, P. Fornasiero, *Appl. Catal. Gen.* **2015**, DOI 10.1016/j.apcata.2015.09.022.
- [61] M. Mohamad, B. Ul Haq, R. Ahmed, A. Shaari, N. Ali, R. Hussain, *Mater. Sci. Semicond. Process.* **2015**, 31, 405–414.
- [62] J. Remón, J. R. Giménez, A. Valiente, L. García, J. Arauzo, *Energy Convers. Manag.* **2016**, 110, 90–112.
- [63] B. M. Bell, J. R. Briggs, R. M. Campbell, S. M. Chambers, P. D. Gaarenstroom, J. G. Hippler, B. D. Hook, K. Kearns, J. M. Kenney, W. J. Kruper, et al., *CLEAN - Soil Air Water* **2008**, 36, 657–661.
- [64] M. O. Sonnati, S. Amigoni, E. P. Taffin de Givenchy, T. Darmanin, O. Choulet, F. Guittard, *Green Chem* **2013**, 15, 283–306.
- [65] J. Esteban, E. Domínguez, M. Ladero, F. Garcia-Ochoa, *Fuel Process. Technol.* **2015**, 138, 243–251.
- [66] A. M. Truscello, C. Gambarotti, M. Lauria, S. Auricchio, G. Leonardi, S. U. Shisodia, A. Citterio, *Green Chem.* **2013**, 15, 625.
- [67] H. Zhang, M. W. Grinstaff, *Macromol. Rapid Commun.* **2014**, 35, 1906–1924.
- [68] D. Tanton, A. Le Borgne, M. Spulchre, N. Spassky, *Macromol. Chem. Phys.* **1994**,

195, 139–148.

- [70] M. A. Carnahan, M. W. Grinstaff, *J. Am. Chem. Soc.* **2001**, *123*, 2905–2906.
- [71] F. Zhang, X. Ren, H. Huang, J. Huang, M. Sudhakar, L. Liu, *Chin. J. Chem. Eng.* **2018**, *26*, 1031–1040.
- [72] A. Galadima, O. Muraza, *J. Taiwan Inst. Chem. Eng.* **2016**, *67*, 29–44.
- [73] G. Srinivasa Rao, N. Pethan Rajan, M. Hari Sekhar, S. Ammaji, K. V. R. Chary, *J. Mol. Catal. Chem.* **2014**, *395*, 486–493.
- [74] J. Barrault, F. Jerome, *Eur. J. Lipid Sci. Technol.* **2008**, *110*, 825–830.
- [75] Y. Gu, A. Azzouzi, Y. Pouilloux, F. Jérôme, J. Barrault, *Green Chem* **2008**, *10*, 164–167.
- [76] J. ten Dam, K. Djanashvili, F. Kapteijn, U. Hanefeld, *ChemCatChem* **2013**, *5*, 497–505.
- [77] M. Akiyama, S. Sato, R. Takahashi, K. Inui, M. Yokota, *Appl. Catal. Gen.* **2009**, *371*, 60–66.
- [78] R. M. Painter, D. M. Pearson, R. M. Waymouth, *Angew. Chem. Int. Ed.* **2010**, *49*, 9456–9459.
- [79] S. Chornaja, E. Sproge, K. Dubencovs, L. Kulikova, V. Serga, A. Cvetkovs, V. Kampars, *Key Eng. Mater.* **2014**, *604*, 138–141.
- [80] X. Cui, C. Zhang, F. Shi, Y. Deng, *Chem. Commun.* **2012**, *48*, 9391.
- [81] V. Calvino-Casilda, M. O. Guerrero-Pérez, M. A. Bañares, *Green Chem.* **2009**, *11*, 939.
- [82] O. Cofré, M. Ramírez, J. M. Gómez, D. Cantero, *Biomass Bioenergy* **2016**, *91*, 37–47.
- [83] H. P. Blaschek, T. C. Ezeji, J. Scheffran, Eds. , *Biofuels from Agricultural Wastes and Byproducts*, Wiley-Blackwell, Ames, Iowa, **2010**.
- [84] A. Pandey, Ed. , *Current Developments in Biotechnology and Bioengineering: Production, Isolation and Purification of Industrial Products*, Elsevier, Amsterdam, **2017**.
- [85] N. Vivek, A. Pandey, P. Binod, *Bioresour. Technol.* **2016**, *213*, 222–230.
- [86] C. Raynaud, P. Sarcabal, I. Meynial-Salles, C. Croux, P. Soucaille, *Proc. Natl. Acad. Sci.* **2003**, *100*, 5010–5015.

- [87] S. Pflügl, H. Marx, D. Mattanovich, M. Sauer, *Bioresour. Technol.* **2012**, *119*, 133–140.
- [88] B. T. Maru, F. López, S. W. M. Kengen, M. Constantí, F. Medina, *Fuel* **2016**, *186*, 375–384.
- [89] N. Stepanov, E. Efremenko, *New Biotechnol.* **2017**, *34*, 54–58.
- [90] N. Vikromvarasiri, S. Haosagul, S. Boonyawanich, N. Pisutpaisal, *Int. J. Hydrog. Energy* **2016**, *41*, 15667–15673.
- [91] J. Hao, R. Lin, Z. Zheng, H. Liu, D. Liu, *World J. Microbiol. Biotechnol.* **2008**, *24*, 1731–1740.
- [92] G. N. Jarvis, E. R. B. Moore, J. H. Thiele, *J. Appl. Microbiol.* **1997**, *83*, 166–174.
- [93] R. Datta, M. Henry, *J. Chem. Technol. Biotechnol.* **2006**, *81*, 1119–1129.
- [94] K. Madhavan Nampoothiri, N. R. Nair, R. P. John, *Bioresour. Technol.* **2010**, *101*, 8493–8501.
- [95] A. Demirci, A. L. Pometto, B. Lee, P. N. Hinz, *J. Agric. Food Chem.* **1998**, *46*, 4771–4774.
- [96] N. Murakami, M. Oba, M. Iwamoto, Y. Tashiro, T. Noguchi, K. Bonkohara, M. A. Abdel-Rahman, T. Zendo, M. Shimoda, K. Sakai, et al., *J. Biosci. Bioeng.* **2016**, *121*, 89–95.
- [97] P. F. F. Amaral, T. F. Ferreira, G. C. Fontes, M. A. Z. Coelho, *Food Bioprod. Process.* **2009**, *87*, 179–186.
- [98] A. Rywińska, P. Juszczak, M. Wojtatowicz, M. Robak, Z. Lazar, L. Tomaszewska, W. Rymowicz, *Biomass Bioenergy* **2013**, *48*, 148–166.
- [99] H. Song, S. Y. Lee, *Enzyme Microb. Technol.* **2006**, *39*, 352–361.
- [100] K.-K. Cheng, X.-B. Zhao, J. Zeng, J.-A. Zhang, *Biofuels Bioprod. Biorefining* **2012**, *6*, 302–318.
- [101] S. S. Yazdani, R. Gonzalez, *Curr. Opin. Biotechnol.* **2007**, *18*, 213–219.
- [102] Z. Liu, Y. Ge, J. Xu, C. Gao, C. Ma, P. Xu, *Bioresour. Technol.* **2016**, *216*, 856–861.
- [103] C. C. Stowers, B. M. Cox, B. A. Rodriguez, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 837–852.
- [104] J. Wallenius, N. Pahimanolis, J. Zoppe, P. Kilpeläinen, E. Master, H. Ilvesniemi, J.

- Seppälä, T. Eerikäinen, H. Ojamo, *Bioresour. Technol.* **2015**, *197*, 1–6.
- [105] A. Zhang, S.-T. Yang, *Process Biochem.* **2009**, *44*, 1346–1351.
- [106] P. K. Dikshit, V. S. Moholkar, *Bioresour. Technol.* **2016**, *216*, 948–957.
- [107] G. S. Kumar, Y. Wee, I. Lee, H. J. Sun, X. Zhao, S. Xia, S. Kim, J. Lee, P. Wang, J. Kim, *Chem. Eng. J.* **2015**, *276*, 283–288.
- [108] Y.-P. Liu, Y. Sun, C. Tan, H. Li, X.-J. Zheng, K.-Q. Jin, G. Wang, *Bioresour. Technol.* **2013**, *142*, 384–389.
- [109] L. Ma, W. Lu, Z. Xia, J. Wen, *Biochem. Eng. J.* **2010**, *49*, 61–67.
- [110] D. Hekmat, R. Bauer, J. Fricke, *Bioprocess Biosyst. Eng.* **2003**, *26*, 109–116.
- [111] R. Bauer, D. Hekmat, *Biotechnol. Prog.* **2006**, *22*, 278–284.
- [112] S. Oh, C. Park, *Enzyme Microb. Technol.* **2015**, *69*, 19–23.
- [113] D. Patil, B. Dev, A. Nag, *J. Mol. Catal. B Enzym.* **2011**, *73*, 5–8.
- [114] P. Goswami, S. S. R. Chinnadayala, M. Chakraborty, A. K. Kumar, A. Kakoti, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 4259–4275.
- [115] T. Uwajima, H. Akita, K. Ito, A. Mihara, K. Aisaka, O. Terada, *Agric. Biol. Chem.* **1980**, *44*, 399–406.
- [116] D. Linke, N. Lehnert, M. Nimtz, R. G. Berger, *Enzyme Microb. Technol.* **2014**, *61–62*, 7–12.
- [117] N. Richter, M. Neumann, A. Liese, R. Wohlgemuth, T. Eggert, W. Hummel, *ChemBioChem* **2009**, *10*, 1888–1896.
- [118] D. Franke, T. Machajewski, C.-C. Hsu, C.-H. Wong, *J. Org. Chem.* **2003**, *68*, 6828–6831.
- [119] K. Isobe, M. Kataoka, J. Ogawa, J. Hasegawa, S. Shimizu, *New Biotechnol.* **2012**, *29*, 177–182.
- [120] K. Isobe, A. Kato, Y. Sasaki, S. Suzuki, M. Kataoka, J. Ogawa, A. Iwasaki, J. Hasegawa, S. Shimizu, *J. Biosci. Bioeng.* **2007**, *104*, 124–128.
- [121] B. Carlos F., W. Yi-Fong, W. Chi-Huey, *J. Am. Chem. Soc.* **1990**, *112*, 2013–2104.
- [122] G. G. Krivov, M. V. Shapovalov, R. L. Dunbrack, *Proteins Struct. Funct. Bioinforma.*

2009, *77*, 778–795.

[123] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781–1802.

[124] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791.

[125] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. B. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662.

[126] C. Vergne-Vaxelaire, F. Bordier, A. Fossey, M. Besnard-Gonnet, A. Debard, A. Mariage, V. Pellouin, A. Perret, J.-L. Petit, M. Stam, et al., *Adv. Synth. Catal.* **2013**, *355*, 1763–1779.

[127] T. Reignier, V. de Berardinis, J.-L. Petit, A. Mariage, K. Hamzé, K. Duquesne, V. Alphand, *Chem Commun* **2014**, *50*, 7793–7796.

[128] M. Steinegger, J. Soeding, **2017**, *Nature Biotechnology*, **2017**, *35*, 1026–1028.

[129] O. C. Hansen, P. Stougaard, *J. Biol. Chem.* **1997**, *272*, 11581–11587.

[130] A. M. Ledebøer, L. Edens, J. Maat, J. W. Bos, C. T. Verrips, Z. Janowicz, M. Eckart, R. Roggenkamp, C. P. Hollenberg, *Oxf. J.* **1985**, *13*, 3063–3082.

[131] K. Sakai, N. Hamada, Y. Watanabe, *Agric. Biol. Chem.* **1985**, *49*, 817–825.

[132] M. Kjellander, K. Götz, J. Liljeruhm, M. Boman, G. Johansson, *Biotechnol. Lett.* **2013**, *35*, 585–590.

[133] S. Upadhyay, G. Torres, X. Lin, *Eukaryot. Cell* **2013**, *12*, 1641–1652.

[134] H.-S. Ko, Y. Yokoyama, N. Ohno, M. Okadome, S. Amachi, H. Shinoyama, T. Fujii, *J. Biosci. Bioeng.* **2005**, *99*, 348–353.

[135] D. P. H. M. Heuts, E. W. van Hellemond, D. B. Janssen, M. W. Fraaije, *J. Biol. Chem.* **2007**, *282*, 20283–20291.

[136] R. T. Winter, D. P. H. M. Heuts, E. M. A. Rijpkema, E. van Bloois, H. J. Wijma, M. W. Fraaije, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 389–403.

[137] T. Uwajima, Y. Shimizu, O. Terada, *J. Biol. Chem.* **1984**, *259*, 2748–2753.

[138] A. M. Klibanov, B. N. Alberti, M. A. Marletta, *Biochem. Biophys. Res. Commun.* **1982**, *108*, 804–808.

- [139] C. H. W. M. R. C. Bhagya, W. Wijesundera Sulochana, Np. Hemamali, *Indian J. Hum. Genet.* **2013**, *19*, 78.
- [140] E. T. Kool, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1–22.
- [141] Z. Zhang, X. Yang, L. Meng, F. Liu, C. Shen, W. Yang, *BioTechniques* **2009**, *47*, 775–779.
- [142] T. G. Mamedov, E. Pienaar, S. E. Whitney, J. R. TerMaat, G. Carvill, R. Goliath, A. Subramanian, H. J. Viljoen, *Comput. Biol. Chem.* **2008**, *32*, 452–457.
- [143] M. Z. Li, S. J. Elledge, *Nat. Methods* **2007**, *4*, 251–256.
- [144] J. H. Lee, P. M. Skowron, S. M. Rutkowska, S. S. Hong, S. C. Kim, *Genet. Anal. Biomol. Eng.* **1996**, *13*, 139–145.
- [145] S. A. Doyle, Ed. , *High Throughput Protein Expression and Purification*, Humana Press, Totowa, NJ, **2009**.
- [146] J. F. Kane, *Curr. Opin. Biotechnol.* **1995**, *6*, 494–500.
- [147] E. Daniel, G. U. Onwukwe, R. K. Wierenga, S. E. Quaggin, S. J. Vainio, M. Krause, *BMC Bioinformatics* **2015**, *16*, DOI 10.1186/s12859-015-0743-5.
- [148] R. Novy, D. Drott, K. Yaeger, R. Mierendorf, *News. Novagen Adv. Prod. Protoc. Mol. Biol. Res.* **2001**, *12*.

Depuis les années 90, la production mondiale de biodiesel ne cesse d'augmenter passant de 15000 à plus de 430000 barils par jour. Lors de la synthèse des biodiesels, « un déchet » est produit à hauteur de 10 %, le glycérol, ce qui représente une manne financière importante. Ce travail de thèse, financé par la SAS PIVERT, a eu pour objectif de valoriser ce glycérol en glyoxal, une molécule plate-forme, en utilisant la catalyse enzymatique. Trois réactions ont été étudiées : une première concernant l'oxydation du glycérol en glycéraldéhyde, une seconde réaction de clivage du glycéraldéhyde en glycolaldéhyde et une troisième réaction d'oxydation du glycolaldéhyde en glyoxal. Les enzymes qui catalysent ces réactions ne sont pas commerciales, nous avons donc développé deux approches pour rechercher ces enzymes. Concernant les oxydases, un programme bio-informatique a été développé pour rechercher une activité enzymatique cible dans la biodiversité. En ce qui concerne la réaction de clivage du glycéraldéhyde, nous avons opté pour une approche de mutagenèse dirigée à partir de la fructose-6-phosphate aldolase basée sur des études de docking. Cette enzyme a été choisie grâce à une étude bibliographique. En effet, elle possédait de nombreux critères indispensables pour notre étude. Ce travail a permis, au sein de notre laboratoire, de développer de nouvelles approches de recherche d'activités enzymatiques dans la biodiversité et une automatisation complète de tout le processus de bio-informatique et de biologie moléculaire.

Since 90s, biodiesel production increase from 15000 to 430000 barrels per day. However, during biodiesel synthesis, one molecule is produced at 10 %, the glycerol which represent an important financeable windfall. This project, funded by SAS PIVERT, was devoted to valorize glycerol into glyoxal, a building block, by enzymatic cascade. Three enzymatic reactions were studied, first glycerol oxidation into glyceraldehyde, second glyceraldehyde cleavage into glycolaldehyde and to finish a last oxidation reaction of glycolaldehyde to give the glyoxal. We haven't found commercial enzymes able to catalyze these reactions. To find it, we have developed two different approaches. A bio-informatic program was created to search oxidase activity in biodiversity. A site directed mutagenesis method was developed from fructose-6-phosphate aldolase based on docking studies. This enzyme have been selected as started point after bibliographic research due to its characteristics. This PhD work has allowed to develop, in our laboratory, new approaches to research enzymes activity in biodiversity and a complete automation of the bio-informatic and molecular biology processes.